

AUTOMATIC METHOD OF MACULAR DISEASES DETECTION USING DEEP CNN-GRU NETWORK IN OCT IMAGES

Paweł POWROZNIK*^{ORCID}, Maria SKUBLEWSKA-PASZKOWSKA*^{ORCID}, Robert REJDAK**^{ORCID}, Katarzyna NOWOMIEJSKA**^{ORCID}

*Faculty of Electrical Engineering and Computer Science, Department of Computer Science, Lublin University of Technology,
Nadbystrzycka 38D, 20-618 Lublin, Poland

**Faculty of Medicine, Chair and Department of General and Pediatric Ophthalmology, Medical University of Lublin,
Chmielna 1, 20-079, Lublin, Poland

p.powroznik@pollub.pl, maria.paszowska@pollub.pl, robertreidak@yahoo.com, katarzyna.nowomiejska@umlub.pl

received 24 July 2023, revised 2 April 2024, accepted 7 April 2024

Abstract: The increasing development of Deep Learning mechanism allowed ones to create semi-fully or fully automated diagnosis software solutions for medical imaging diagnosis. The convolutional neural networks are widely applied for central retinal diseases classification based on OCT images. The main aim of this study is to propose a new network, Deep CNN-GRU for classification of early-stage and end-stages macular diseases as age-related macular degeneration and diabetic macular edema (DME). Three types of disorders have been taken into consideration: drusen, choroidal neovascularization (CNV), DME, alongside with normal cases. The created automatic tool was verified on the well-known Labelled Optical Coherence Tomography (OCT) dataset. For the classifier evaluation the following measures were calculated: accuracy, precision, recall, and F1 score. Based on these values, it can be stated that the use of a GRU layer directly connected to a convolutional network plays a pivotal role in improving previously achieved results. Additionally, the proposed tool was compared with the state-of-the-art of deep learning studies performed on the Labelled OCT dataset. The Deep CNN-GRU network achieved high performance, reaching up to 98.90% accuracy. The obtained results of classification performance place the tool as one of the top solutions for diagnosing retinal diseases, both early and late stage.

Key words: Drusen, Deep CNN-GRU, AMD classification, OCT, deep learning

1. INTRODUCTION

Age-related macular degeneration (AMD) is a major cause of visual impairment in elderly population of well-developed countries [26]. AMD involves dysfunction of choriocapillaris and retinal pigment epithelium (RPE) [13]. Drusen are hallmarks of early and intermediate AMD thus they are key to the diagnosis of AMD [7]. Advanced (late) stages of AMD may lead to severe vision loss due to choroidal neovascularization (CNV) – abnormal growth of vessels from the choroidal vasculature to the neurosensory retina through the Bruch's membrane. The abundance of drusen is the largest and best documented intraocular risk factor for AMD progression [29]. Drusen are extracellular deposits between the RPE basal lamina and the inner collagenous layer of Bruch's membrane are dome-shaped, lipid-rich, and often continuous with a thin layer of the same material (basal linear deposit) [6, 37]. Diabetic macular edema (DME) is defined as retinal thickening caused by the accumulation of intraretinal fluid, primarily in the inner and outer plexiform layers of the retina and may be present at any stage of diabetic retinopathy. DME remains the most common cause of vision loss among diabetic patients [39].

In clinical practice, the detection of macular disease, both AMD and DME, is typically performed by optical coherence tomography (OCT), providing cross-sectional images of the retina. OCT is a non-invasive, imaging technology used to visualize the cross-sectional retinal structure [45].

Recent advances in multimodal imaging, as OCT, have al-

lowed one to improve our ability to characterize the AMD phenotype. To help improve our understanding of drusen and their associations, large datasets are essential. Spectral-domain optical coherence tomography (SD-OCT) has been shown to have a much higher sensitivity and specificity detecting subretinal drusenoid deposits (reticular pseudodrusen) compared with the blue channel of color fundus photographs (CFPs), infrared reflectance, fundus autofluorescence, near-infrared fundus autofluorescence, confocal blue reflectance, and indocyanine green angiography [48]. Developing new methods for detection of drusen may inspire new approaches for the clinical practice.

The increasing development of Deep Learning (DL) mechanism allowed ones to create semi-fully or fully automated diagnosis software solutions for medical imaging diagnosis. Machine learning (ML) algorithms have been shown to be powerful tools in the automatic quantification of retinal biomarkers identified in OCT [24] making them ideal for the detection of drusen. Automated algorithms for drusen volume quantification are available including a software for the high definition - OCT Cirrus.

The motivation for this study lies in its potential to revolutionize ophthalmic care. Early and accurate diagnosis is the cornerstone of effective treatment, yet rare eye diseases often present diagnostic challenges due to their low prevalence and the complexity of their symptoms. This can lead to misdiagnosis, delayed treatment or irreversible damage to vision. By developing robust algorithms capable of discerning subtle patterns and anomalies in diagnostic images, clinicians may be equipped with powerful tools

to recognize these conditions early and accurately.

Moreover, the integration of artificial intelligence in ophthalmology can assure access to expert-level screening, particularly in under-resourced regions where specialist knowledge is scarce. With the ability to process vast datasets and learn from each new case, these systems can continuously improve, becoming more precise and reliable over time.

Undertaking this study also offers a unique opportunity to contribute to the broader field of medical AI. Rare diseases, with their unique manifestations, provide a rich and challenging dataset for developing advanced machine learning techniques. Success in this area not only benefits patients with rare eye conditions but also enhances the AI methodologies that can be applied to a myriad of other medical applications.

The main aim of this study is to propose a new network, Deep CNN-GRU for classification of early-stage and end-stages AMD diseases. Three types of disorders have been taken into consideration: drusen, CNV, DME, alongside with normal cases. The created automatic tool was verified on the well-known Labelled Optical Coherence Tomography dataset.

The rest of the paper is organised as follows: Section 2 describes the related works about deep learning approaches for retinal disorders classifications, Section 3 presents the method of conducting experiments, Section 4 shows the obtained results for classification, and finally Section 5 concludes the study and gives future research directions.

2. RELATED WORKS

In recent years, an increasing trend of investigating new automatic methods involving artificial intelligence (AI) may be observed [11]. This kind of automatic software may fasten the diagnosis as well as indicate the type of the classified disease. Moreover, the progression may be indicated [2]. These methods extensively applied for OCT analysis may be classified into: single-tasks involving classification or segmentation [2, 44] and multiple-tasks that combine above-mentioned issues [11].

