Gümüşhane University Journal of Science

GUFBD / *GUJS* (2023) 13(2): 256-269 doi: 10.17714/gumusfenbil.1168842 Research Article

Cataract disease classification from fundus images with transfer learning based deep learning model on two ocular disease datasets

İki göz hastalığı veri seti üzerinde transfer öğrenme tabanlı derin öğrenme modeli ile fundus görüntülerinden katarakt hastalığı sınıflandırması

Halit ÇETİNER[*]

Isparta University of Applied Sciences, Vocational School of Technical Sciences, 32000, Isparta

• Received: 31.08.2022 • Accepted: 20.01.2023

Abstract

Cataract is one of the most serious eye diseases that can blind if left untreated. Detection of the disease in the early stages rather than in the advanced stages can prevent the patient from being blind. At this point, suspected patients should be constantly checked. Continuous control and follow-up of patients is a tiring and laborious process. For the reasons stated, two different deep learning models are proposed in this article that can be used in the diagnosis and detection of cataracts to assist the work and procedures of ophthalmologists. The proposed deep learning models were run on a fundus dataset with normal and cataract symptoms. The proposed deep learning models provide automatic classification of normal and cataract images. Fine-tuning and layer additions were performed on the upper layer using a pre-trained deep learning model called MobileNet V3 Small. A basic MobileNet V3 Small model has also been created to evaluate the performance of the model, which has been enriched by fine-tuning and adding layers to its upper layers. The difference between the proposed model and the basic model is demonstrated by comparing the classification performances of cataract and normal images with accuracy and complexity matrix measurements. According to the best results obtained in the performance comparisons made by separating the training and test data according to the KFold option, the proposed model gave a more successful result graph of 8.26% than the basic model. Finally, the proposed MobileNet V3 model has also been tested on images composed of two different datasets. On average, the proposed MobileNet V3 model on the combined dataset reached 96.62% accuracy.

Keywords: Cataract, Deep learning, MobileNet V3 Small, Transfer learning

Öz

Katarakt, tedavi edilmediği takdirde kör edebilen en ciddi göz hastalıklarından biridir. Hastalığın ileri aşamalarından ziyade erken aşamada tespit edilmesi hastanın kör olmasının önüne geçebilmektedir. Bu noktada şüphe duyulan hastaların sürekli olarak kontrol edilmesi gerekmektedir. Sürekli olarak hastaların kontrol ve takip edilmesi ise yorucu ve zahmetli bir işlemdir. Belirtilen sebeplerden dolayı bu makalede katarakt tanı ve tespitinde kullanılabilecek göz doktorlarının yaptıkları iş ve işlemlere yardımcı iki farklı derin öğrenme modeli önerilmiştir. Önerilen derin öğrenme modelleri normal ve katarakt belirtilerine sahip fundus veri seti üzerinde çalıştırılmıştır. Önerilen derin öğrenme modelleri normal ve kataraktlı görüntülerin otomatik olarak sınıflandırmasını sağlamaktadır. MobileNet V3 Small adlı önceden eğitilmiş derin öğrenme modeli kullanılarak üst katmanda ince ayarlamalar ve katman eklemeleri gerçekleştirilmiştir. Üst katmanlarında ince ayarlamalar ve katman eklemeleri yapılarak zenginleştirilen modelin performans değerlendirmesini yapabilmek için temel bir MobileNet V3 Small modeli de oluşturulmuştur. Önerilen model ile temel model arasındaki fark katarakt ve normal görüntülerin sınıflandırma performanslarını karşılaştırılarak doğruluk ve karmaşıklık matris ölçümleri ile gösterilmiştir. K Katlı seçeneğine göre eğitim ve test verileri ayrılarak yapılan performans karşılaştırmalarında elde edilen en iyi sonuçlara göre önerilen model, temel modelden 8.26% daha başarılı bir sonuç grafiği vermiştir. Son olarak, önerilen MobileNet V3 modeli, iki farklı veri setinin birleşmesinden oluşan görüntüler üzerinde de test edilmiştir. Ortalama olarak birleştirilmiş veri seti üzerinde önerilen MobileNet V3 modeli ile %96.62 doğruluk oranına ulaşmıştır.

***Anahtar kelimeler**: Katarakt, Derin öğrenme, MobileNet V3 Small, Transfer öğrenme*

[*] Halit ÇETİNER; halitcetiner@isparta.edu.tr

1. Introduction

The opacity and turbidity of the inner lens of the eye are among the symptoms of cataract disease. At the same time, blindness can be experienced in cataract treatments where medical intervention is delayed (B. R. Kumar & Shimna, 2017). Ophthalmology is the field that deals with any eye disease such as cataracts (Grewal et al., 2018). Cataract disease, which is the most common cause of blindness worldwide, is examined and analyzed by ophthalmologists. Ophthalmologists directly or indirectly visualize the eye and its surroundings to help diagnose and diagnose any eye disease such as cataracts. They aim to identify lesions with pattern recognition after visualization. Diagnostic technologies used in the field of ophthalmology are very suitable for the application of computer vision techniques through digital display. In this article, the newest and rapidly spreading deep learning models in ophthalmology and machine learning are combined. Deep learning methods have adapted very quickly to the field of ophthalmology (Lee et al., 2017). For the stated reasons, a model that helps ophthalmologists has been developed by applying deep learning technologies together with computer vision techniques.

