Erciyes Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 40, Sayı 1, 2024

Erciyes University
Journal of Institue Of Science and Technology
Volume 40, Issue 1, 2024

Detection of Retinal Diseases from Fundus Images Using Deep Learning and Adaptive Histogram Equality

Ali Emre Gök[*1] , **Şakir Taşdemir**[2]

[*1] Nevşehir Hacı Bektaş Veli University, Faculty of Engineering and Architecture, Computer Engineering, NEVŞEHİR
[2] Selcuk University, Faculty of Technology, Computer Engineering, KONYA

(Alınış / Received: 29.01.2024, Kabul / Accepted: 01.04.2024, Online Yayınlanma / Published Online: 30.04.2024)

Keywords
Deep Learning,
Diabetic Retinopathy,
Cataracts,
Glaucoma,
Image Classification

Abstract: Recently, various eye diseases such as cataracts, diabetic retinopathy, glaucoma, macular edema, myopia, and astigmatism have been seen frequently. Cataracts, diabetic retinopathy, and glaucoma cause blurred vision, loss of vision, and blindness in cases where they are left untreated and undiagnosed. Lack of experts and equipment, hardware problems, and erroneous decisions made by experts cause problems in the diagnosis process. Because of these reasons, computer-aided diagnosis systems that can diagnose accurately are required. Deep learning algorithms performed well in the field of health, recently. These results show that deep learning algorithms can be used in the diagnosis of eye diseases. In this study, various CNN models were used for classifying eye diseases such as cataracts, diabetic retinopathy, and glaucoma from fundus images. In the image preprocessing stage, the Contrast Limited Adaptive Histogram Equalization method was used. Experimental results demonstrate that VGG16 was the most successful model among the evaluated models in this study and the Contrast Limited Adaptive Histogram Equalization method increased the performance.

Derin Öğrenme ve Adaptif Histogram Eşitleme Kullanarak Retinal Hastalıkların Fundus Görüntülerinden Tespiti

Anahtar Kelimeler
Derin öğrenme,
Diyabetik retinopati,
Katarakt,
Glokom,
Görüntü sınıflandırma

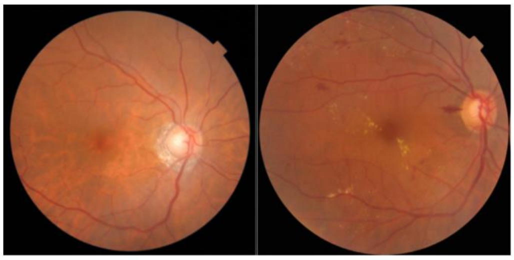
Öz: Günümüzde, katarakt, diyabetik retinopati, glokom, maküler ödem, miyop ve astigmat gibi çeşitli göz hastalıkları sıklıkla görülmektedir. Bu hastalıklardan katarakt, diyabetik retinopati ve glokom, teşhis ve tedavi edilmedikleri durumlarda, bulanık görmeye, görme kaybına ve hatta körlüğe neden olmaktadır. Uzman ve donanım eksikliği, donanımsal problemler ve uzmanlarca verilen hatalı kararlar gibi çeşitli nedenlerle, teşhis aşamasında sorunlarla karşılaşılmaktadır. Bu nedenlerden dolayı, Bilgisayar Destekli Teşhis sistemlerine ihtiyaç duyulmaktadır. Son zamanlarda, derin öğrenme algoritmalarıyla gerçekleştirilen çalışmalarda, başarılı sonuçlar elde edilmiştir. Bu başarılı sonuçlar, derin öğrenmenin, göz hastalıklarının teşhisinde kullanılabileceğini göstermektedir. Bu çalışmada, çeşitli CNN modelleri kullanılarak, fundus görüntüleri üzerinden, katarakt, diyabetik retinopati ve glokom gibi göz hastalıklarının sınıflandırılması gerçekleştirilmiştir. Fundus görüntülerinde, Kontrast Sınırlı Adaptif Histogram Eşitleme yöntemi kullanılmıştır. Deneysel sonuçlar, VGG16 modelinin, bu üç model arasında en başarılı model olduğunu ve Kontrast Sınırlı Adaptif Histogram Eşitleme yönteminin de performansı arttırdığını göstermektedir.

*İlgili Yazar, email: aliemregok@nevsehir.edu.tr

1. Introduction

Diabetes is an increasingly common disease today that poses a risk, especially for people between the ages of 40-59. It is a disorder caused by irregular changes in the amount of glucose in the blood, which provides energy for the activities in the body. The disease affects every individual in developed or developing countries with various levels of healthcare. It is estimated that the number of diabetic patients will reach 629 million around the world by 2045. High levels of blood sugar cause this excess glucose to accumulate in the vessels and block blood flow to the organs and result in various diseases [1].

Diabetic retinopathy (DR) is one of the diseases caused by diabetes and is becoming common day by day [1, 2]. It is estimated that DR, which is among the most important eye diseases causing visual impairment, will be seen in two hundred million people in 2040 [3, 4]. DR, a microvascular disease that develops over time [2, 5], is seen in people who have had diabetes for at least 10 years and have not had a proper eye examination during this period [1]. DR results from the loss of function of the cells in the back of the eye, which are responsible for sensing light and sending signals to the brain, due to damage to the blood vessels in that area because of diabetes [6-8]. In general, the vessels become occluded and become abnormally visible on the retina, while conditions like swelling or leakage of damaged vessels are seen in large cases [7]. Depending on the damage, visual disturbances or loss of vision are experienced [6, 7]. The most obvious and general symptoms used in the diagnosis of DR are exudates. Exudate is secreted at the points where bleeding occurs due to damage to the vessels. The size, shape, etc. characteristics of these secreted exudates indicate the stage of DR [7]. Fundus images of DR and healthy eyes are shown in Figure 1.