Drusen is stated to be a very important factor of early AMD pathology. That is why various segmentation as well as classification methods applying AI have been developed. In [2] the segmentation method was applied for indicating the outer boundary of the retinal pigment epithelium (OBRPE) and the Bruch's membrane (BM) based on OCT images (166 and 200 volumes). A multitask segmentation network was proposed to capture the area between OBRPE and BM which benefited indicating the characterisation of the drusen as well as non-pathological regions where OBRPE and BM were overlapped. In [4] another segmentation method for OCT images was proposed using a new approach, called Multi-scale Transformer Global Attention Network (MsTGANet). This method utilized encoder-decoder architecture together with multi-semantic global channel and spatial joint attention module (MsGCS) to learn the model multi-semantic global contextual information as well as multi-scale transformer non-local module (MsTNL) for capturing multi-scale non-local features. The studies were performed on 8616 retinal OCT B-scans, collected from the UCSD dataset. In [2] for the drusen segmentation a U-Net architecture with Pyramid Layer was proposed, which was applied on the feature maps before passing it between encoder and decoder part of U-Net. All tests were conducted on OCT data gathered from 38 participants. For training and testing purposes

B-scan images were applied. A Generalized Dice Coefficient as loss function were used.

There are many studies concerning classification of the retinal diseases with new or modified models of CNN. In [44] Opti-Net deep learning method was applied for indicating the AMD areas in spectral-domain SD-OCT images with great success. Two datasets were used consisting 267 AMD, 115 controls and 337 AMD and 46 control cases, respectively. In [25] the new deep neural approach, called Perturbed Composite Attention Model (PCAM), was proposed for classification of macular diseases such as AMD, DME, and CNV based on OCT images. Two attention mechanism were specified. The multilevel perturbed spatial attention (MPSA) and multidimension attention (MDA) were applied for indicating the relevant contextual information in the spatial and channel domains, respectively.

In [36] a multi-scale convolutional tool using VGG16 and feature pyramid network was developed for AMD related pathologies with great success. The study was performed on two datasets: the Noor Eye Hospital (NEH) and the University of Californian San Diego (UCSD) containing OCT B-scans. In order to enlarge the datasets the augmentation methods, including rotation, shearing, brightness change, zoom change, and horizontal flipping were applied. For the NEH dataset the drusen, CNV and normal cases were classified, while for the UCSD drusen, AMD, DME and normal ones.

In [42] healthy cases, CNV and ones with drusen were classified utilizing VGG19 deep network based on 1396 OCT images. The network was pre-trained on non-medical ImageNet dataset to medical domain using an adapted densely connected classifier. The data were collected at the Ophthalmology Department of Intercommunal Hospital Center of Créteil, France.

In [46] both early-stage (drusen) and end-stages (neovascular and geographic atrophy – GA) of AMD forms were classified alongside with normal cases. The pre-trained VGG19 model on ImageNet was proposed. In the model three dense layers were added. The experiments were performed based on OCT images collected from Northwestern Memorial Hospital.

In [30] six retinal diseases: AMD, Central Serous Retinopathy (CSR), Diabetic Retinopathy (DR), CNV, DME, and drusen as well as the normal cases were detected using the proposed an enhanced deep ensemble convolutional neural network based on OCT images. The network was created using EfficientNetV2-B0 and Xception models as well as a capsule network. The verification of the proposed solution was performed on two datasets, containing 108,312 and 572 OCT images, respectively.

In [16] a tool, Label Smoothing Generative Adversarial Network (LSGAN), was proposed for classification drusen, CNV, DME alongside with normal cases. This solution consisted of three parts. Firstly, the generator, created synthetic images imitating OCT. Second, the discriminator, was used to differentiate the real OCT image with the generated one. Final, the classifier, gave the recommendation about the retinal diseases. The classification part involved various types of models, like InceptionV3, ResNet50 and DenseNet121. The evaluation of the tool was performed on two datasets: UCSD and HUCM, consisting of 84,484 and 8,904 OVT B-scan images, respectively.

The objective [33] was to employ OCT images and deep learning techniques for the classification of dry and wet AMD. This goal was achieved by utilizing two deep convolutional neural network architectures, namely AlexNet and ResNet, pretrained by ImageNet dataset. A transfer learning for fine-tune the VGG-16 network was presented in [23] for AMD classification. In the first

step for learning purposes, the ImageNet dataset was used. Subsequently, model tuning was performed using the OCT dataset.

A great number of studies were performed based on the Labelled Optical Coherence Tomography dataset containing 84,495 OCT images [18]. In [31] classification of the retinal diseases, like: CNV, DME and drusen, alongside with normal cases was presented. The CNN with 12 layers, ReLU activation function was proposed.

In [28] a hybrid system for classification retinal diseases provided high accuracy as well as needed small amount of computing load was proposed. The image features were extracted using image preprocessing as well as pre-trained VGG16 and DenseNet121. The Firefly algorithm was applied for selecting the best features. For the classification purposes SVM, Logistic Regression (LR) and Random Forest (RF) were utilized. The experiments were performed on two datasets: Labelled Optical Coherence Tomography and Srinivasan [38], containing 723 images. Various classifications were performed: between normal cases and ones with AMD, between AMD and DME, and between CNV and drusen. The pre-trained VGG16 network was also applied in [23] for the same type of classification.

Detection of CNV, DME, drusen, and normal conditions was performed utilizing CNN model with batch normalization for creating a web application [14]. The CNN model was pre-trained based on ImageNet dataset. It consisted of the following networks: ResNet, Inception, and ResNeXt.

In [5] two CNN networks were proposed for classification of retinal diseases, such as DME, AMD, drusen, and CNV as well as normal cases. The study involved pre-trained networks: Inception V3, VGG16 and modified VGG16 by adding two convolutional layers.

In [47] the same retinal diseases with normal cases were classified utilizing a Multi-branch hybrid attention network (MHA-Net). This deep learning approach involved parallel channel attention and spatial attention mechanisms for identifying the relevant characteristic features. The results showed that the proposed attention mechanism improved the performance of classification of the retinal diseases.

In [41] a hybrid artificial intelligence system, OCTNet, was proposed for AMD classification focusing on CNV, DME, and Drusen, and normal cases. The following networks were used to build the system: Support Vector Machine with Linear kernel (LSVM), Support Vector Machine with Radial Basis Function kernel (RBF SVM), Artificial Neural Network (ANN), k-Nearest Neighbor (kNN), Random Forest (RF), Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA), and Naïve Bayes (NB).

In [12] a new approach of deep learning, called Iterative fusion CNN (IFCNN), was proposed that combine features from current and previous convolutional layers to gain high accuracy.

Deep Multi-scale Fusion Convolutional Neural Network (DMF-CNN) was proposed for encoding the retina disease characteristics that were then combined for reliable and high classification [10].

Two deep learning approaches were presented for CNV, DME, drusen and normal cases classification [19]. The first one was developed based on CNN and the second combined the following models: VGG16, VGG19, ResNet50, ResNet151, DenseNet121, as well as Inception V3.

A deep residual network, ResNet50, was applied for DME, CNV, drusen and normal cases classification [3]. A fully connected block was added to the network that both improved the accu-

racy and eliminated the overfitting issues.

In [43] the classification of CNV, DME, drusen and healthy OCT images was presented utilizing various architectures of CNN approach. The solution with the highest accuracy, the seven-layer CNN, was recommended for retinal disease classification.