Ophthalmologists are struggling to control the ever-increasing number of cataract patients. Symptoms of cataract disease include opacity. Although the opacity is in the form of a point in the initial stage, it covers the entire lens in the later stages (Zhang et al., 2017). As a result of this, blurring and hazy images occur in the images with light scattering. This makes it difficult to use any vehicle (Fraser et al., 2013). The presence of cataract blocks lens light, causing it to not reach the lens (Junayed et al., 2021). It is a serious disease that does not prevent vision in the initial stage, but causes vision loss especially over the age of 40 (Rana & Galib, 2017). Early treatment of cataract disease, instead of costly and painful surgeries, closes the road to blindness (Pascolini & Mariotti, 2012). It is estimated that the number of people with severe visual impairment due to cataracts worldwide will exceed 40 million by 2025 (Cao et al., 2020). The number of people with severe visual impairment due to cataracts is estimated to exceed 40 million worldwide by 2025 (Cao et al., 2020). One third of visually impaired cases occur due to cataract disease (Organization, 1998). It is a serious disease that does not prevent vision in the initial stage, but causes vision loss especially over the age of 40 (Rana & Galib, 2017). The presence of cataract blocks lens light, causing it to not reach the lens (Junayed et al., 2021). As a result of this, with the light scattering, it becomes difficult to drive in any way (Fraser et al., 2013). For those who have cataracts and are over the age of 40, driving a vehicle becomes very risky. Early treatment of cataract disease paves the way for blindness instead of costly and painful surgeries (Pascolini & Mariotti, 2012). However, ophthalmologists are unable to control the ever-increasing number of cataract patients. At this point, technological tools are needed to assist ophthalmologists.

Retinal diseases such as cataract, trachoma, diabetic retinopathy, corneal opacity, glaucoma can cause blindness. Among the diseases mentioned, cataract is one of the most common eye disorders that causes vision distortion (Junayed et al., 2021). The types of cataracts of nuclear cataract (Hu et al., 2020), cortical cataract (Hu et al., 2020) and posterior subcapsular cataract (Yang et al., 2016) are divided into three main groups according to the location and region of development. The main causes of the different types of cataracts mentioned are thought to be caused by environmental factors such as diabetes, smoking and aging (Yang et al., 2016). In order to prevent blindness, the detection of cataract in an early stage with rapid and risk-free methods is of vital importance in treatment. Cataract detection is a challenging problem due to different reasons such as cataract scale, shape and location, patient's age, gender, eye type. In recent years, different studies have been carried out for automatic cataract detection. Automatic cataract detection and classification can be performed using ultrasound, slit lamp, retro-illumination, or fundus images (Junayed et al., 2021). Fundus images, which are widely used by experts among these cataract imaging systems, are of great interest in this field (Raju et al., 2016). Imaging systems such as slit lamps other than fundus images can only be used by experienced ophthalmologists. For this reason, the scarcity of ophthalmologists, especially in underdeveloped or developing countries, delays the treatment process (Yang et al., 2016). There is a strong need for an automatic cataract detection system based on fundus images to assist ophthalmologists in the delayed treatment process.

Although many hospitals have traditional eye imaging systems, the diagnosis may be mistaken due to reasons such as user error or sensor failure (Wang et al., 2021). There is a trend towards computer-aided diagnostic systems to solve the problems arising from the shortage of ophthalmologists and the inadequacies of different equipment (Doi, 2007; Gao et al., 2015; Liu et al., 2017). There are artificial intelligence supported systems based on different basic features such as discrete cosine transformation (DCT) (Fan et al., 2015), deep features, local features (Xiong et al., 2017) for cataract classification. In the artificial intelligence supported systems in the literature, limitations and computational costs are high due to the low test accuracies and the number of model parameters in general. At the same time, it is seen that methods based on classical machine learning are generally used in the diagnosis of cataracts (Fan et al., 2015; Manchalwar & Warhade, 2017; Qiao et al., 2017; Xiong et al., 2017). In classical machine learning methods, unlike automatic feature extraction such as deep learning, manual feature extraction is performed. Instead of using a method with low discrimination, a new method with high accuracy is needed for cataract classification. A method with sufficient competence to assist ophthalmologists will contribute to the literature. Flaxman et al. state that automatic control of fundus images of cataract diseases with artificial intelligence supported software is important in terms of ophthalmology (Flaxman et al., 2017). Ophthalmologists can diagnose and detect cataracts by examining the fundus images of the right and left eyes separately. This process needs to be facilitated by a deep learning-based system. As a result of facilitation, the material and moral costs caused by cataract disease will be avoided. Automated systems are needed to reduce workloads so that ophthalmologists can fully catch up with patients. There is a significant spread of diagnosing diseases with the help of remote health services (Bakator & Radosav, 2018; Ertuğrul et al., 2021).

Transfer learning and CNN-based methods have been compared in different ways in the literature. Imran et al. stated that transfer learning-based methods tend to be more subjectivity than CNN-based methods. They declare that transfer learning-based methods can be trained on large datasets such as ImageNet, with fine-tuning of hyperparameters (Imran et al., 2021). The learned features can be transferred with the transfer learning method. In transfer learning-based methods, some of the model weights can be kept constant in order to be used in target tag estimation, and some can be updated and used in target tag estimation. As a result of these processes, adjustable fast and effective cataract classification and prediction processes can be performed using large data sets such as ImageNet.

The contributions of the article prepared with the specified scopes to the literature are presented below.

- Classification of images of normal and cataract disease is provided by using a deep learning model trained using large datasets such as ImageNet.
- A deep learning model based on MobileNet V3 Small is presented for the classification of images of cataracts and normal disease.
- A basic MobileNet V3 Small model was created to compare the performance of the presented MobileNet V3 Small based model.
- Training and test data are separated according to the KFold technique to compare the performance of the proposed and basic model.
- An average of 96.62% accuracy was obtained in the data set combined with the proposed MobileNet V3 model. This rate is important in that it is 0.62% higher than the data set obtained in a single data set. The data set has been expanded to improve the performance criteria of the proposed deep learning model.