(a) (b)

Figure 1. Fundus images of healthy (a) and DR (b) eyes [9]

Cataracts is the most common eye disease causing vision loss and it is estimated that forty million people will lose their vision due to cataracts by 2025 [10, 11]. The lens in the eye, consisting of water, protein, and ectodermal tissue, allows the light to fall onto the retina thanks to its transparent structure and shape. The loss of the transparency of the lens due to several reasons, preventing the passage of light to the retina is called cataracts [12-14]. Cataracts, which mostly occurs due to age [15], is also caused by other diseases like diabetes, trauma, congenital factors, and smoking [13, 16]. In addition to the most common type of cataracts, nuclear cataracts, there are also cortical and posterior types. These species may be seen separately or together [12, 13]. Fundus images of cataracts patients and healthy eyes are given in Figure 2.

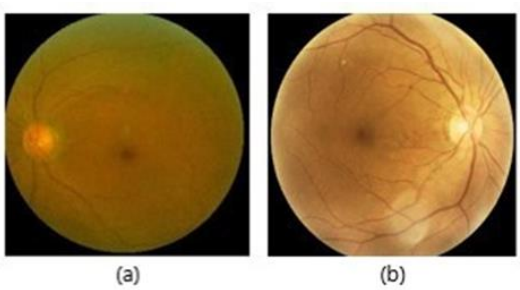


Figure 2. Fundus images of cataracts (a) and healthy (b) eyes [17]

Another important eye disorder, glaucoma is a disease that causes loss of vision because of severe damage to optical nerves. Since the disease is asymptomatic until it progresses, it is called silent vision thief [18-20]. Glaucoma is an irreversible discomfort due to the damage it causes to the OD and nerves. Although there is no cure for glaucoma, in cases of early diagnosis, its progression can be slowed down by various methods and serious consequences such as blindness can be prevented [18, 20, 21]. Figure 3 shows the fundus images of glaucoma and healthy eyes.

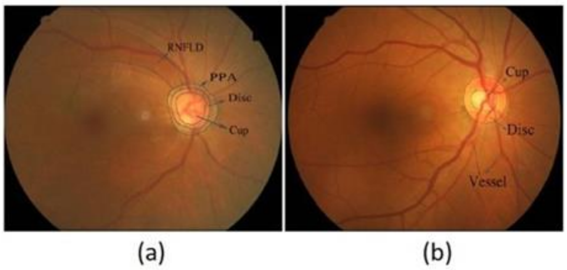


Figure 3. Fundus images of glaucoma (a) and healthy (b) [19]

Diagnosis of various eye diseases is made by ophthalmologists by inspecting the fundus images. In rural areas, number of experts is insufficient. Besides, intense workloads or lack of expertise caused ophthalmologists to make misdiagnoses. In addition, ophthalmologists with different experiences may make different diagnoses on the same fundus images. Because of these reasons, subjective results are obtained. Therefore, Computer Aided Diagnosis systems are needed for fast and accurate diagnosis. Various approaches have been used for diagnosing diseases from fundus images in the literature.

In a study conducted by B. Tymchenko et al. in 2020, a deep learning (DL) method that performs DR diagnosis on fundus images was proposed. In the proposed system created by EfficientNet-B4, EfficientNet-B5, and SE-ResNetXt50 models, several datasets are used together: EyePACs 2015 dataset, which is an open-source dataset that contains 35126 fundus images, IDRiD dataset with 413 fundus images, and Messidor dataset contains 1200 fundus images. At the end of the study, 98% accurate results were obtained [22]. In another study by S. Alrajjou et al. in 2022, it was aimed to diagnose DR by combining DL approaches. In the system using an improved PSO-NDAE-based CNN classifier, fundus images of 35126 healthy and DR patients obtained from the Kaggle database were used. According to the results of the study, it was observed the suggested system had the highest accuracy rate of 85.78% in comparison with other algorithms [23]. A DL model aimed at DR diagnosis via fundus images was proposed by A. Ayala et al. in 2021. The model was created using DenseNet121 architecture and transfer learning methodology. At the end of the study performed with the APTOS dataset with 5590 fundus images and the Messidor dataset with 1744 fundus images, it was observed that the proposed model produced 97.78% accurate results [24].

In a study performed by E. Acar et al. in 2021, cataracts diagnosis on fundus images was aimed. 548 cataracts and 5535 healthy images were selected from the dataset which contains 6392 fundus photographs in total, belonging to five thousand different patients, taken from the Kaggle Ocular Disease Recognition Database. At the end of the study that was carried out using VGGNet and DenseNet models, results with an accuracy of 97.94% were obtained [25]. In another study conducted by B. Kalyani et al. in 2023, a system for diagnosing cataracts disease was proposed using a combination of CNN and LTSM models. It was observed that the proposed system produced 98% correct results at the end of the study, which was carried out on a dataset consisting of UMN, WEB data, and data from the internet and containing a total of 4630 fundus images, including 2615 cataracts and 2015 healthy images [26]. In 2021, M. S. Junayed et al. proposed a DL network named CataractNet that aimed to diagnose cataracts through fundus images. To reduce running cost and average runtime, loss and activation functions are arranged to train the network with small kernels, fewer training layers, and parameters. According to the results of the study, which was performed using 2679 cataracts and 2067 healthy fundus images, the proposed CataractNet produced 99.13% accurate results. The obtained results showed that the proposed system is more successful than other developed cataracts detection systems [27]. In 2020, a study by F. Li et al. aimed to diagnose glaucoma with a DL approach based on ResNet101. A total of 34279 fundus images were used, of which 12618 were glaucoma, 1114 were suspicious, and 12853 were healthy, labeled by eight different glaucoma experts. Independently of this dataset, 3481 fundus images (1524 healthy, 1442 glaucoma, 515 suspected) were used to evaluate the proposed model. At the end of the study, accuracy, sensitivity, and specificity rates were observed as 95% [28].

In another study carried out by M. N. Bajwa et al. in 2019, a 2-stage system was proposed for glaucoma diagnosis. In the first stage of the proposed system, the CNN model was used, and optical disc detection was performed. In the second stage, glaucoma diagnosis was made using a deep CNN model. At the end of the study, performed on seven different datasets, namely ORIGA, HRF, OCT&CFI, DIARETDB1, DRIVE, DRIONS-DB, and Messidor, it was observed that the algorithm developed for OD detection produced better results in different datasets than its counterparts. In the diagnosis of glaucoma which was performed in the second stage of the system, it was observed that 79.67% correct results were produced [29]. In a study in 2019 by M. Kim et al. a DL system was proposed for diagnosing glaucoma using fundus images. Using the CNN and gradient-weighted class activation mapping (Grad-CAM) models results with 96% accuracy, 96% sensitivity, and 100% specificity were produced in the study, which was carried out with the Dataset-Optic-Disc dataset containing 1903 fundus images [30].

In 2019, Hongyan Zhang et al. developed an algorithm for automatically diagnosing and grading cataracts using fundus images. ResNet18 and gray-level co-occurrence matrix were used for feature extraction, while two support vector machines and a fully connected neural network were used for cataracts detection. A total of 1352 fundus images with images of various levels of cataracts were used in the study. Results of the study showed that the proposed algorithm has 92.66% accuracy for six-level grading and 94.75% for four-level grading for cataracts [31]. In another study performed in 2019, Tao Li et al. collected 13673 fundus images to evaluate DL models' performance in clinical applications for DR diagnosis. Various models were used in the study for DR grading, lesion segmentation, and lesion detection. From the results of the study, 0.83% accurate results were obtained for DR classification. On the other hand, the models used performed poorly in segmentation and detection operations [9]. Table 1 shows the summary of the literature.

Table 1. Summary of the literature

Study Author(s)	Year	Focus	Method/Model Used	Datasets Used	Accuracy
B. Tymchenko et al.	2020	DR Diagnosis	EfficientNet-B4, EfficientNet-B5, SE-ResNetXt50	EyePACs 2015, IDRiD, Messidor	98%
S. Alrajjou et al.	2022	DR Diagnosis	PSO-NDAE-based CNN	Kaggle (35,126 fundus images)	85.78%
A. Ayala et al.	2021	DR Diagnosis	DenseNet121, Transfer Learning	APTOS, Messidor	97.78%
E. Acar et al.	2021	Cataracts Diagnosis	VGGNet, DenseNet	Kaggle Ocular Disease Recognition Database	97.94%
B. Kalyani et al.	2023	Cataracts Diagnosis	CNN, LTSM	UMN, WEB, Internet Data	98%
M. S. Junayed et al.	2021	Cataracts Diagnosis	CataractNet	2,679 cataracts and 2,067 healthy images	99.13%
F. Li et al.	2020	Glaucoma Diagnosis	ResNet101	34,279 images (12,618 glaucoma, 1,114 suspicious, 12,853 healthy)	95%
M. N. Bajwa et al.	2019	Glaucoma Diagnosis	CNN (2-stage system for OD detection and diagnosis)	ORIGA, HRF, OCT&CFI, DIARETDB1, DRIVE, DRIONS-DB, Messidor	79.67%
M. Kim et al.	2019	Glaucoma Diagnosis	CNN, Grad-CAM	Dataset-Optic-Disc (1,903 fundus images)	96%
Hongyan Zhang et al.	2019	Cataracts Grading	ResNet18, gray-level co-occurrence matrix, SVMs, FCNN	1,352 fundus images	92.66%-94.75%
Tao Li et al.	2019	DR Diagnosis & Grading	Various DL models for grading, segmentation, detection	13,673 fundus images	0.83%

Although reliable results are obtained from current studies in the literature, this field is still an active research topic. Considering the pros and cons of the current studies in the literature, we proposed a robust approach for detecting diseases from fundus images to evaluate the effects of the CLAHE method with models consist different blocks of varied sizes. Firstly, we used the Contrast Limited Adaptive Histogram Equality (CLAHE) method to improve histogram levels of fundus images. Then we used VGG16, MobileNet, and DenseNet169 models as classifiers. The main reason we selected these models is that the number of parameters is small, medium, and large scale. In addition, the VGG model uses stack convolution, the MobileNet model uses residual blocks and the DenseNet169 model uses densely connected blocks. Thus, different CNN models which use various blocks of varied sizes have been evaluated. The comparison in this study is based on the performance of these three models.

2. Material and Method

2.1. Dataset

There are 1038 cataracts, 1098 DR, 1007 glaucoma, and 1074 healthy fundus images in the dataset consisting of 4217 images in total. Images in the dataset were obtained from various sources such as IDRiD, Oculur recognition, and HRF [17]. For the training of DL models used, the dataset was divided into 8:1:1 ratio. 80% of data was used in training process and 10% was used for cross validation during this process. The test process was conducted with the data that was not used in the training process. The number of fundus images for each class is given in Table 2. Examples of fundus images from each class are also given in Figure 4.

Table 2. Number of fundus images

Class	Number of Fundus Images
Cataracts	1038
Diabetic Retinopathy	1098
Glaucoma	1007
Healthy	1074
Total	4217

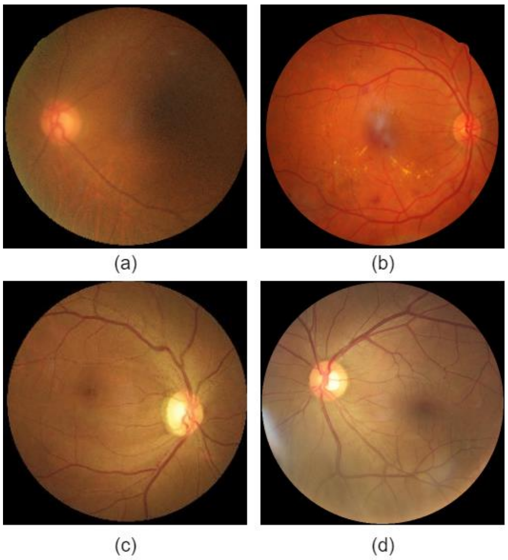


Figure 4. Cataracts (a), DR (b), Glaucoma (c) and Healthy (d) fundus images [17]

2.2. Convolutional Neural Networks

Deep learning is a type of multi-layered artificial neural network that enables computers to learn and operate on their own, similar to the human way of thinking. The concept of DL aims to enable computers to learn complex models in large datasets and to produce desired results when necessary [32, 33]. Computers have difficulties processing data when raw data is presented as input. DL eliminates this problem by dividing the data to process it in different layers [33]. DL, which has developed rapidly in recent years, owes this development to the software libraries developed, the increasing datasets, and the ease of access to these sets. Additionally, the ability to design deeper and more complex artificial neural networks with the developing technology has also contributed to the development of DL [34].