There are several studies about classification using multi-tasks for retinal diseases purposes. In [11] a new approach was proposed for drusen, CNV and normal retina classification. This dual guidance network involved classification using convolutional neural network (CM-CNN) and segmentation done based on U-Net network (CAM-UNet). The OCT images were analysed. The experiments were performed using two datasets: the UCSD and the other created for the purpose of the studies. The latter one consisted of macular edema and healthy cases.

In [27] non-AMD, early AMD, and intermediate AMD classes were classified utilizing Residual-Attention-UNET model with attention mechanism for segmentation drusen and end-to-end CNNs for final classification. The 2D network was created with three AI solutions: VGG16, EfficientNetB3, and DenseNets. Dataset consisting 366 eyes of 120 subjects divided into: no-AMD (40), early AMD (40) and intermediate AMD (40). OCT scans 512×128 were used. In order to increase the images an augmentation was applied based on image dilation and erosion.

Reticular pseudodrusen (RPD) and drusen were classified utilizing a deep learning framework with 3D Inception-V [32]. It consisted of three methods: Ungradable Classification Model and Outlier Model Development for detection ungradable scans and Drusen/RPD Classification Model for drusen, RPD and healthy cases classification. The experiments were performed on the UK Biobank dataset containing 1284 participants. The OCT images were indicated as with drusen, RPD, both drusen and RPD and the control group.

The use of Fully Connected Convolutional Neural Networks for AMD segmentation and classification is discussed in detail in [40]. This solution allows to map the characteristic features into a vector, which gives the possibility of classification. The proposed tool allowed for the segmentation of the retinal region, and then the classification of age-related disorders. The used network was trained on two datasets, the THOCT dataset and the Duke dataset, containing a total of over 3000 OCT images.

The collected results of scientific research on the classification of retinal diseases using deep learning approaches clearly show that the proposed structure of the Deep CNN-GRU network in this paper has not been studied before.

3. MATERIAL AND METHODS

3.1. OCT Retina Dataset

The Labelled Optical Coherence Tomography dataset is a widely used collection in the field of medical image analysis. It was firstly introduced in 2018 in [18]. This dataset consists of 84,495 images of a diverse collection of OCT images acquired from different clinical settings and devices. The dataset is organized into three main subsets: a training, validation and testing. Each of them consists of images grouped into four categories:

- NORMAL: This subset includes OCT images of healthy retinas. These images serve as the baseline for comparison with diseased retinas. They showcase the normal anatomical structures and characteristics of the retina, allowing research-

ers to differentiate between healthy and pathological conditions.

- CNV (Choroidal Neovascularization): This subset contains OCT images of retinas affected by Choroidal Neovascularization. CNV is a condition characterised by the abnormal growth of new blood vessels beneath the retina. The OCT images in this subset highlight the presence of these abnormal blood vessels and associated retinal changes.
- DME (Diabetic Macular Edema): This subset comprises OCT images of retinas affected by DME. This disease is a complication of diabetic retinopathy and involves the accumulation of fluid in the macula, the central part of the retina. The task of this area is to ensure adequate visual acuity. The OCT images in this subset reveal the presence of macular thickening, fluid accumulation, and other characteristic features of DME.
- DRUSEN: This subset includes OCT images of retinas with Drusen deposits. Drusen are yellowish-white deposits that accumulate under the retina, commonly associated AMD. The OCT images in this subset demonstrate the presence, size, and distribution of Drusen, aiding in the diagnosis and monitoring of AMD.

Images collected in the dataset came from adult patients from 5 research centers: the Shiley Eye Institute of the University of California San Diego, the California Retinal Research Foundation, Medical Center Ophthalmology Associates, the Shanghai First People's Hospital, and Beijing Tongren Eye Center. Details regarding demographics are presented in Tab. 1.

Tab. 1. Characteristic of patients from OCT Retina dataset [18]

| Diagnosis | DME | CNV | Drusen | Normal |
|------------------|-----------|-----------|-----------|-----------|
| Mean age | 57(20-90) | 83(58-97) | 82(40-95) | 60(21-86) |
| Male | 38.3% | 54.2% | 44.4% | 59.2% |
| Female | 61.7% | 45.8% | 55.6% | 40.8% |
| Caucasian | 42.6% | 83.3% | 85.2% | 59.9% |
| Asian | 23.4% | 6.3% | 8.6% | 21.1% |
| Hispanic | 23.4% | 8.3% | 4.9% | 10.2% |
| African American | 4.3% | 2.1% | 1.2% | 1.4% |
| Mixed or other | 10.6% | 0% | 0% | 7.5% |

The example images of the above-mentioned dataset group are depicted in Fig. 1.

OCT images were preprocessed to enhance quality and normalize them for analysis, it involves a series of following steps:

- Noise reduction techniques, such as filtering algorithms, were applied to remove unwanted noise and improve image clarity.
- Contrast enhancement can help improve the visibility of important features and structures within the images. Techniques like histogram equalization or contrast stretching were used to enhance the visual quality of the images.
- Normalization involves standardizing the intensity values of the pixels in the images to a consistent scale. This step ensures that the images are comparable and have a consistent brightness level, which is important for accurate analysis and comparison.

3.2. Deep CNN-GRU

In order to ensure high quality classification of changes in OCT images, the Deep CNN-GRU classifier has been proposed. It consisted of five consecutive blocks. Each of them contained two convolution and one max-pooling layers. The data from the last block was processed by Gated Recurrent Units (GRU) elements, then flattened, processed by a fully-connected layer and subjected to classification process using the Softmax function. The structure of proposed model was presented in Fig. 2 and Tab. 2.

3.2.1. Convolutional Network

Convolutional networks have been extensively studied and have demonstrated remarkable performance in numerous domains. The pioneering work of [22] introduced the concept of CNNs and their application to handwritten digit recognition, known as the LeNet-5 architecture [22]. Since then, numerous advancements have been made, including the popularisation of deeper architectures such as AlexNet [20], VGGNet [35], and ResNet [15].

CNNs are specifically designed for processing structured grid-like data, such as images or audio. Moreover, CNNs have been highly successful in various computer vision tasks, including image classification, object detection, and segmentation.

Structurally, CNNs are composed of several layers that are designed to effectively handle grid-based information, such as images. The main layers in a typical CNN architecture include convolutional layers, activation functions, pooling and fully connected layers.

3.2.2. Convolutional Layer

The first layer is typically a convolutional layer, which is responsible for learning and extracting local spatial patterns from the input data. Each convolutional layer consists of a set of learnable filters. Feature maps are generated by sliding filters across the input data and executing element-wise multiplications and summations. The filters capture different features, such as edges, textures, or shapes, by convolving with the input data. The quantity and dimensions of filters within each layer can be adjusted based on the task's complexity or the preferred network structure [34].