The rest of the article consists of 3 sections. In the first section, detailed information is given about the data sets in which the performance tests of deep learning models are made. In the second section, the performance results obtained from different models are presented comparatively. In the last section, the study is concluded.

2. Material and methods

2.1. Material

In this study, two different models were defined. The main ones are MobileNet V3, proposed MobileNet V3. A dataset containing normal and cataract images (Ocular Disease Recognition, 2021) was used to test the basic and proposed MobileNet V3 model. The images in the dataset were obtained from 5,000 different people. The images used include different diseases in addition to cataract disease. These diseases are expressed as diabetes, glaucoma, hypertension, myopia, amd, and other abnormalities. In the article study, only normal images with cataracts and not related to any disease were selected from these diseases. There are 1088 images in total in the data set created by selection. 500 of the selected 1088 images represent normal images, while the rest represent images of the eyes with cataracts. Triyadi et al. using the data set (Ocular Disease Recognition, 2021) in their study, they mention 300 normal images and 313 images with cataract.

When these images are evaluated in terms of two different eyes, right and left, it is reported that there are 500 normal images and 594 cataracted images in the data set (Triyadi et al., 2022). In different sources in the literature (Ocular Disease Recognition, 2021), the number of data in the data set is expressed in different numbers (Khan et al., 2022). Experimental studies, if the "full_df.csv" file in the data set is analyzed as left cataract and right cataract images, the correct number of normal and cataract images can be reached. While the number of images with left cataract is 304, the number of images with right cataract is 290. 6 out of a total of 594 images are not used due to image quality. A total of 588 cataract images are used. On the other hand, the proposed MobileNet V3 model was also tested on a data set consisting of combining both the (Ocular Disease Recognition, 2021) data set and the (Chen, 2022) data set. The (Chen, 2022) data set was combined with the other (Ocular Disease Recognition, 2021) data set to measure the effect of the proposed model on performance. The (Chen, 2022) dataset includes normal, cataract, glaucoma, retinal disease type images. In the (Chen, 2022) data set, there are 300, 100, 100, 100 images of normal, cataract, glaucoma, retinal disease type images, respectively. Among them, only normal and cataract images were used in performance tests.

When testing model performance using the images in the data set, training and test data are separated according to the KFold technique. Image examples showing normal and cataract images in the data set are given in Figure 1. Images in Figure 1 are images with indexes 1011, 309, 523, 1008, 461, 69, 244, 993, 104 and 242 from top left to bottom right corner. In Figure 1, the image with index 1011 represents the image with cataract, while the image with index 1008 represents the normal image. mages with index 1011, 309, 523, 461, 69, 104, 242, respectively, are images with cataracts. The images with index 1008, 244, 993 represent normal images. The images used in Figure 1 belong to the (Ocular Disease Recognition, 2021) dataset.

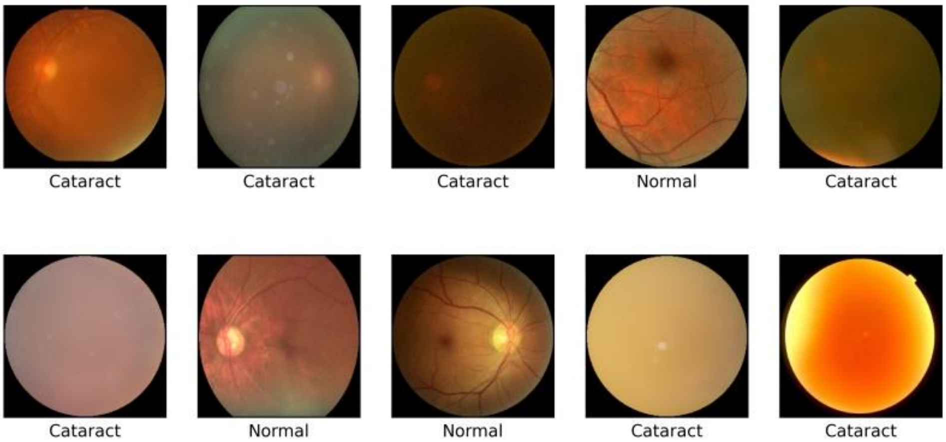

Figure 1. Examples of normal and cataract images in the data set used

2.2. MobileNet V3 Small

With the recent increase in portable devices, there is a great interest in transfer learning methods such as ImageNet that are pre-trained with millions of data. One of the architectures where this interest is concentrated is the MobileNet architectures. There are three different versions of MobileNet architectures, MobileNet V1 (A. G. Howard et al., 2017), MobileNet V2 (Sandler et al., 2018) and MobileNet V3. The MobileNet V2 architecture was developed by changing the bottleneck structure on MobileNet V1, which was developed as the first architecture of the version series (Sandler et al., 2018). There are two different versions of the MobileNet V3 architecture, small and large. The development of the proposed and basic model for the classification of cataract and normal images is based on the MobileNet V3 Small architecture. This architecture works in an optimized way with the help of NetAdapt and NAS search algorithms. In MobileNet V3 architectures, the hard swish function (h-swish) activation function is used instead of the ReLU activation function, which is predominantly used in the MobileNet V2 version (Qian et al., 2021). Like the ReLU activation function, the swish activation function is also an activation function. Swish activation function includes sigmoid function. The sigmoid function is used in the classification layers of deep learning models. This function is actively used in estimation, binary classification and logistic regression models (Mercioni &

Holban, 2020). Sigmoid is also used to calculate swish activation functions. The mathematical equation of the swish function is given in Equation 1.

$$swish(x) = x\sigma(x) \ = x\frac{1}{1+e^{-x}} \tag{1}$$

In Equation 1, where the sigmoid activation function is used, the computational cost is high. In order to reduce this computational cost, the sigmoid function has been converted to a h-swish (Avenash & Viswanath, 2019). This transformation is shown in Equation 2.