Convolutional Neural Networks (CNN) is the most widely used DL architecture. CNN has the ability to learn features on images without any feature extraction preprocess. CNN consists of three main layers: convolutional, pooling, and fully connected. The convolutional layer uses filters of varied sizes for feature extraction from images. Filter matrices are used over the images for the convolution process and to produce feature maps. The pooling layer makes the larger feature maps smaller to produce new and small feature maps while keeping the dominant features unchanged. There are some kinds of pooling methods such as tree pooling, average poling, max pooling, global average pooling (GAP), etc. The third of the main layers of CNN is a fully connected layer. This layer is placed as the last layer and produces the output of the system. Every neuron in the fully connected layer is connected to the previous layer [35-38]. CNN architecture is given in Figure 5.

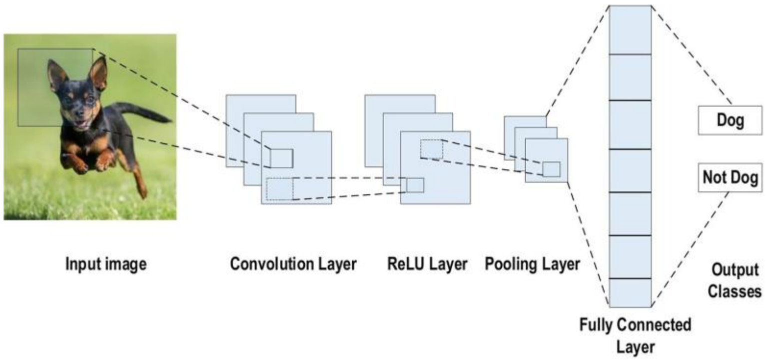


Figure 5. Image classification example with CNN architecture [35]

In this study, three different DL models were trained separately, and their performances were compared. The GAP layer was added to the last layer of the models and the Dense layer consisting of 128, 256, and 512 neurons, respectively, was added to the continuation. The first model of these models is the MobileNet model. This model was developed by Google based on deeply separable filters. MobileNet is faster and smaller in size than other models. It was developed to perform DL applications on mobile devices [39]. Another model used in the study is the VGG16 model. This model consists of a total of sixteen layers, 13 convolutional and 3 fully connected layers. VGG16 is used in image classification and object recognition applications [40]. VGG16 architecture is given in Figure 6.

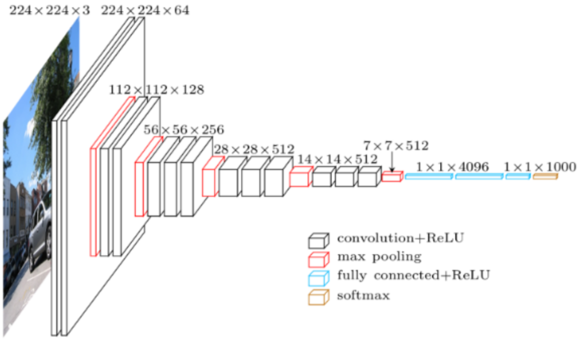


Figure 6. Architecture of VGG16 [41]

The third and last DL model used in this study is the DenseNet169 model. This model was first announced by Gao et al. in the article "Densely Connected Convolutional Networks". Each layer of this model is connected to the previous layer and thanks to this connection, each layer can also use the information of the previous layers. In this way, the network shows better learning performance. DenseNet169 model is used for image classification, object detection, and face recognition [42]. An example of DenseNet169 architecture is given in Figure 7.

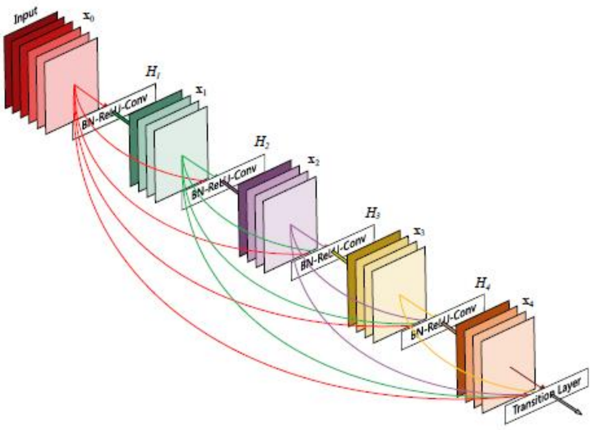


Figure 7. DenseNet169 with 5 blocks [42]

2.3. Adaptive Histogram Equalization with Contrast Limit

Contrast Limited Adaptive Histogram Equalization is a histogram equalization technique used in visual analysis. It is used to improve the contrast and highlight the details in the image. Histogram equalization changes the pixel density values of an image to make it more evenly distributed in terms of contrast. However, classical histogram equalization techniques sometimes over-amplify and generate noise. CLAHE divides the image into small regions and performs separate histogram equalization in each region to fix this problem. The CLAHE technique is used in applications where details such as retinal diseases are important, especially in areas such as medical imaging. Within the scope of this study, the training process was conducted by applying the CLAHE method to all images in the dataset.

3. Experimental Setup

CNN models were trained for 50 epochs using a batch size of thirty-two and a binary cross entropy loss function. The optimization algorithm was executed for fifty iterations. The computer used in the study had 128 GB of RAM and a 3060 Ti GPU. The Python programming language was used with the Keras and Scikit-learn libraries.

3.1. Performance Metrics

In this study, the results of each model were compared using four different performance evaluation metrics: accuracy, precision, recall, and F1-score. Equations of the metrics are given in Table 3. In the equations, TP represents true positive, TN represents true negative, FP represents false positive, and FN represents false negative. The accuracy is the ratio of the correctly predicted examples to all the examples. The precision indicates ratio of positive estimations to all positive examples. The recall shows the ratio of positive estimations to real positive samples. Finally, harmonic mean of the precision and the recall is called F1-score [43].

Table 3. Performance evaluation metrics

Metric	Mathematical Expression
Accuracy	$\frac{TP+TN}{TP+TN+FN+FP}$
Precision (P)	$\frac{TP}{TP+FP}$
Recall (R)	$\frac{TP}{TP+FN}$
F1-Score	$\frac{2\times P\times R}{P+R}$

4. Results

The results obtained from each model are arranged as tables. The accuracy, precision, recall, and F1-score values of the MobileNet model are given in Table 4. The accuracy of the MobileNet model was calculated as 92.5%.

Table 4. Performance of the MobileNet Model in Disease Classification

Class	Precision	Recall	F1-Score
Cataracts	0.917	0.943	0.930
DR	0.991	1	0.996
Glaucoma	0.880	0.863	0.871
Healthy	0.905	0.889	0.897
Weighted Average	0.924	0.925	0.925
Standard Deviation	0.041	0.052	0.039

The results and values of the VGG16 model, which is another model in the study, are given in Table 5. The results obtained have an accuracy rate of 91%.

Table 5. Performance of the VGG16 Model in Disease Classification

Class	Precision	Recall	F1-Score
Cataracts	0.949	0.886	0.916
DR	1	1	1
Glaucoma	0.793	0.902	0.844
Healthy	0.911	0.852	0.880
Weighted Average	0.915	0.911	0.912
Standard Deviation	0.092	0.055	0.058

Finally, the results and values of the DenseNet169 model are given in Table 6. The results of this model are 91.7% accurate.

Table 6. Performance of the DenseNet169 Model in Disease Classification

Class	Precision	Recall	F1-Score
Cataracts	0.917	0.943	0.930
DR	1	0.991	0.995
Glaucoma	0.952	0.784	0.860
Healthy	0.823	0.944	0.879
Weighted Average	0.923	0.918	0.917
Standard Deviation	0.065	0.078	0.052

The results of the study showed that the precision, recall, and F1-score values for cataracts diagnosis were 91.7%, 94.3%, and 93% for MobileNet, 94.9%, 88.6%, and 91.6% for VGG16 model, and 91.7%, 94.3%, and 93% for DenseNet169 model, respectively. These values showed that the DenseNet169 and MobileNet models were more successful than the VGG16 model in diagnosing cataracts.

When the performances of the models in DR diagnosis were examined, precision, recall, and F1-score values of the MobileNet model were measured as 99.1%, 100%, and 99.6%, respectively. The VGG16 model's values were measured as 100% for all three metrics. The DenseNet169 model had values of 100%, 99.1%, and 99.5%, respectively. The results showed that the VGG16 model was more successful than the MobileNet and DenseNet169 models in diagnosing DR.

The accuracy, recall, and F1-score values for glaucoma diagnosis were 88%, 86.3%, and 87.1% for the MobileNet model, 79.3%, 90.2%, and 84.4% for the VGG16 model, and 95.2%, 78.4%, and 86% for the DenseNet169 model, respectively. According to these values, it can be said that the MobileNet model may be a better choice for diagnosis of glaucoma.

Finally, the performances of the models in classifying healthy fundus images were examined. According to the results, the precision, recall, and F1-score values of the MobileNet model were measured as 90.5%, 88.9%, and 89.7%, the values of the VGG16 model were measured as 91.1%, 85.2%, and 88% and the values of the DenseNet169 model were measured as 82.3%, 94.4%, and 87.9%, respectively. These values indicated that the MobileNet model was more performant than the other models when it comes to classifying healthy fundus images.

5. Discussion

The confusion matrices, accuracy graphs, and loss graphs of the specified models are given in Figures 8-13. When Figure 8 is examined, it is seen that 9 data in total belonging to the healthy class were predicted to be glaucoma by the MobileNet model. When it comes to the VGG16 model, in Figure 10, out of a total of 14 data from the healthy class, 14 were predicted to be in the glaucoma class. Finally, in Figure 12, it is seen that 16 of the glaucoma images were predicted as healthy. These three common eye diseases are often confused with each other in terms of symptoms, each one has vision blur or loss, risk factors such as age, diabetes, or genetics, and being asymptomatic until the progress.

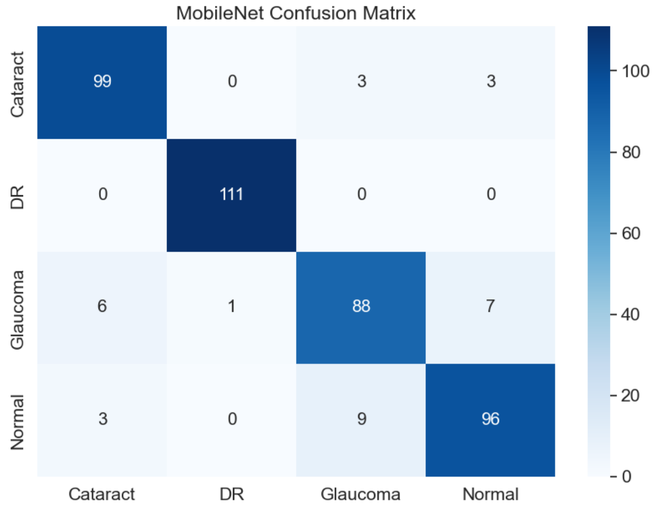


Figure 8. Confusion matrix derived from the classification of diseases with the MobileNet

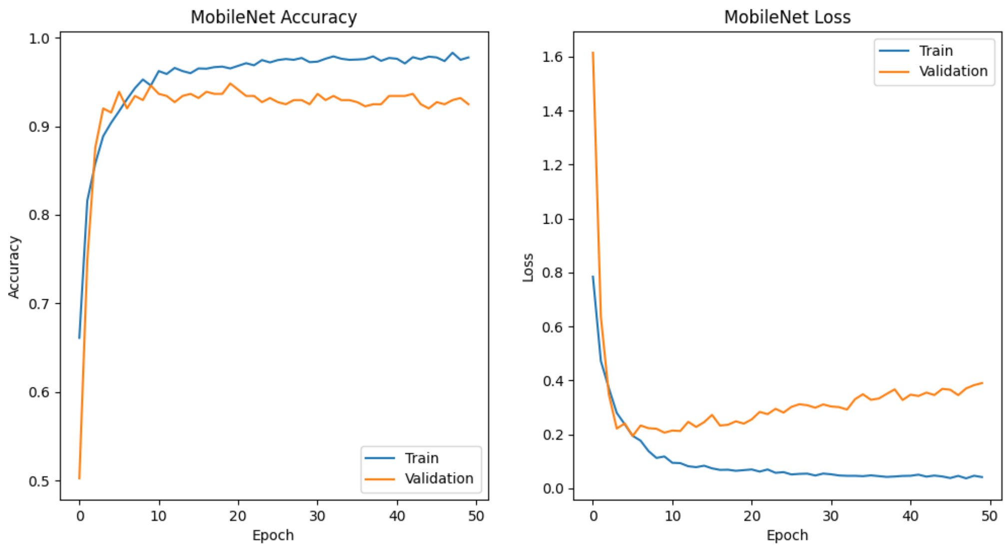


Figure 9. Accuracy and loss graphs of the MobileNet

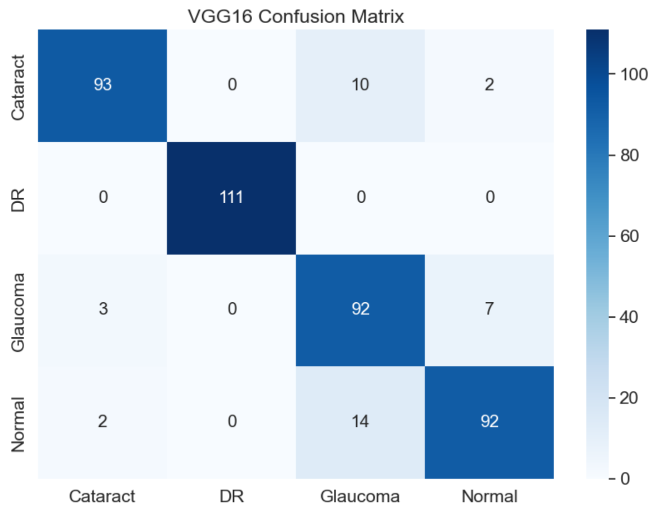


Figure 10. Confusion matrix derived from the classification of diseases with the VGG16

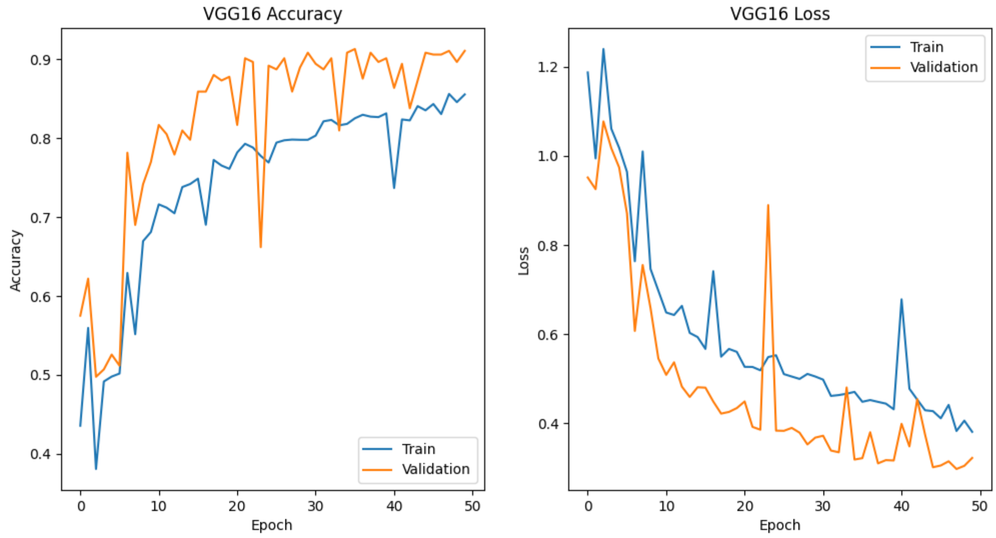


Figure 11. Accuracy and loss graphs of the VGG16

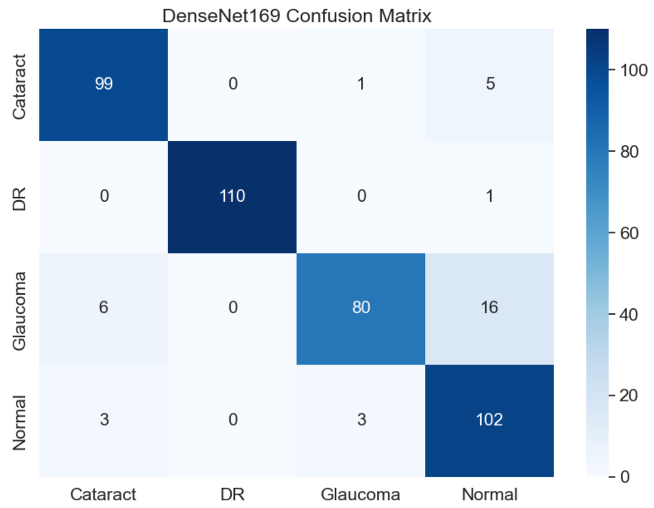


Figure 12. Confusion matrix derived from the classification of diseases with the DenseNet169