3.2.3. Kernel Layer

The kernel, a small matrix comprised of real values, plays a pivotal role in image processing. It operates on an input matrix, typically an image, in a patch-wise manner, with dimensions equal to that of the kernel. By performing a dot product between the patch and kernel values, a single entry in the feature map is generated. The patch selection process moves systematically across the input image, either horizontally or vertically, based on the chosen stride. This process continues until the entire image has been processed. During training, the kernel values are adjusted iteratively, undergoing changes after each iteration. This dynamic adaptation helps the model strive for optimal accuracy and minimize optimization loss. Ultimately, this operation allows to learn

diverse features such as edges or color-related characteristics, contributing to its ability to analyse and understand images effectively.

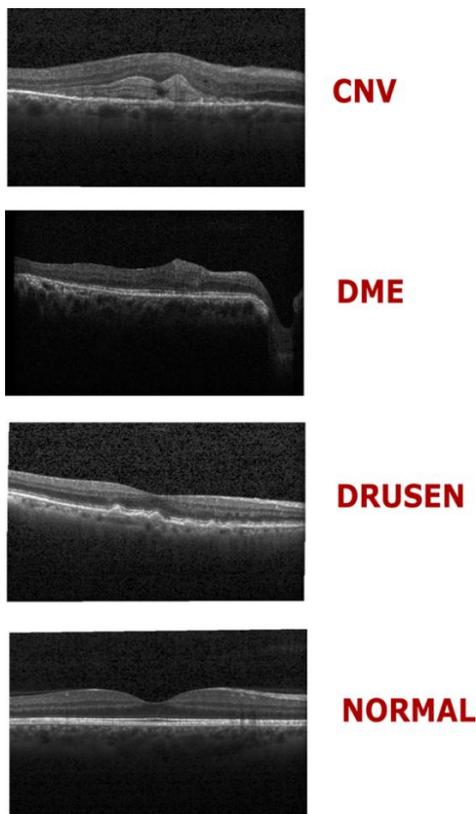


Fig. 1. Example images from OCT Retina dataset

3.2.4. Activation function

This dynamic adaptation following the convolution operation, usually, a non-linear transformation is performed on each element by applying an activation function. The Rectified Linear Unit (ReLU) is the prevailing choice for activation functions in CNNs. ReLU effectively replaces negative values with zeros while preserving positive values unaltered, thereby introducing non-linearities into the network. ReLU helps the network learn complex relationships between the input data and the desired output by introducing non-linear transformations. It was also applied in this study. Mathematically ReLU can be expressed as follow:

$$ReLU(x) = \max(0, x) \tag{1}$$

3.2.5. Pooling

Pooling operations are used to down sample the spatial dimensions of the feature maps while retaining the most important information. Commonly for pooling, within each region, are used such operators like: max pooling or average pooling, which can divide the feature maps into non-overlapping regions and select the maximum or average value, respectively. Pooling helps reduce the computational complexity and the number of parameters in the network while enhancing translation invariance and providing some degree of spatial invariance. In case of this study the max pooling actions were performed.

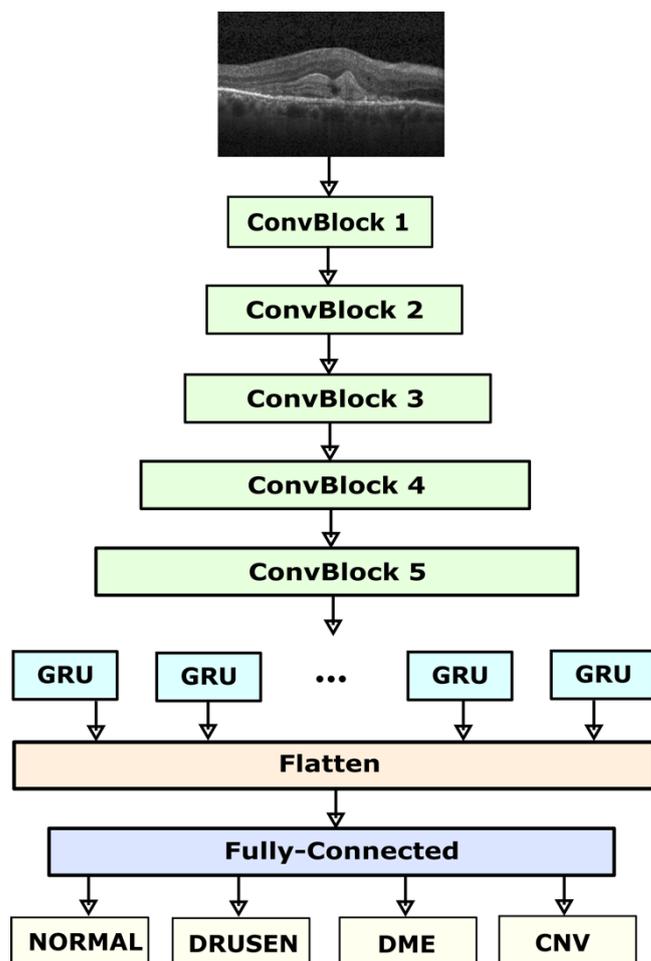


Fig. 2. Deep CNN-GRU architecture

Tab. 2. The summary of Deep CNN-GRU network

| Layer | Conv Block No | Type | Kernel size | Features No | Input size |
|-------|---------------|---------|-------------|-------------|-------------|
| 1. | 1 | Conv2D | 3x3 | 64 | 3x224x224 |
| 2. | | Conv2D | 3x3 | 64 | 64x224x224 |
| 3. | | Pooling | 2x2 | - | 64x112x112 |
| 4. | 2 | Conv2D | 3x3 | 128 | 128x112x112 |
| 5. | | Conv2D | 3x3 | 128 | 128x112x112 |
| 6. | | Pooling | 2x2 | - | 128x56x56 |
| 7. | 3 | Conv2D | 3x3 | 256 | 256x56x56 |
| 8. | | Conv2D | 3x3 | 256 | 256x56x56 |
| 9. | | Pooling | 2x2 | - | 256x28x28 |
| 10. | 4 | Conv2D | 3x3 | 512 | 512x28x28 |
| 11. | | Conv2D | 3x3 | 512 | 512x28x28 |
| 12. | | Pooling | 2x2 | - | 512x14x14 |
| 13. | 5 | Conv2D | 3x3 | 512 | 512x14x14 |
| 14. | | Conv2D | 3x3 | 512 | 512x14x14 |
| 15. | | Pooling | 2x2 | - | 512x7x7 |
| 16. | - | GRU | - | - | 512x49 |
| 17. | - | FC | - | 64 | 25088 |
| 18. | - | Output | - | 4 | 64 |

3.2.6. Fully-connected Layer

After several convolutional and pooling operations, the high-level features are usually flattened into a vector and passed through fully-connected layers (one or more). The idea of this layer is to connect each neuron with neurons in previous and subsequent layers. In this way the network is able to learn complex relationships between the extracted features and the target output. These layers are often used in the final stages of the network and can have varying sizes depending on the specific task e.g., Softmax for classification.

3.2.7. Regularisation

Deep neural networks that possess elevated learning parameters and are trained on lower-quality or noisy data may encounter challenges related to overfitting. This denotes a situation in which the model demonstrates superior performance on the training data, yet struggles to accurately classify new test instances from the same problem domain. To mitigate this problem, a dropout strategy is employed. During the training process, dropout randomly deactivates neurons in fully-connected layers with a probability. The application of dropout can be represented by the following equation [17]:

$$y_k = \sum_{K \in K^*} \text{Probability}(x) y_k^K \quad (2)$$

where:

y_k denotes predicted unit k , K^* is the set of all narrowed networks, y_k^K represents the output from unit K .

3.2.8. Loss function

The loss function quantifies the discrepancy between the predicted outputs of the network and the true labels associated with the input data. It measures an error or a distance between the predicted output and the ground truth, providing a single scalar value that indicates how well the network is performing. The objective of the CNN during the training process is to minimize this loss value.

The choice of a specific loss function depends on the nature of the problem being solved. Some commonly used loss functions in the fully-connected layer of a CNN include: mean squared error (MSE), categorical, binary or sparse categorical cross-entropy [17, 34].

During the training process, the loss function is used to calculate the gradient of the loss with respect to the model's parameters. This gradient is then utilized in the optimization algorithm (e.g., stochastic gradient descent) to update the weights and biases of the fully-connected layer, enabling the network to iteratively improve its predictions and minimize the loss. In this study categorical cross-entropy as the loss function was used.

3.2.9. Gated Recurrent Unit

The GRU model is a variant of recurrent neural network (RNN). That architecture is widely used in machine learning and natural language processing tasks. It was firstly introduced in [21] as a modification of the traditional RNN and has gained popularity

due to its effectiveness and computational efficiency.

The GRU addresses some of the limitations of the traditional RNN and the Long Short-Term Memory (LSTM) architecture. It is designed to capture and model long-term dependencies in sequential data. One of its main advantages over traditional RNNs is the mitigation of the vanishing gradient problem [9, 21].

GRU also incorporates gating mechanisms to control the flow of information within the network. However, it uses a simplified architecture with two gates: the update and the reset. Its role is to decide about the extent of information preservation from prior time steps and how much new information is to incorporate from the current time step.

The update gate in the GRU controls the information flow from the previous to current hidden state. It decides whether to update the hidden state based on two information: previous value of hidden state and current input value. By selectively updating the hidden state, the GRU can remember or forget information from previous time steps, allowing it to capture long-term dependencies [8].

The reset gate determines the amount of previous information that will be forgotten and how much will go into the current calculation. It acts as a filter, allowing the GRU to adaptively choose which past information is relevant [8].

The GRU's ability to selectively update and reset information makes it effective in wide ranges of tasks strictly connected with medical image processing.

3.2.10. Feature extraction

The feature extraction includes the processing of input data through successive layers of convolution, pooling, processing by the GRU, finally flattening and classification. To begin, each OCT image undergoes a series of convolutional layers, which transform it into multiple dimensions, represented by feature maps. The initial convolutional block takes input images of size 224×224 with three red, green, and blue channels. This block generates 64 feature maps with dimensions of 122×122 , subsequently reduced by a max-pooling layer to $112 \times 112 \times 64$.

Likewise, the second convolutional block receives 122×122 input with a dimension of 64, producing feature maps with dimensions of $11 \times 112 \times 128$. These feature maps are further reduced through a second max-pooling layer to $56 \times 56 \times 128$. Following the same pattern, the feature maps pass through the 3rd, 4th, and 5th convolutional blocks. Eventually, the final feature maps are obtained with dimensions of $7 \times 7 \times 512$. These feature maps are subsequently fed into a GRU for the purpose of classification.

4. RESULTS

4.1. Classifier evaluation

The evaluation of the proposed model incorporated the following metrics [4]: Accuracy (eq. 3), Precision (eq. 4), Recall (eq. 5), and F1 score (eq. 6). To ensure robustness, a set of experiments were conducted, involving a random data split into training, validation, and testing sets, with proportions of 60%, 20%, and 20%, respectively. To enhance result consistency, the experiments were independently repeated for 10 iterations.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (3)$$

$$Precision = TP/(TP + FP) \quad (4)$$

$$Recall = TP/(TP + FN) \quad (5)$$

$$F1 = 2 \cdot (Precision \cdot Recall)/(Precision + Recall) \quad (6)$$

In case of DRUSEN class, TP refers to accurately predicted DRUSEN cases, FP refers to cases misclassified as DRUSEN (while they are actually NORMAL or DME or CNV by the proposed system), TN represents correctly classified NORMAL or DME or CNV cases, and FN denotes DRUSEN cases misclassified as NORMAL or DME or CNV cases. The same strategy was performed for all analyzed classes.

Tab. 3. shows the accuracy results of the proposed Deep CNN-GRU model for classifying eye diseases based on OCT images. This metric reflects the model's ability to differentiate between healthy cases and those with diseases. The results obtained affirm that utilizing deep learning for eye disease recognition is highly effective. With a mean accuracy surpassing 95%, the classification approach has demonstrated remarkable success. For more detailed insights, Tab. 4. provides accuracy results for the four specified classes: healthy eyes, DME, CNV, and DRUSEN cases found in OCT images. The model achieved a minimum accuracy greater than 92% and nearly reached 99% at its peak performance.

Tab. 3. Accuracy results for Deep CNN-GRU

| Class | Mean | Max | Min | ±SD |
|-------|--------|--------|--------|-------|
| all | 95.43% | 98.90% | 92.20% | 3.86% |

Tab. 4. Accuracy results for individual classes

| Class | Mean | Max | Min | ±SD |
|--------|--------|--------|--------|-------|
| NORMAL | 95.56% | 98.90% | 92.24% | 2.18% |
| CNV | 94.05% | 98.85% | 94.00% | 2.48% |
| DME | 95.77% | 98.89% | 92.20% | 2.89% |
| DRUSEN | 95.45% | 98.14% | 92.23% | 2.73% |

Tab. 5. displays the Precision results for the developed tool. The average precision for each class surpasses 96%, ranging between 92% and 98%. Such high values indicate that the network performs numerous correct classifications and only a few misclassifications (Fig. 3.).

Tab. 5. Precision results for individual classes

| Class | Mean | Max | Min | ±SD |
|--------|--------|--------|--------|-------|
| NORMAL | 96.28% | 98.31% | 93.63% | 2.51% |
| CNV | 96.29% | 98.31% | 93.28% | 2.52% |
| DME | 96.51% | 98.37% | 94.09% | 2.14% |
| DRUSEN | 96.33% | 98.48% | 92.46% | 2.56% |

The Recall metric evaluates the model's ability to accurately classify positive instances. The results, showcased in Tab. 6., demonstrate exceptional performance for the developed tool. The average Recall surpassed 97%, with individual values falling within the range of 92.06% to 99.82%.