$$h - swish(x) = x\frac{ReLU6(x+3)}{6} \tag{2}$$

In Equation 2, which was developed to reduce the computational cost, the ReLU6 function was used instead of the sigmoid activation function. The difference in Equation 2 is using the ReLU6 function rather than a custom constant. $h - swish$ helps with network regulation as it solves the saturation problem of output neurons (Roy et al., 2022). Details of Equation 1 and Equation 2 equations are presented in the study (Howard et al., 2019). In the MobileNet V2 version, features are extracted from a high-dimensional space with minimal loss with linear bottleneck and inverted residual features. Linear bottleneck acts as a layer in the MobileNet V2 architecture by reducing the input size. These layered structures are also seen to be preserved in the MobileNet V3 architecture (Qian et al., 2021). Inverted residual block structures reduce memory costs in MobileNet V2 architectures. There are also shortcuts between bottleneck structures to avoid obstacles such as gradient loss and bursts. The MobileNet V3 architecture, on the other hand, takes the features of the MobileNet V2 architecture forward with the help of Nas and NetAdapt optimization based network exploration methods. An effective model is created by using the h-swish activation function instead of ReLU at some points of the MobileNet V3 architecture.

2.3. Proposed Model

Figure 2 shows the proposed deep learning model for automatic classification of cataract images. The proposed model consists of 14 steps. In the first step, normal and cataract images are presented as input. In the second step, fine adjustments were made on the upper layers of the basic MobileNet V3 Small architectural layers without using the weight values. In the third step, a convolution layer with 256 3x3 windows was added. In the fourth step, the maximum pooling layer with a 3x3 window was added. In the fifth step, a two-dimensional convolution layer with 256 filters was added, similar to the third step. In the sixth step, similar to the fourth step, although maximum pooling was used, a window size of 2x2 was used as the window size.

In the seventh layer, a dropout layer was applied that performs 0.8 neuron dropout. With this layer, memorization of the model is prevented. In the eighth layer, by smoothing the model, the created feature size of the model was reshaped. In the ninth step, the batch normalization layer was added, which normalizes the inputs between the layers. In the tenth step, a dense layer with 512 neurons was added. In the eleventh step, a second dense layer with 256 neurons was added. In the twelfth layer, a fully connected layer with swish activation function has been applied. In the thirteenth layer, classification layer with softmax activation function is applied. In the last step, the fourteenth step, the class estimation was performed and the process steps of the model were completed. It is also important to evaluate the performance of the model obtained as a result of these steps.

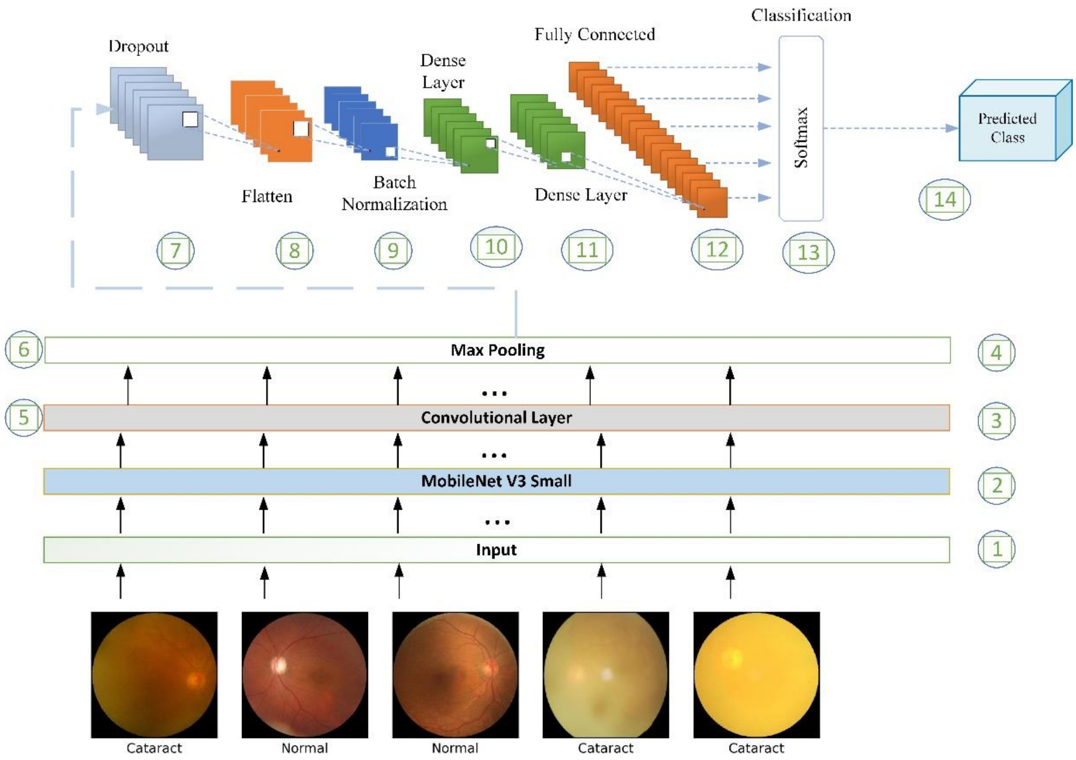

Figure 2. Proposed deep learning model for cataract classification

The proposed model, which was created as a result of the combination of the steps, and the model called the basic MobileNet V3 model were compared according to performance metrics. The basic MobileNet V3 model is the model created in the upper or lower layers of the MobileNet V3 architecture without adding or removing any layers. The structure of the proposed model has emerged by determining the different layers in addition to the basic model by successive experimental studies. Formulas defined in Equations 3-6 were used to compare the proposed model with the basic MobileNet V3 model. The formulas in Equations 3-6 are measurement metrics used in the literature in general terms (Goutte & Gaussier, 2005).