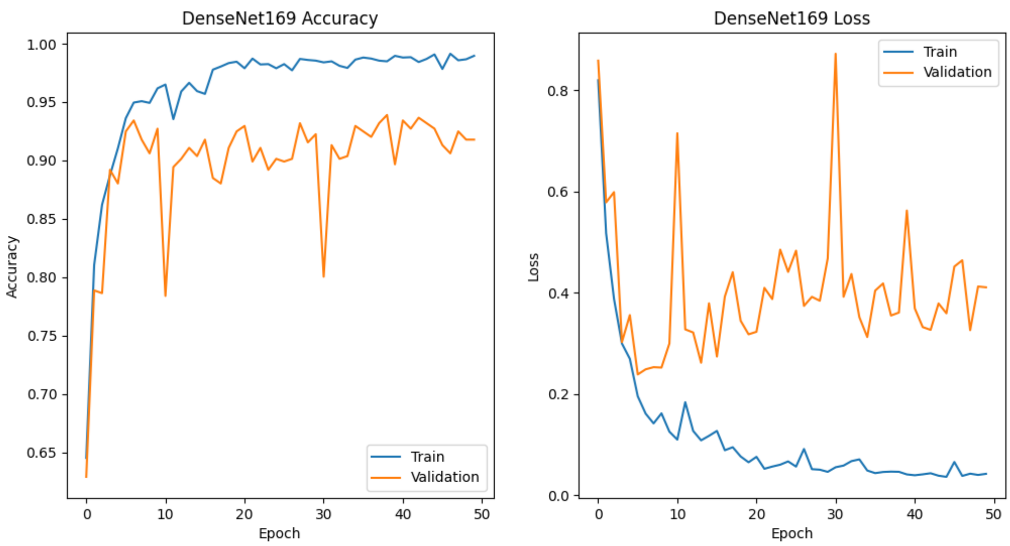


Figure 13. Accuracy and loss graphs of the DenseNet169

Results obtained in this study and literature are given in Table 7. When the values in the table are inspected, the results of this study are better than some other studies in the literature. T. Li et al used various models for DR detection and obtained an accuracy value of 82.84% while S. Alrajjou et al obtained results with an accuracy of 85.78%. The accuracy rates for the four classes obtained from this study is 92.5% with the MobileNet, 91% with the VGG16 and 91.7% with the DenseNet169. In another study carried out by M. N. Bajwa et al, accuracy rates up to 79.67% were obtained for glaucoma diagnosis while we obtained higher accuracy values. When it comes to the studies performed with the same dataset used in this study, D. Jose et al obtained a maximum accuracy value of 87.3% while B. Şener and E. Sümer achieved accuracy rates of 98.47%, 85.85%, and 80.7% with EfficientNetB0, VGG16, and VGG19 models, respectively. Our VGG16 model has an average accuracy rate of 91% outperforming their VGG16 model. This difference may result from using the CLAHE method.

Table 7. Performance comparison with other studies in literature

Reference	Year	Authors	Models	Dataset	Results
[9]	2019	T. Li et al.	VGG16, ResNet-18, GoogLeNet, DenseNet-121, SE-BN Inception	A new dataset consisting of 13673 fundus images of various levels of DR	82.84% accuracy for DR detection
[23]	2022	S. Alrajjou et al.	J48, KNN, Random Forest, Enhanced CNN	Kaggle dataset including a total of 10024 fundus images belonging to healthy and four different grades of DR	Proposed enhanced CNN has achieved an accuracy rate of 85.78% for DR detection
[29]	2019	M. N. Bajwa et al.	CNN with random training and cross validation	ORIGA, HRF, OCT & CFI, DIARETDB1, DRIVE, DRIONS-DB, Messidor	79.67% accuracy with random training and 79.39% with cross validation for glaucoma diagnosis
[44]	2023	D. Jose et al.	LeNet, DenseNet and ManualNet	A total of 4217 fundus images belonging Cataracts, DR, Glaucoma and Healthy classes	87.3% accuracy with LeNet, 59.91% accuracy with DenseNet and 80.7% accuracy with ManualNet
[45]	2023	B. Şener and E. Sümer	EfficientNetB0, VGG16 and VGG19	A total of 4217 fundus images belonging Cataracts, DR, Glaucoma and Healthy classes	Accuracy rates of 98.47% with EfficientNetB0, 85.85% with VGG16 and 92.6% with VGG19 for four different classes
This study			MobileNet, VGG16 and DenseNet169	A total of 4217 fundus images belonging four different classes	92.5% with MobileNet, 91% with VGG16 and 91.7% with DenseNet169 for four different classes

6. Conclusion

In conclusion, this study, it was aimed to compare the performances of MobileNet, VGG16, and DenseNet169 models in classifying fundus images and to determine the best-performed model. The study was conducted with a dataset consisting of 224x224x3 pixels, and 4217 fundus images including images belonging to these four classes: glaucoma, cataracts, DR, and healthy. In the models used, the epoch and batch size values were 50 and 32, respectively. 80% of the dataset was used as training set, 10% as the test set, and the rest for cross-validation. According to the results obtained from this study, the VGG16 and MobileNet models were more performant than the DenseNet169 model by a small margin when it comes to the diagnosis of DR. The MobileNet model was more successful in classification of all four classes than the other two models. The results of the study demonstrate that using DL models such as VGG16, MobileNet, and DenseNet169 with the CLAHE method may help ophthalmologists diagnose cataracts, DR, and glaucoma with lower error rates in less time.

References

[1] Gadekallu, T. R. and Khare, N., et al. 2020. Deep neural networks to predict diabetic retinopathy. Journal of Ambient Intelligence and Humanized Computing. (2020), 1-14.

[2] Wong, T. Y. and Sun, J., et al. 2018. Guidelines on diabetic eye care: the international council of ophthalmology recommendations for screening, follow-up, referral, and treatment based on resource settings. Ophthalmology. 125 (2018), 1608-1622.

[3] Ting, D. S. W. and Pasquale, L. R., et al. 2019. Artificial intelligence and deep learning in ophthalmology. British Journal of Ophthalmology. 103 (2019), 167-175.

[4] Yau, J. W. and Rogers, S. L., et al. 2012. Global prevalence and major risk factors of diabetic retinopathy. Diabetes care. 35 (2012), 556-564.

[5] Tan, G. S. and Cheung, N., et al. 2017. Diabetic macular oedema. The lancet Diabetes & endocrinology. 5 (2017), 143-155.

[6] Beede, E. and Baylor, E., et al. 2020. A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy. Proceedings of the 2020 CHI conference on human factors in computing systems, 1-12.

[7] Mansour, R. F. 2018. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. Biomedical engineering letters. 8 (2018), 41-57.

[8] Gök, A. E., and Taşdemir, Ş. 2023. Santral Seröz Koryoretinopati Hastalığının Derin Öğrenme ile Teşhisi: Derleme Makalesi, Online, 112-122.

[9] Li, T. and Gao, Y., et al. 2019. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. Information Sciences. 501 (2019), 511-522.

[10] Li, J. and Xu, X., et al. 2018. Automatic cataracts diagnosis by image-based interpretability. 2018 IEEE international conference on systems, man, and cybernetics (SMC), 3964-3969.

[11] Xu, X. and Zhang, L., et al. 2019. A hybrid global-local representation CNN model for automatic cataracts grading. IEEE journal of biomedical and health informatics. 24 (2019), 556-567.

[12] Asbell, P. A. and Dualan, I., et al. 2005. Age-related cataracts. The Lancet. 365 (2005), 599-609.

[13] Gao, X. and Lin, S. Wong, T. Y. 2015. Automatic feature learning to grade nuclear cataracts based on deep learning. IEEE Transactions on Biomedical Engineering. 62 (2015), 2693-2701.

[14] Kanski, J. J., and Bowling, B. 2011. Clinical ophthalmology: a systematic approach.

[15] Luo, Y. and Chen, K., et al. 2020. Dehaze of cataractous retinal images using an unpaired generative adversarial network. IEEE Journal of Biomedical and Health Informatics. 24 (2020), 3374-3383.

[16] Pratap, T., and Kokil, P. 2019. Computer-aided diagnosis of cataracts using deep transfer learning. Biomedical Signal Processing and Control. 53 (2019), 101533.

[17] 2022. Eye Diseases Classification. Kaggle, (2022),

[18] Chen, X. and Xu, Y., et al. 2015. Glaucoma detection based on deep convolutional neural network. 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), 715-718.

[19] Chai, Y. and Liu, H. Xu, J. 2018. Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. Knowledge-Based Systems. 161 (2018), 147-156.

[20] Schacknow, P. N., and Samples, J. R. 2010. The glaucoma book: a practical, evidence-based approach to patient care.

[21] Haleem, M. S. and Han, L., et al. 2013. Automatic extraction of retinal features from colour retinal images for glaucoma diagnosis: a review. Computerized medical imaging and graphics. 37 (2013), 581-596.

[22] Tymchenko, B. and Marchenko, P. Spodarets, D. 2020. Deep learning approach to diabetic retinopathy detection. arXiv preprint arXiv:2003.02261. (2020),

[23] Alrajjou, S. and Boahen, E. K., et al. 2022. An enhanced interpretable deep learning approach for diabetic retinopathy detection. 2022 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), 127-135.

[24] Ayala, A. and Ortiz Figueroa, T., et al. 2021. Diabetic retinopathy improved detection using deep learning. Applied Sciences. 11 (2021), 11970.

[25] Acar, E. and Türk, Ö., et al. 2021. Employing deep learning architectures for image-based automatic cataracts diagnosis. Turkish Journal of Electrical Engineering and Computer Sciences. 29 (2021), 2649-2662.

[26] Kalyani, B. and Hemavathi, U., et al. 2023. Smart cataracts detection system with bidirectional LSTM. Soft Computing. (2023), 1-9.

[27] Junayed, M. S. and Islam, M. B., et al. 2021. CataractNet: An automated cataracts detection system using deep learning for fundus images. IEEE Access. 9 (2021), 128799-128808.

[28] Li, F. and Yan, L., et al. 2020. Deep learning-based automated detection of glaucomatous optic neuropathy on color fundus photographs. Graefe's Archive for Clinical and Experimental Ophthalmology. 258 (2020), 851-867.

[29] Bajwa, M. N. and Malik, M. I., et al. 2019. Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. BMC medical informatics and decision making. 19 (2019), 1-16.

[30] Kim, M. and Han, J. C., et al. 2019. Medinoid: computer-aided diagnosis and localization of glaucoma using deep learning. Applied Sciences. 9 (2019), 3064.

[31] Zhang, H. and Niu, K., et al. 2019. Automatic cataracts grading methods based on deep learning. Computer methods and programs in biomedicine. 182 (2019), 104978.

[32] Salvi, M. and Acharya, U. R., et al. 2021. The impact of pre-and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis. Computers in Biology and Medicine. 128 (2021), 104129.

[33] Goodfellow, I. and Bengio, Y. Courville, A. 2016. Deep learning.

[34] Lundervold, A. S., and Lundervold, A. 2019. An overview of deep learning in medical imaging focusing on MRI. Zeitschrift für Medizinische Physik. 29 (2019), 102-127.

[35] Alzubaidi, L. and Zhang, J., et al. 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. Journal of big Data. 8 (2021), 1-74.

[36] Öztürk, B., and Bilgiç, G. 2022. Yenilenebilir Enerji Tahmini İçin Derin Öğrenme Yaklaşımları. Teknobilim-2022: Enerji Krizi ve Yenilenebilir Enerji. (2022), 111.

[37] LeCun, Y. and Bengio, Y. Hinton, G. 2015. Deep learning. nature. 521 (2015), 436-444.

[38] Kübra, U., and Taşdemir, Ş. 2021. Detection and classification of leucocyte types in histological blood tissue images using deep learning approach. Avrupa Bilim ve Teknoloji Dergisi. (2021), 130-137.

[39] Howard, A. G. and Zhu, M., et al. 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861. (2017),

[40] Simonyan, K., and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. (2014),

[41] Sugata, T., and Yang, C. 2017. Leaf App: Leaf recognition with deep convolutional neural networks. IOP Conference Series: Materials Science and Engineering, 012004.

[42] Huang, G. and Liu, Z., et al. 2017. Densely connected convolutional networks. Proceedings of the IEEE conference on computer vision and pattern recognition, 4700-4708.

[43] Bülbül, M.A. 2023Optimization of artificial neural network structure and hyperparameters in hybrid model by genetic algorithm: iOS–android application for breast cancer diagnosis/prediction. The Journal of Supercomputing. (2023), 1-21.

[44] Jose, D. 2023. Classification of EYE Diseases Using Multi-Model CNN. 2023 IEEE 4th Annual Flagship India Council International Subsections Conference (INDISCON), 1-6.

[45] Şener, B., and Sümer, E. Classification of Eye Disease from Retinal Images Using Deep Learning.