Tab. 6. Recall results for individual classes

| Class | Mean | Max | Min | ±SD |
|--------|--------|--------|--------|-------|
| NORMAL | 99.31% | 99.82% | 97.37% | 1.07% |
| CNV | 95.79% | 98.08% | 92.82% | 2.69% |
| DME | 95.10% | 97.75% | 91.57% | 3.01% |
| DRUSEN | 95.35% | 97.85% | 92.06% | 2.91% |

The F1 score, calculated based on Precision and Recall, yielded high results with a mean value exceeding 94%. The proposed tool demonstrated its highest accuracy in recognizing NORMAL cases. It performed slightly worse for cases with DRUSEN, CNV and DME. However, the differences in mean performance were minimal, up to 0.79%. Notably, all measures (Tab. 3-7) exhibited very low standard deviation, indicating that the observations were closely clustered around the mean, making the results highly repeatable.

Tab. 7. F1 score results for individual classes

| Class | Mean | Max | Min | ±SD |
|--------|--------|--------|--------|-------|
| NORMAL | 99.78% | 99.06% | 95.59% | 1.95% |
| CNV | 96.04% | 98.19% | 93.22% | 2.60% |
| DME | 95.41% | 98.06% | 92.81% | 2.71% |
| DRUSEN | 95.83% | 98.17% | 92.42% | 2.75% |

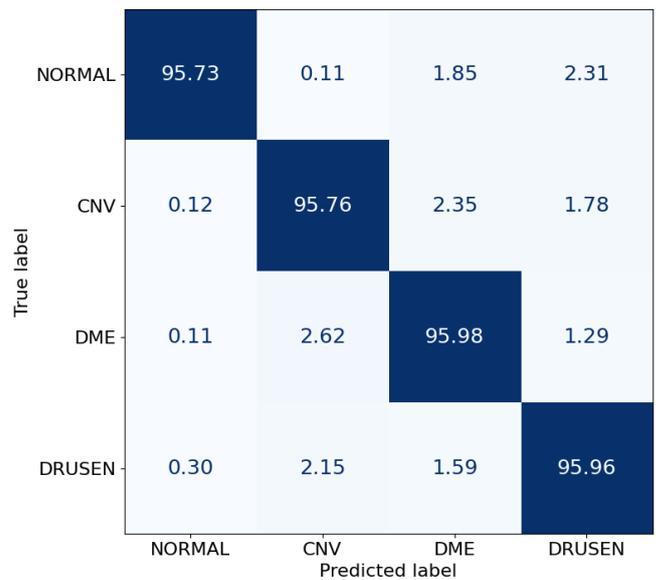


Fig. 3. Confusion matrix

Additionally, in Fig. 4. and 5., the performance evaluation of the Deep CNN-GRU classifier, presented in graphical form, can be observed. These figures depict accuracy and cross-entropy (loss) during both the training and validation phases. Reaching epoch 150, the accuracy achieved in training is 98.7%, while in validation, it is 96.4%. Similarly, the corresponding training and validation loss values for this architecture are 0.54 and 0.76, respectively.

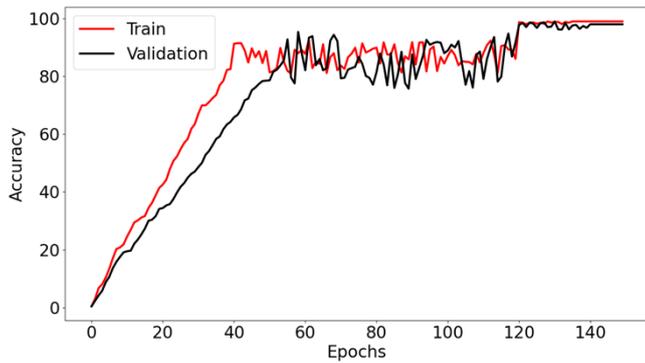


Fig. 4. Accuracy model for Deep CNN-GRU

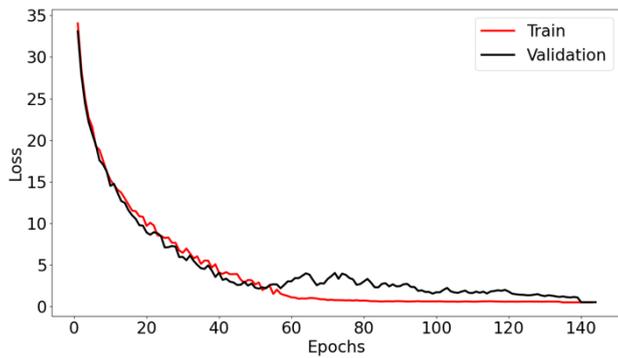


Fig. 5. Loss model for the Deep CNN-GRU

In order to guarantee the precision of the developed model Leave-One-Out Cross-Validation (LOOCV) was conducted. Despite its computational intensity, this method yields dependable and impartial insights into the model's performance. Utilizing LOOCV allowed us to calculate the root mean squared error (RMSE) – Tab. 8.

Tab. 8. Leave-One-Out Cross-Validation result

| RMSE | ±SD |
|-------|-------|
| 6.27% | 4.18% |

4.2. Comparison with the state-of-the-art

The state-of-the-art of deep learning studies performed on the Labelled Optical Coherence Tomography dataset are gathered in Tab. 9. To a large extent they concern well known convolutional approaches that were adjusted to the classification of retinal disorders with great success. The obtained accuracy is in range 85.15% and 99.69%. The proposed tool in this study obtained maximal accuracy equal to 98.90%, which is higher than for classifiers presented in [5, 10, 12, 19, 23, 31, 47]. The proposed Deep CNN-GRU achieved slightly worse results in comparison to studies described in [1] and [3]. The obtained performance for the dataset allows us to draw conclusion that the proposed tool is the proper choice for the DRUSEN, CNV, DME and NORMAL cases classification. It is worth mentioning that the Deep CNN-GRU was trained using one dataset. This approach does not need additional training on other dataset, like ImageNet, and then adjusting it to retinal diseases.

Tab. 9. State-of-the-art classification retinal diseases based on OCT Retina Dataset

| Classifiers/method | Accuracy | Study |
|----------------------------|---------------|------------|
| VGG16 | 92.19% | [47] |
| RepVGG | 94.45% | |
| ResNet50 | 95.31% | |
| Res2Net50 | 95.47% | |
| SENet | 95.24% | |
| SKNet | 95.40% | |
| MHA-Net | 96.51% | |
| CNN | 94.35% | [31] |
| OCCNet | 99.69% | [1] |
| IFCNN | 85.15% | [12] |
| DMF-CNN | 96.03% | [10] |
| VGG16, VGG19, Inception V3 | 98.30% | [19] |
| VGG16 | 98.60% | [23] |
| ResNet50 | 99.40% | [3] |
| CNN | 96.50% | [43] |
| CNN | 98.65% | [5] |
| Deep CNN-GRU | 98.90% | Own |

5. CONCLUSIONS AND FUTURE WORKS

Nowadays, the DL models have been applied for classification of various retina disorders with great success. These fully- or semi-automated tools are key elements for eye specialists for detection, accelerate treatment diagnosis as well as delay the progress of the diseases.

The proposed classifier, the Deep CNN-GRU, for recognizing rare eye diseases such as drusen, CNV, DME and detecting healthy people obtained very good results. The gained accuracy result of over 98% ensures that the developed tool is suitable for diagnosing retinal eye diseases based on OCT images. The effectiveness of the created solution using deep learning methods turned out to be more appropriate tool for classifying retinal diseases than the methods widely applied in scientific studies (Tab. 9).

Despite achieving high efficiency in detecting retinal diseases, the described method involving the use of a Deep CNN-GRU for the classification of OCT images carries several inherent limitations.

One of it can be the complexity and computation cost. The described architecture is complex with multiple convolutional layers followed by GRU units. This complexity can lead to high computational costs, which can be a limiting factor, especially when deploying in real-time clinical environments or in settings with limited computational resources.

The second restriction might be connected with dataset limitations. The performance of the CNN-GRU model is highly dependent on the quality and diversity of the dataset. If the OCT Retina Dataset lacks variability in terms of devices, patient demographics, or is not representative of the global population, the model may not generalize well to external datasets. Moreover, the quality of the labels in the training dataset is critical. Any mistakes in the annotations of the OCT images can lead to incorrect learning and thus affect the performance of the classifier.

The third limitation might be connected with interpretability and explainability. Deep learning models, including CNNs and GRUs, are often criticized for their lack of interpretability. Medical practitioners may be hesitant to trust and rely on the model's predictions without understanding the rationale behind its decisions.

The last limitation handles of sequential data. While GRUs are designed to deal with sequential data, OCT images are not inherently sequential. If the temporal dynamics of disease progression are not considered or relevant, the benefits of GRUs may not be fully realised.

AI has ability to make new insights into the vast amount of medical digital data. Ethical considerations regarding the practical use of the presented tool should cover many aspects. Firstly, those related to the risk of unauthorised access to medical data. Automated diagnosis tools require access to vast amounts of patient data, which raises concerns about the security of this information and the risk of data breaches.

Moreover, there is a need to ensure that patient information remains confidential and is not shared without consent is a critical ethical obligation. Another aspect which has to be taken under consideration is anonymization. When using patient data to train diagnostic algorithms, it's crucial to anonymize the data to protect patient identities.

Future studies can take three directions. First, it would be interesting to apply our tool to other medical images. Second, the further works may be performed on improving the performance of the proposed classifier. Third, a tool can be extended to the automatic system supporting ophthalmologists.

REFERENCES

- Alqudah A, Alqudah A. M. Artificial intelligence hybrid system for enhancing retinal diseases classification using automated deep features extracted from OCT images. *International Journal of Intelligent Systems and Applications in Engineering*. 2021; 9(3): 91-100.
- Asgari R, Orlando J I, Waldstein S, Schlanitz F, Baratsits M, Schmidt-Erfurth U, Bogunović H. Multiclass segmentation as multi-task learning for drusen segmentation in retinal optical coherence tomography. *Medical Image Computing and Computer Assisted Intervention*. 2019; 11764: 192–200.
- Asif S, Amjad K. Deep residual network for diagnosis of retinal diseases using optical coherence tomography images. *Interdisciplinary Sciences: Computational Life Sciences*. 2022; 14(4): 906-916.
- Baratloo A, Hosseini M, Negida A, El Ashal, G. Part 1: simple definition and calculation of accuracy, sensitivity and specificity. *Emergency*. 2015; 3(2): 48-49.
- Berrimi M, Moussaoui A. Deep learning for identifying and classifying retinal diseases. In *2020 2nd International Conference on computer and information sciences (ICIS)*. 1-6.
- Chen L, Messinger JD, Sloan KR, et al. Abundance and multimodal visibility of soft drusen in early age-related macular degeneration: a clinicopathologic correlation. *Retina*. 2020; 40: 1644–1648.
- Coleman HR, Chan CC, Ferris FL, Chew EY. Agerelated macular degeneration. 2008; 372: 11.
- Chollet F. *Deep Learning with Python*. Manning Publications. 2017.
- Chung J, Gulcehre C, Cho K, Bengio Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. 2014; arXiv preprint arXiv:1412.3555.
- Das V, Dandapat S, Bora P. K. Automated Classification of Retinal OCT Images Using a Deep Multi-Scale Fusion CNN. *IEEE Sensors Journal* 2021; 21:23256–23265.
- Diao S, Su J, Yang C, Zhu W, Xiang D, Chen X, Peng Q, Shi F. Classification and segmentation of OCT images for age-related macular degeneration based on dual guidance networks. *Biomedical Signal Processing and Control*. 2023; 84: 104810.
- Fang L, Jin Y, Huang L, Guo S, Zhao G, Chen X. Iterative fusion convolutional neural networks for classification of optical coherence tomography images. *Journal of Visual Communication and Image Representation*. 2019; 59: 327-333.
- Flaxman SR, Bourne RRA, Resnikoff S, et al. Global causes of blindness and distance vision impairment 1990-2020: a systematic review and meta-analysis. *Lancet Glob Health*. 2017; 5: e1221–e1234.
- Haq A, Fariza A, Ramadijanti N. Automatic Detection of Retinal Diseases in Optical Coherence Tomography Images using Convolutional Neural Network. In *2021 International Electronics Symposium (IES)*. 343-348.
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR)*. 2016; 770-778.
- He X, Fang L, Rabhani H, Chen X, Liu Z. Retinal optical coherence tomography image classification with label smoothing generative adversarial network. *Neurocomputing*. 2020; 405: 37-47.
- Islam M Z, Islam M M, Asraf A. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in medicine unlocked*. 2020; 20: 100412.
- Kermany D S, Goldbaum M, Cai W, Valentim C CS, Liang H, Baxter S L, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*. 2019; 172(5):1122–1131. <https://doi.org/10.1016/j.cell.2018.02.010>.
- Kim J, Tran L. Retinal Disease Classification from OCT Images Using Deep Learning Algorithms. In *Proceedings of the 2021 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. Melbourne Australia. 2021.
- Krizhevsky A, Sutskever I, Hinton G. E. ImageNet classification with deep convolutional neural networks. *Advances in neural information processing systems*. 2012; 25: 1097-1105.
- Kyunghyun C, et al. Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation. 2014; arXiv preprint arXiv:1406.1078.
- LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*. 1998; 86(11): 2278-2324.
- Li F, Chen H, Liu Z, Zhang X, Wu Z. Fully automated detection of retinal disorders by image-based deep learning. *Graefe's Archive for Clinical and Experimental Ophthalmology*. 2019; 257: 495–505.
- Liefers B, Taylor P, Alsaedi A, et al. Quantification of key retinal features in early and late age-related macular degeneration using deep learning. *American Journal of Ophthalmology*. 2021; 226: 1–12.
- Mishra S. S, Mandal B, Puhan N B. Perturbed composite attention model for macular optical coherence tomography image classification. *IEEE Transactions on Artificial Intelligence*. 2021; 3(4): 625-635.
- Mitchell P, Liew G, Gopinath B, Wong TY. Age-related macular degeneration. *The Lancet*. 2018; 392: 1147–1159.
- Moradi M, Chen Y, Du X, Seddon J M. Deep ensemble learning for automated non-advanced AMD classification using optimized retinal layer segmentation and SD-OCT scans. *Computers in Biology and Medicine*. 2023; 154: 106512.
- Özdaş M B, Uysal F, Hardalaç F. Classification of Retinal Diseases in Optical Coherence Tomography Images Using Artificial Intelligence and Firefly Algorithm. *Diagnostics*. 2023; 13(3): 433.
- Pollreis A, Reiter GS, Bogunovic H, Baumann L, Jakob A, Schlanitz FG, Sacu S, Owsley C, Sloan KR, Curcio CA, Schmidt-Erfurth U. Topographic Distribution and Progression of Soft Drusen Volume in Age-Related Macular Degeneration Implicate Neurobiology of Fovea. *Investigative Ophthalmology & Visual Science*. 2021; 62(2): 26-26.

30. Rahimzadeh M, Mohammadi M R. ROCT-Net: A new ensemble deep convolutional model with improved spatial resolution learning for detecting common diseases from retinal OCT images. In 2021 11th International Conference on Computer Engineering and Knowledge (ICCKE). 85-91.
31. Saraiva A A, Santos D B S, Pimentel P M C, Sousa J V M, Ferreira N M, Batista Neto J D E S, Soares S, Valente A. Classification of optical coherence tomography using convolutional neural networks. In Proceedings. 2020.
32. Schwartz R, Khalid H, Liakopoulos S, Ouyang Y, de Vente C, González-Gonzalo C, Lee A Y, Guymier R, Chew E Y, Egan C, Wu Z, Kumar H, Farrington J, Müller P L, Sánchez C I, Tufail, A. A Deep Learning Framework for the Detection and Quantification of Reticular Pseudodrusen and Drusen on Optical Coherence Tomography. *Translational Vision Science & Technology*. 2022; 11(12): 3-3.
33. Serener A, Serte S. Dry and wet age-related macular degeneration classification using oct images and deep learning. In 2019 Scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT). 2019; 1-4.
34. Shah P M, Ullah F, Shah D, Gani A, Maple C, Wang Y, Shah Abrar M, Islam S. U. Deep GRU-CNN model for COVID-19 detection from chest X-rays data. *IEEE Access*. 2021; 10: 35094-35105.
35. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv* 2014; preprint arXiv:1409.1556.
36. Sotoudeh-Paima S, Jodeiri A, Hajizadeh F, Soltanian-Zadeh H. (Multi-scale convolutional neural network for automated AMD classification using retinal OCT images. *Computers in biology and medicine*. 2022; 144: 105368.
37. Spaide RF, Curcio CA. Drusen characterization with multimodal imaging. *Retina*. 2010; 30: 1441-1454
38. Srinivasan P P, Kim L A, Mettu P S, Cousins S W, Comer G M, Izatt J A, Farsiu S. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomedical optics express*. 2014; 5(10): 3568-3577.
39. Sun J K, Lin M M, Lammer J, Prager S, Sarangi R, Silva P S, Aiello L P. Disorganization of the retinal inner layers as a predictor of visual acuity in eyes with center-involved diabetic macular edema. *JAMA Ophthalmology*. 2014; 132: 1309-1316.
40. Sun Z, Sun Y. Automatic detection of retinal regions using fully convolutional networks for diagnosis of abnormal maculae in optical coherence tomography images. *Journal of biomedical optics*. 2019; 24(5): 056003-056003.
41. Sunija A P, Kar S, Gayathri S, Gopi V P, Palanisamy P. Octnet: A lightweight CNN for retinal disease classification from optical coherence tomography images. *Computer methods and programs in biomedicine*. 2021; 200: 105877.
42. Taibouni K, Miere A, Samake A, Souied E, Petit E, Chenoune Y. Choroidal neovascularization screening on OCT-angiography Choriocapillaris images by convolutional neural networks. *Applied Sciences*. 2021; 11(19): 9313.
43. Tayal A, Gupta J, Solanki A, Bisht K, Nayyar A, Masud M. DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases. *Multimedia Systems*. 2022; 28:1417-1438.
44. Tvenning A O, Hanssen S R, Austeng D, Morken T S. Deep learning identify retinal nerve fibre and choroid layers as markers of age-related macular degeneration in the classification of macular spectral-domain optical coherence tomography volumes. *Acta Ophthalmologica*. 2022; 100(8): 937-945.
45. Wang M, Zhu W, Shi F, Su J, Chen H., Yu K, Zhou Y, Peng Y, Chen Z, Chen, X. MsTGANet: Automatic drusen segmentation from retinal OCT images. *IEEE Transactions on Medical Imaging*. 2021; 41(2): 394-406.
46. Wang Y, Lucas M, Furst J, Fawzi A A, Raicu D. Explainable deep learning for biomarker classification of oct images. In 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE) 2020; 204-210.
47. Xu L, Wang L, Cheng S, Li Y. MHANet: A hybrid attention mechanism for retinal diseases classification. *Plos one*. 2021; 16(12): e0261285.
48. Zweifel SA, Imamura Y, Spaide TC, Fujiwara T, Spaide RF. Prevalence and significance of subretinal drusenoid deposits (reticular pseudodrusen) in age-related macular degeneration. *Ophthalmology*. 2010;117(9):1775-1781.

This project has been done within the “Staž za miedzą” restricted grant funded by the Medical University in Lublin, Poland (Chair and Department of General and Pediatric Ophthalmology, Medical University of Lublin, Poland and Faculty of Electrical Engineering and Computer Science, Department of Computer Science, Lublin University of Technology, Lublin, Poland). The study was carried out as a part of the project “Lubelska Unia Cyfrowa – Wykorzystanie rozwiązań cyfrowych i sztucznej inteligencji w medycynie – projekt badawczy”, no. MEiN/2023/DPI/2194.

Paweł Powroźnik:  <https://orcid.org/0000-0002-5705-4785>

Maria Skublewska-Paszowska:  <https://orcid.org/0000-0002-0760-7126>

Robert Rejdak:  <https://orcid.org/0000-0003-3321-2723>

Katarzyna Nowomiejska:  <https://orcid.org/0000-0002-5805-8761>



This work is licensed under the Creative Commons BY-NC-ND 4.0 license.