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 = 2x\frac{PrecisionxRecall}{Precision + Recall} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

TP, FP, TN and FN in Equations 3-6 represent correctly predicted cataract, incorrectly predicted cataract, correctly predicted normal image and incorrectly predicted normal image, respectively. Recall, F1, Precision and accuracy values, which can be used comparatively in the literature, are calculated for the specified measurements.

3. Results and discussion

Performance measurements of the proposed model and the basic MobileNet V3 model were obtained using the equation connections between Equations 3-6. The proposed model achieved a test accuracy rate of 97%, while the basic MobileNet V3 Small model achieved a test accuracy rate of 89%. The difference between the proposed model and the basic model is 8%. In experimental studies, 20 epoch values were defined in the training of models. The training and testing rate was determined to be 80% and 20%, respectively. Training and test data of cataract and normal images in the dataset were separated according to the KFold technique.

Table 1. Model performance results according to KFold options

KFold	Model	Train Accuracy	Test Accuracy
1	MobileNet V3 Small model	0.78	0.79
2		0.86	0.89
3		0.90	0.95
4		0.78	0.87
5		0.64	0.61
Average		0.79	0.82
1	Proposed model	0.99	0.96
2		0.99	0.95
3		0.99	0.97
4		0.99	0.94
5		0.99	0.97
Average		0.99	96.0

With the KFold 5 technique, 5 different results were obtained. The KFold group with the most successful test result was preferred. Performance results of the basic MobileNet V3 Small model are plotted based on the KFold 2 option, while the performance plots of the proposed model are drawn based on the KFold 3 option. The graphs of the accuracy and loss rate obtained according to the selected models are shown in Figure 3.

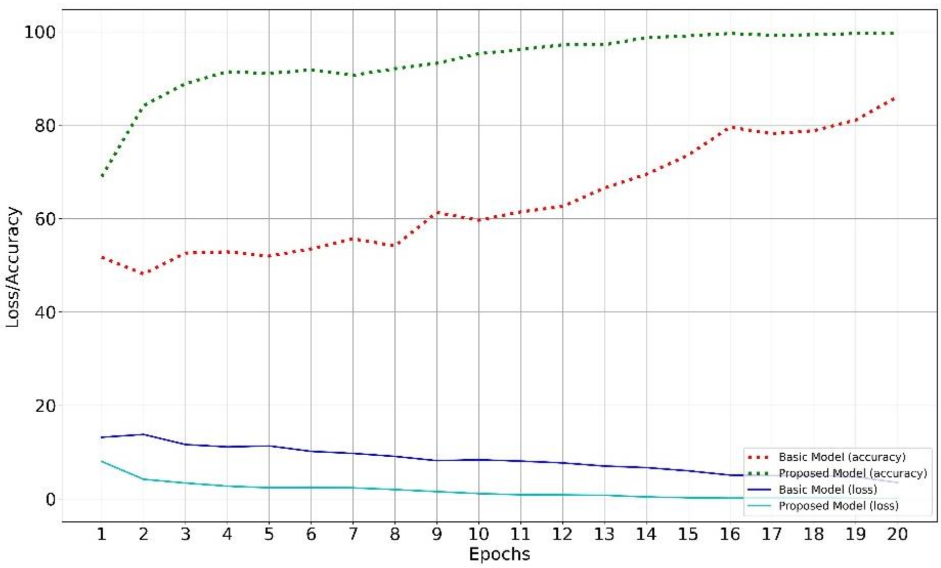

Figure 3. Training accuracy and loss performance of the proposed and basic MobileNet V3 models

According to the data from Figure 3, the loss value of the proposed model is reduced to 0.012%. According to the same chart, the basic MobileNet V3 Small model training loss rate reaches 0.35%. There is a difference of 0.33% between the basic model and the proposed model. The difference between the loss values of both models is very large. According to the results given in Table 1, the proposed model reached the highest 97% accuracy rate in terms of test accuracy, while the basic MobileNet V3 Small model reached 95%. Compared to the KFold option with the highest accuracy rate, the difference between models is around 2%.

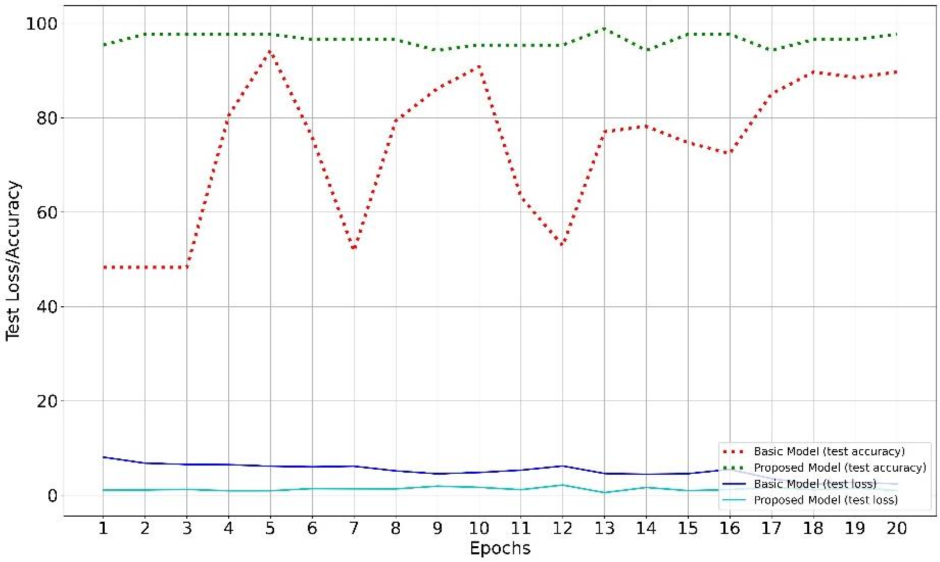

Figure 4. Validation accuracy and loss performance of basic MobileNet V3 and proposed models

According to Figure 4 data, the proposed model and basic model validation accuracy rates were 97.71% and 89.45%, respectively. Proposed model and basic MobileNet V3 Small validation loss rates decrease to 0.11% and 0.240%, respectively.

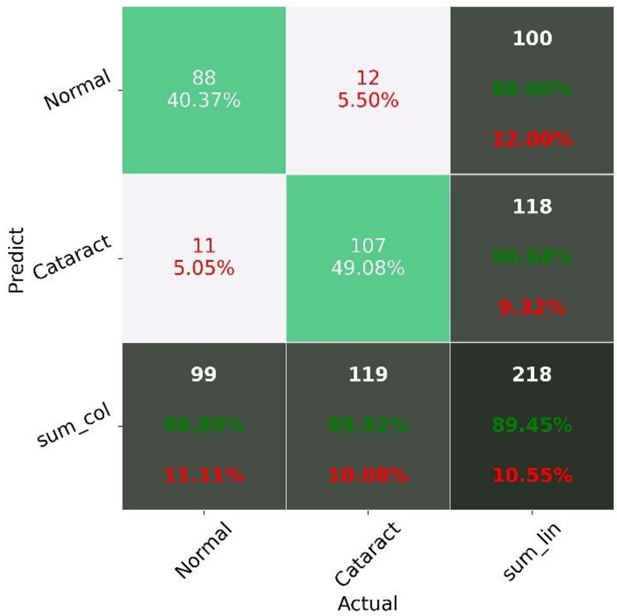

Figure 5. The confusion matrix of the basic MobileNet V3 Small model

Data from the confusion matrix of the basic MobileNet V3 Small model is presented in Figure 5. According to the basic model, the cataract class is predicted with a 90.68% accuracy rate, while the normal class is predicted with an 88% accuracy rate. According to the basic model, 12 errors and 88 correct determinations were made from the base class. According to the basic model, 107 correct detections and 11 incorrect detections were made in the cataract class.

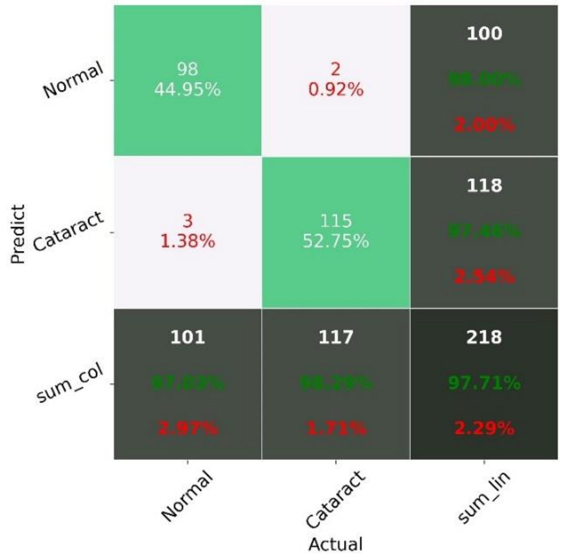

Figure 6. The confusion matrix of proposed model

The confusion matrix data of the proposed model is presented in Figure 6. According to the proposed model, the cataract class is predicted with an accuracy rate of 97.46%, while the normal class is estimated with a 98% accuracy rate. According to the proposed model, 2 errors and 98 correct detections were made from the normal class. According to the same proposed model, 115 correct detections and 3 incorrect detections were made in the cataract class. On average, an accuracy rate of 97.71% was achieved.

Table 2. Deep learning models performance results

Models	Type	Precision	Recall	F1 score	Accuracy
Proposed Model	Normal	0.97	0.98	0.97	0.97
	Cataract	0.98	0.97	0.98	
Basic MobileNet V3 Small	Normal	0.92	0.87	0.90	0.89
	Cataract	0.89	0.93	0.90	

In addition to the confusion matrix results given in Figure 5 and Figure 6 of the proposed model and the basic MobileNet V3 Small model in Table 2, precision, recall, F1 score, accuracy values are presented collectively. According to the presented data, it is seen that the proposed model is ahead in terms of performance not only in accuracy values, but also in other precision, recall, and F1 score values.

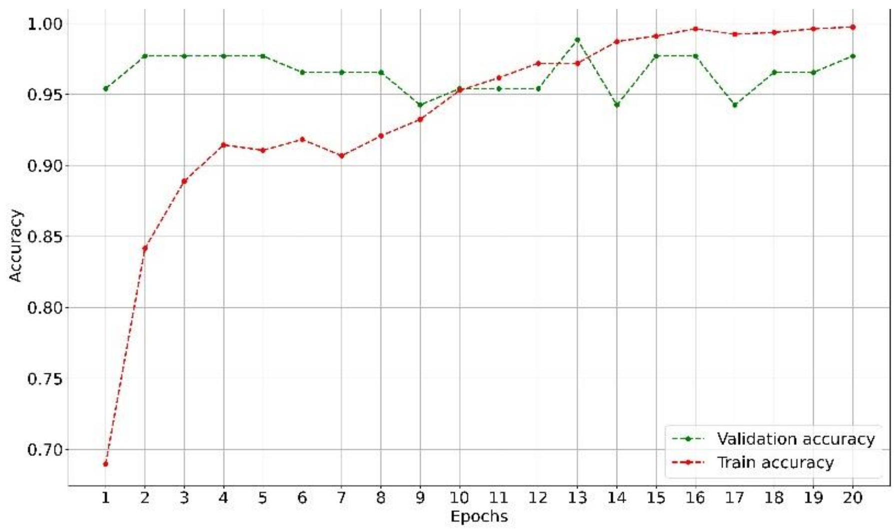

Figure 7. The accuracy graph of the proposed MobileNet V3 model tested on the combined dataset

In Figure 7, train accuracy and validation accuracy graphs of the proposed MobileNet V3 model in the extended data set are presented. Train accuracy graph reaches 98.31% accuracy at the 20th epoch. Validation accuracy rate reached 96.62% accuracy rate.

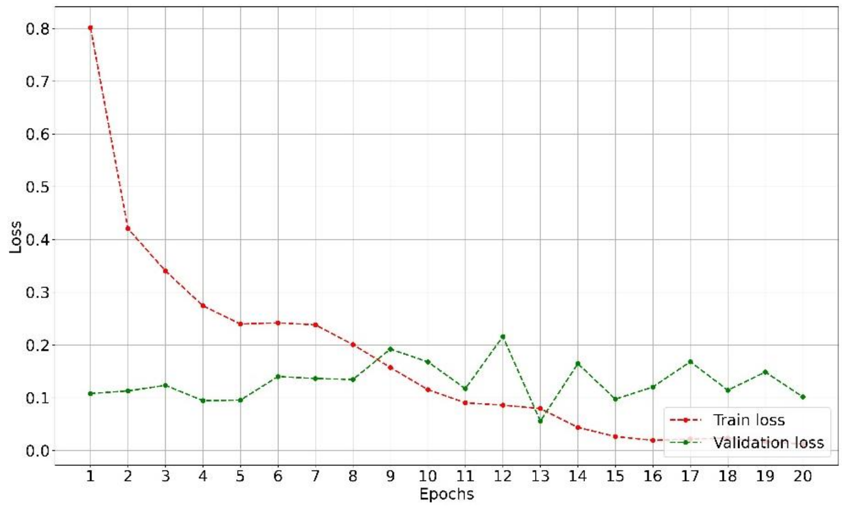

Figure 8. The loss graph of the proposed MobileNet V3 model tested on the combined dataset

Figure 8 shows the loss graphs obtained on the combined data set. Validation loss values decrease up to 0.12% at the 20th epoch. Train loss values decrease up to 0.05% at the 20th epoch.

4. Discussion

Apart from this article, a DenseNet201-based deep learning model has been developed for cataract detection. In this study, which was developed differently from the DenseNet201 model study (Çetiner & Çetiner, 2022), an architectural model that can work on light devices was preferred. With the model based on the MobileNet V3 model, the classification of cataract disease has been carried out with high accuracy. In addition to these, the data set used in the study (Çetiner & Çetiner, 2022) was expanded. Only the (Ocular Disease Recognition, 2021) data set was used in the previously proposed DenseNet201 model. The proposed MobileNet V3 model in this study has not been tested only on the (Ocular Disease Recognition, 2021) dataset. The proposed MobileNet V3 model was also tested on the combined version of the (Ocular Disease Recognition, 2021) and (Chen, 2022) dataset. Despite the expansion of the data set, there was no decrease in the accuracy rate. Compared to the DenseNet201 architecture and the MobileNet V3 architecture, the DenseNet201 architecture is both older and has more layers. Working performance measures (Çetiner & Çetiner, 2022), developed based on DenseNet201 architecture, are 3% more accuracy than this study, but they are not lighter. Today, not only applications that work in desktop environments, but also applications that can easily work in mobile environments are preferred. For the reasons stated, this article was carried out as a separate study.

In Table 3 data, the results of the proposed model test performance were compared with studies in the literature. The data set used by the studies in the literature was used when making comparisons. As a result of performance comparisons, the accuracy, F1 score, recall and precision values are listed in Table 3. According to these listed data, there is a 7.5% difference in accuracy and 16% in precision compared to the (Jayachitra et al., 2021) study in determining the proposed model cataract class. The proposed model has a difference of 4.43% in accuracy compared to the (K S et al., 2021) study in detecting the cataract class. According to studies in the literature (Jayachitra et al., 2021) and (K S et al., 2021), the proposed model gave a good result in determining the cataract class. In normal class determination, on the other hand, 7.5% difference in accuracy and 15% difference in precision compared to the proposed model (Jayachitra et al., 2021) study. There was a 4.43% difference in normal class detection compared to another study in the literature (K S et al., 2021). (Y. Kumar & Gupta, 2022) achieved 92.6% accuracy in detecting cataracts in the Inception V3 model. In normal images, they reached 96.4% accuracy. (Y. Kumar & Gupta, 2022) achieved 90.9% cataract classification accuracy on the DenseNet 121 model. In the normal image, they reached 98.9%

accuracy. The accuracy rate they achieve in the classification of normal images is better than the classification success they achieve in images with cataracts.

Table 3. Comparison results with similar data sets

Class	Model	Accuracy (%)	F1 score (%)	Recall (%)	Precision (%)
Cataract	(Jayachitra et al., 2021)	89.5	--	--	82
	(K S et al., 2021)	92.56	--	--	--
	(Y. Kumar & Gupta, 2022)'s Inception V3	92.6	--	--	--
	(Y. Kumar & Gupta, 2022)'s DenseNet 121	90.9	--	--	--
	Proposed	97.0	98.0	97.00	98.00
Normal	(Jayachitra et al., 2021)	89.50	--	--	82
	(K S et al., 2021)	92.56	--	--	--
	(Y. Kumar & Gupta, 2022)'s Inception V3	96.4	--	--	--
	(Y. Kumar & Gupta, 2022)'s DenseNet 121	98.9	--	--	--
	Proposed	97.0	97.0	98.00	97.00

5. Conclusions

Recently, deep learning-based studies have been carried out for automatic classification of eye diseases. The main purpose of these studies is to increase the processing speed by reducing the number of ophthalmologists. The main benefit stated will be a reduction in workload in public health institutions and hospitals. Recent studies in the literature support that deep learning-based studies show higher performance in classification of ophthalmological diseases compared to classical classifiers. A deep learning model based on MobileNet V3 Small is presented for the automatic classification of cataract disease, which is defined as a serious eye disease and can cause blindness, with the stated motivations. According to the results given, the proposed model is a very useful model for the literature because it can be used in portable devices and is light. In the classification of the cataract class, which was obtained as a performance criterion, accuracy, F1 score, recall, precision values were 97%, 98%, 97%, 98%. In the classification of non-Cataract normal images, accuracy rates of 97%, 97%, 98%, and 97% were achieved. In future studies, innovative approaches for real-time detection of cataract regions can be created.

Author contribution

The entire article was written by the corresponding author.

Declaration of ethical code

The authors of this article declare that the materials and methods used in this study do not require ethical committee approval and/or legal-specific permission.

Conflicts of interest

The authors declare that there is no conflict of interest.

References

Avenash, R., & Viswanath, P. (2019). Semantic segmentation of satellite images using a modified CNN with hard-swish activation function. *14th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, VISIGRAPP 2019* (pp. 413–420), Czech Republic.

Bakator, M., & Radosav, D. (2018). Deep learning and medical diagnosis: a review of literature. In *Multimodal Technologies and Interaction, 2*(3). https://doi.org/10.3390/mti2030047.

Cao, L., Li, H., Zhang, Y., Zhang, L., & Xu, L. (2020). Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. *Information Fusion*, *53*, 196–208.

Çetiner, H., & Çetiner, İ. (2022). Classification of cataract disease with a Densenet201 based deep learning model. *Journal*, *12*(3), 1264–1276. https://doi.org/10.21597/jist.1098718.

Chen, Y. (2022). *Retina Dataset*. https://github.com/yiweichen04/retina_dataset.

Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, *31*(4–5), 198–211.

Ertuğrul, Ö. F., Acar, E., Aldemir, E., & Öztekin, A. (2021). Automatic diagnosis of cardiovascular disorders by sub images of the ECG signal using multi-feature extraction methods and randomized neural network. *Biomedical Signal Processing and Control*, *64*, 102260. https://doi.org/10.1016/j.bspc.2020.102260.

Fan, W., Shen, R., Zhang, Q., Yang, J.-J., & Li, J. (2015). Principal component analysis based cataract grading and classification. *2015 17th International Conference on E-Health Networking, Application & Services (HealthCom)* (pp. 459–462), USA. https://doi.org/10.1109/HealthCom.2015.7454545.

Flaxman, S. R., Bourne, R. R. A., Resnikoff, S., Ackland, P., Braithwaite, T., Cicinelli, M. V, Das, A., Jonas, J. B., Keeffe, J., & Kempen, J. H. (2017). Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *The Lancet Global Health*, *5*(12), e1221–e1234.

Fraser, M. L., Meuleners, L. B., Lee, A. H., Ng, J. Q., & Morlet, N. (2013). Vision, quality of life and depressive symptoms after first eye cataract surgery. *Psychogeriatrics*, *13*(4), 237–243.

Gao, X., Lin, S., & Wong, T. Y. (2015). Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering*, *62*(11), 2693–2701.

Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *Lecture Notes in Computer Science*, *3408*, 345–359. https://doi.org/10.1007/978-3-540-31865-1_25.

Grewal, P. S., Oloumi, F., Rubin, U., & Tennant, M. T. S. (2018). Deep learning in ophthalmology: a review. *Canadian Journal of Ophthalmology*, *53*(4), 309–313. https://doi.org/10.1016/j.jcjo.2018.04.019.

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: efficient convolutional neural networks for mobile vision applications. *ArXiv Preprint ArXiv:1704.04861*.

Howard, A., Sandler, M., Chu, G., Chen, L.-C., Chen, B., Tan, M., Wang, W., Zhu, Y., Pang, R., & Vasudevan, V. (2019). Searching for Mobilenetv3. *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 1314–1324), Korea.

Hu, S., Wang, X., Wu, H., Luan, X., Qi, P., Lin, Y., He, X., & He, W. (2020). Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images. *IEEE Access*, *8*, 174169–174178.

Imran, A., Li, J., Pei, Y., Akhtar, F., Mahmood, T., & Zhang, L. (2021). Fundus image-based cataract classification using a hybrid convolutional and recurrent neural network. *The Visual Computer*, *37*(8), 2407–2417. https://doi.org/10.1007/s00371-020-01994-3.

Jayachitra, S., Nitheesh Kanna, K., Pavithra, G., & Ranjeetha, T. (2021). A novel eye cataract diagnosis and classification using deep neural network. *Journal of Physics: Conference Series*, *1937*(1). https://doi.org/10.1088/1742-6596/1937/1/012053.

Junayed, M. S., Islam, M. B., Sadeghzadeh, A., & Rahman, S. (2021). CataractNet: an automated cataract detection system using deep learning for fundus images. *IEEE Access*, *9*, 128799–128808. https://doi.org/10.1109/ACCESS.2021.3112938.

K S, Y., Mithra, N. M., KS, V., & R, K. (2021). Classification of diabetic retinopathy through identification of diagnostic keywords. *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 716–721), India. https://doi.org/10.1109/ICIRCA51532.2021.9544621.

Khan, M. S., Tafshir, N., Alam, K. N., