International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 9-No.2 (2023) pp. 62-67

http://dergipark.org.tr/en/pub/ijcesen

ISSN: 2149-9144

Research Article

Using Machine Learning to Detect Different Eye Diseases from OCT Images

Şükrü AYKAT[1]*, Sibel SENAN[2]

[1] Mardin Artuklu University, Department of Computer Technologies, 47510, Mardin-Turkey
* **Corresponding Author** : **Email:** sukruaykat@gmail.com - **ORCID :** 0000-0003-1738-3696

[2] Istanbul University-Cerrahpasa, Department of Computer Engineering, 34320, Istanbul-Turkey
Email: ssenan@iuc.edu.tr - **ORCID :** 0000-0001-6773-0428

Article Info:

DOI: 10.22399/ijcesen.1297655
Received : 16 May 2023
Accepted : 05 June 2023

Keywords

Retinal Disease
Optical Coherence Tomography (OCT)
Machine Learning
Deep Learning
DenseNet

Abstract:

Diseases or damage to the retina that cause adverse effects are one of the most common reasons people lose their sight at an early age. Today, machine learning techniques, which give high accuracy results in a short time, have been used for disease detection in the biomedical field. Optical coherence tomography is an advanced tool for the analysis, detection and treatment of retinal diseases by imaging the retinal layers. The aim of this study is to detect eight retinal diseases that can occur in the eye and cause permanent damage as a result, using machine learning from eye tomography images. For this purpose, hyperparameter settings were applied to six deep learning models, training was performed on the OCT-C8 dataset and performance analyzes were made. The performance of these hyperparameter-tuned models was also compared with previous eye disease detection studies in the literature, and it was seen that the classification success of the hyperparameter-tuned DenseNet121 model presented in this study was higher than the success of the other models discussed. The fine-tuned DenseNet121 classifier achieved 97.79% accuracy, 97.69% sensitivity, and 97.79% precision for the OCT-C8 dataset.

1. Introduction

One of the five sense organs, which has a very important place in human life, is the eye. The retina, on the other hand, is an important layer of the eye consisting of color and light-sensitive cells and nerve fibers that provide vision directly related to the brain. In order to make sense of the perceived visual signals, it is transferred to the brain to complete the visual event [1]. Diseases that occur in the retina often cause blindness and severe vision loss. Early detection of the disease is of great importance in order to avoid incurable eye damage caused by these diseases [2]. Many imaging modalities are used to evaluate and monitor retinal abnormalities. Optical coherence tomography (OCT) and Fundus imaging are widely used.

Recently, artificial intelligence has taken an important place in disease detection and diagnosis in the biomedical field. Artificial intelligence applications provide great convenience in disease diagnosis by performing classification processes that take a long time and are tiring for experts [3], [4]. Machine learning is a subfield of artificial intelligence designed to mimic human intelligence by learning from data [5]. Deep learning, on the other hand, is a subset of machine learning based on artificial neural networks architecture. In general, artificial neural networks consist of input layer, several hidden layers and output layers. Techniques based on machine learning have shown success in many industries, including finance, pattern recognition, entertainment, biomedical and medical applications. Liu et al. [6] in estimation of stock price, Sarkar et al. [7] and Tripathi in the recommendation system of music and Sharma [8] in diagnosis of brain tumor and skin cancer used machine learning methods. In the literature, machine learning models trained with retinal images obtained from different imaging techniques have achieved a high success rate in the diagnosis of retinal diseases.

Kermany et al. [1] applied the transfer learning model to a dataset of OCT images. The model's success in classifying diabetic macular edema and age-related macular degeneration demonstrated comparable performance to ophthalmologists.

Islam et al. [9] examined how to use deep transfer learning using OCT images to detect diabetic retinopathy. In their study, they explored how to optimize the models by retraining existing deep learning models. The proposed method outperformed existing methods in terms of accuracy and training time. Tayal et al. [10] presented a deep learning framework for the classification of four different retinal diseases. In their study, multiple eye deformity estimation was performed using three different convolutional neural network models. Noise removal, retinal layer removal and brightness enhancement preprocesses were applied to OCT images. With this study, classification accuracy was 96.5%, sensitivity 96% and specificity 98.6%. Lu et al. [11] developed a new intelligent model for retinal disease detection based on deep learning from OCT images. In the proposed system, sensitivity was 94%, specificity was 97.3%, and average accuracy was 95.9%. In the presented study, Subramanian et al. The OCT-C8 dataset [12], which is openly published on Kaggle and consists of 8 different classes, was used by [13]. These classes include central serous retinopathy (CSR), age-related macular degeneration (AMD), choroidal neovascularization (CNV), diabetic macular edema (DME), diabeticretinopathy (DR), drusen and macular hole (MH) diseases and their normal appearance. Six different deep learning models with hyperparameter tuning were applied on the dataset. The performances of the models discussed are compared with each other and with other studies in the literature, and the results are presented. The materials and methods used in this study are described in Section 2. The performances of the machine learning models discussed in Section 3 are examined. The general results of the study are given in Section 4.

2. Material and Methods

2.1 Dataset

The OCT-C8 dataset [12] consists of 8 classes as Normal, CSR, DME, DR, AMD, CNV, DRUSEN, and MH. This dataset is divided into three sections, Training, Validation, and Testing. The distributions are shown in Table 1. The OCT-C8 dataset, which was created from images collected from various sources such as Open-ICPSR and Kaggle, was subjected to data preprocessing. As preprocessing, cropping, filling, horizontal rotation and image enhancement techniques were used. The number of samples in the data set reached 24,000 images after image augmentation [13]. Sample images from the dataset are shown in Figure 1.

Table 1. *OCT-C8 dataset distribution*

Classes	Number of	Train	Validation	Test
AMD	3.000	2.300	350	350
CNV	3.000	2.300	350	350
CSR	3.000	2.300	350	350
DME	3.000	2.300	350	350
DR	3.000	2.300	350	350
DRUSEN	3.000	2.300	350	350
MH	3.000	2.300	350	350
Normal	3.000	2.300	350	350
Total	**24.000**	**18.400**	**2.800**	**2.800**

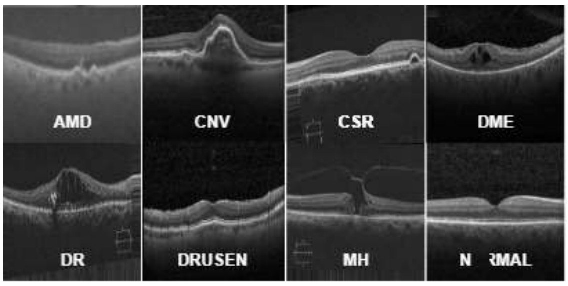

Figure 1. *Sample images of the OCT-C8 dataset*

2.2 Convolutional Neural Network (CNN)

CNN is one of the most widely used deep learning architectures for efficient training through multiple layers [14]. Figure 2 depicts the overall architecture of the CNN. The Convolutional, pooling, and fully connected layers are the basic layers of the CNN architecture. In the convolution layer, a filter is moved on the

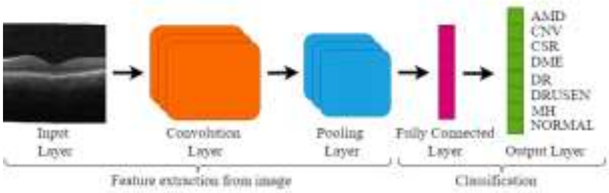

Figure 2. *General architecture of convolutional neural network*

given input variables by a user-specified number of steps. As a result of the convolution layer, feature maps are formed and used as the input of the next layer. The pooling layer is used to reduce the size of a feature. The output size of the layer is smaller than the previous layer. The output of the pooling layer is given to the fully connected layer as one-dimensional input [15].

2.3 DenseNet121

DenseNet121 algorithm proposed by Huang et al. [16] is the current architecture of CNN used for visual object recognition with fewer parameters. DenseNet combines previous layer output with subsequent layers with combined point attributes.

2.4 EfficientNetV2S

Tan and Le [17] proposed a convolutional neural network called EfficientNet in May 2019. The authors used a multidimensional hybrid model method to improve both accuracy and speed of the model. EfficientNet improves accuracy and provides speed by balancing network width, depth and resolution with integrated scaling. EfficientNet is seven models modified from the most basic model B0 to B7 in terms of channels, layers and resolution.

2.5 InceptionV3

Inception network is a pre-trained ESA model introduced by Google in 2014 [18]. This network consists of 22 layers with filters of different sizes used for maximum pooling and extracting features at various scales. Small filters are used to save time in calculations. In 2015, Google released the 48-layer InceptionV3 [19] to reduce parameters in the Inception model.

2.6 MobileNet

MobileNet is a new type of convolutional neural network. With its high efficiency and low power consumption features, MobileNet provides a high accuracy in image classification and recognition. MobileNet optimizes the standard convolution layer and divides the standard convolution process into two parts. These are deep convolution and point convolution. Deep convolution applies a single convolution kernel to the single input channel of each feature map for convolution computation. Point convolution is a standard convolution with a 1x1 convolution kernel [20].

2.7 VGG16

VGG16 is a neural network based model with approximately 138 million parameters and 16 layers. ReLU activation at the end of each convolution layer and maximum pooling at the end of all blocks are used to minimize dimensions [21]. It is among the best image classification models of neural networks.

2.8 Xception

It is a CNN architecture introduced by Chollet [22], which consists entirely of deeply separated convolution layers. In order to reduce the size of the problem in the Xception model, 1×1 convolution and deeply separable convolutions, which are a combination of deep convolution and point convolution, are used. There are approximately 23 million parameters in the Xception architecture.

2.9 Proposed Hyperparameter Tuned Model

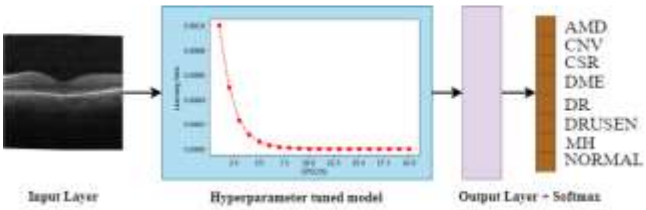

Figure 3. *Proposed hyperparameter tuned model architecture*

In this study, convolutional neural network models DenseNet121, EfficientNetV2S, InceptionV3, MobileNet, VGG16 and Xception were used. The architecture of the proposed hyperparameter tuned model is shown in Figure 3. The OCT images applied to the input layer of the models are 224x224x3 in size. Before training on the data set, these models were hyperparameter tuned.

Model trainings were conducted during 20 epochs. While the learning rate was 0.001 at the beginning of the training, the learning rate was halved as the number of epochs increased. The reason we do this is that if the learning rate is too small at the beginning, it may get stuck at the optimum local value and cause the optimum global value to never be reached. In the last layer, OCT images were classified using the softmax function.

3. Results and Discussions

3.1 Confusion Matrix

The confusion matrix is used to measure the performance of a model. As seen in Table 2, in the confusion matrix, the rows show the actual sample numbers in the test set, and the columns show the sample numbers predicted by the model.

Table 2. *Confusion Matrix*

<table>
<tr><td colspan="2" rowspan="2"></td><th colspan="2">Predicted Class</th></tr>
<tr><th>correct</th><th>incorrect</th></tr>
<tr><td rowspan="2">Actual Class</td><td>correct</td><td>True Positive - [TP]</td><td>False Negative - [FN]</td></tr>
<tr><td>incorrect</td><td>False Positive - [FP]</td><td>True Negative - [TN]</td></tr>
</table>

Accuracy

It is the ratio of the number of correct and incorrect samples found to be correct to the total number of samples.

Accuracy = (TP+TN) / (TP+TN+FP+FN) (1)

Sensitivity

It is the rate at which samples in the real class are guessed correctly.

Sensitivity = (TP) / (TP+FN) (2)

Precision

It is the ratio of the number of correctly predicted samples to the number of samples that are actually correct.

Precision = (TP) / (TP+FP) (3)

The accuracy, sensitivity and precision metric values of the CNN learning models are shown in Table 3. Compared to other learning models used in this study, the DenseNet121 model has the best performance with an overall accuracy rate of 97.79%. Among the ESA learning models, the lowest performing models were VGG16 with an overall accuracy of 96.14%.

Table 3. *Performance metric values of hyperparameter-tuned deep learning models*

Models	Accuracy (%)	Precision (%)	Sensitivity (%)
DenseNet121	**97.79**	**97.79**	**97.69**
EfficientNetV2S	97.57	97.57	97.57
InceptionV3	97.71	97.71	97.60
MobileNet	96.36	96.42	96.32
VGG16	96.14	96.24	96.04
Xception	97.39	97.38	97.30

The resulting confusion matrix for the DenseNet121 model is shown in Figure 4. The diagonal elements of the confusion matrix represent the correct classifications. The remaining items are misclassifications. The rows show the predicted classes, the columns the actual classes. Accuracy, precision, and sensitivity class criteria are determined using the confusion matrix. Table 4 shows the comparison of the best performing hyperparameter tuned DenseNet121 model in our study with the performance of other studies in the literature. When Table 4 is examined, it is seen that our model is classified with the highest

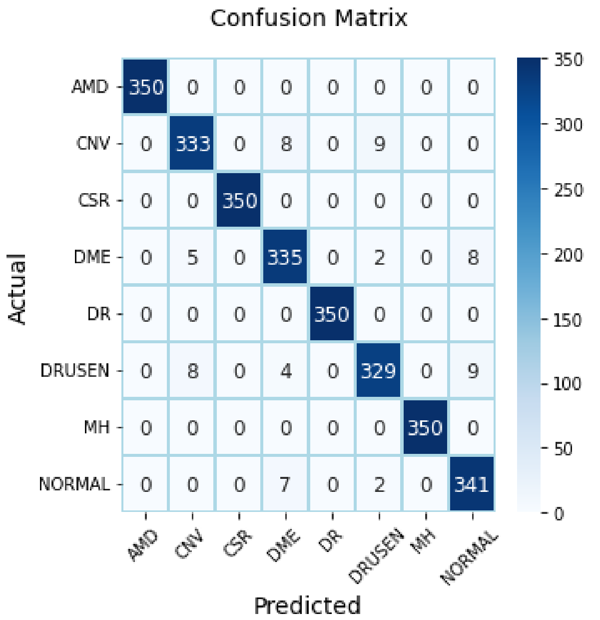

Figure 4. *Confusion matrix of the DenseNet121 model*

Table 4. *Comparison results of Hyperparameter tuned DenseNet121 model with performances of other studies*

Study	Model	Number of Classes	Accuracy (%)
Kermany et al. [1]	InceptionV3	4	96.60
Islam et al. [9]	DenseNet 201	4	98.66
Tayal et al. [10]	DL-CNN	4	96.54
Lu et al. [11]	ResNet	4	95.90
Subramanian et al. [13]	VGG16	8	97.21
Hyperparameter-Tuned DenseNet121		***8***	***97.79***

performance of 8-class data. At the same time, our model performed better than most models that classify 4-class data. The high success of our proposed model is the result of changing the hyperparameter setting of existing deep learning models.

4. Conclusions

In this period when retinal diseases increase and negatively affect human life, we propose new results in deep learning models to classify retinal patients into eight classes using OCT images. Six hyperparameter-tuned deep learning models were used to classify AMD, CSR, CNV, DR, DME, DRUSEN, and MH patients and healthy individuals. It has been observed that the hyperparameter tuned deep learning models have good performances on the OCT dataset. The classification success of the hyperparameter tuned DenseNet121 model is 97.79%. This result demonstrates the effectiveness of the deep learning approach in detecting retinal diseases.

Within the scope of this study, applications can be made that can be used by ophthalmologists in the future and can give ideas to the experts. In addition, mobile applications can be developed to help physicians in the detection of retinal diseases.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] D. S. Kermany *et al.*, (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5);1122-1131 doi: 10.1016/J.CELL.2018.02.010.

[2] K. L. Pennington and M. M. DeAngelis (2016). Epidemiology of age-related macular degeneration (AMD): associations with cardiovascular disease phenotypes and lipid factors. *Eye Vis. (London, England)*, 3(1) doi: 10.1186/S40662-016-0063-5.

[3] X. Yang, T. Zhang, C. Xu, S. Yan, M. S. Hossain, and A. Ghoneim, (2016). Deep relative attributes. *IEEE Trans. Multimed.*, 18(9);1832–1842, doi: 10.1109/TMM.2016.2582379.

[4] M. S. Hossain, G. Muhammad, and A. Alamri, (2019). Smart healthcare monitoring: a voice pathology detection paradigm for smart cities, *Multimed. Syst.*, 25(5);565–575, doi: 10.1007/S00530-017-0561-X/FIGURES/9.

[5] I. El Naqa and M. J. Murphy, (2015). What Is Machine Learning?," in Machine Learning in Radiation Oncology: Theory and Applications, *I. El Naqa, R. Li, and M. J. Murphy, Eds. Cham: Springer International Publishing*, pp. 3–11. doi: 10.1007/978-3-319-18305-3_1.

[6] K. Liu, J. Zhou, and D. Dong, (2021). Improving stock price prediction using the long short-term memory model combined with online social networks. *J. Behav. Exp. Financ.*, 30; 100507, doi: 10.1016/J.JBEF.2021.100507.

[7] M. Sarkar, A. Roy, Y. Badr, B. Gaur, and S. Gupta, (2022). An Intelligent Music Recommendation Framework for Multimedia Big Data: *A Journey of Entertainment Industry*, pp. 39–67, doi: 10.1007/978-981-16-3828-2_3.

[8] S. Tripathi and N. Sharma, (2021). Computer-Based Segmentation of Cancerous Tissues in Biomedical Images Using Enhanced Deep Learning Model. 39(5);1208–1222, doi: 10.1080/02564602.2021.1994044.

[9] K. T. Islam, S. Wijewickrema, and S. O'Leary (2019). Identifying diabetic retinopathy from OCT images using deep transfer learning with artificial neural networks. *Proc. - IEEE Symp. Comput. Med. Syst.*, 281–286, doi: 10.1109/CBMS.2019.00066.

[10] A. Tayal, J. Gupta, A. Solanki, K. Bisht, A. Nayyar, and M. Masud, (2022). DL-CNN-based approach with image processing techniques for diagnosis of retinal diseases. *Multimed. Syst.*, 28(4);1417–1438, doi: 10.1007/S00530-021-00769-7/FIGURES/13.

[11] W. Lu, Y. Tong, Y. Yu, Y. Xing, C. Chen, and Y. Shen, (2018). Deep Learning-Based Automated Classification of Multi-Categorical Abnormalities From Optical Coherence Tomography Images. *Transl. Vis. Sci. Technol.*, 7(6);41–41, doi: 10.1167/TVST.7.6.41.

[12] Retinal OCT - C8, (2022). https://www.kaggle.com/datasets/obulisainaren/retinal-oct-c8/code (accessed May 09, 2023).

[13] M. Subramanian, K. Shanmugavadivel, O. S. Naren, K. Premkumar, and K. Rankish, (2022). Classification of Retinal OCT Images Using Deep Learning, *Int. Conf. Comput. Commun. Informatics, ICCCI* , doi: 10.1109/ICCCI54379.2022.9740985.

[14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, (1998). Gradient-based learning applied to document recognition, *Proc. IEEE*, 86(11);2278–2323, doi: 10.1109/5.726791.

[15] M. M. Islam, T. N. Poly, B. A. Walther, H. C. Yang, and Y. C. Li, (2020). Artificial Intelligence in Ophthalmology: A Meta-Analysis of Deep Learning Models for Retinal Vessels Segmentation, *J. Clin. Med. 9;(*4);1018, doi: 10.3390/JCM9041018.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, (2017). Densely Connected Convolutional Networks. *Conf. Comput. Vis. Pattern Recognit.*, pp. 4700–4708, Accessed: Sep. 11, 2022. [Online].: https://github.com/liuzhuang13/DenseNet.

[17] M. Tan and Q. V. Le, (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *36th Int. Conf. Mach. Learn. ICML 2019*, 2019;10691–10700, doi: 10.48550/arxiv.1905.11946.

[18] C. Szegedy et al., (2015). Going deeper with convolutions; Going deeper with convolutions. doi: 10.1109/CVPR.2015.7298594.

[19] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, (2016). Rethinking the Inception Architecture for

Computer Vision; Rethinking the Inception Architecture for Computer Vision. *IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.

[20] A. G. Howard *et al.*, (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications.

[21] K. Simonyan and A. Zisserman, (2015). VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION. *Comput. Vis. Pattern Recognit.*, 2015, Accessed: Apr. 26, 2022. [Online]. Available: http://www.robots.ox.ac.uk/

[22] F. Chollet, (2017). Xception: Deep Learning with Depthwise Separable Convolutions. doi: 10.1109/CVPR.2017.195.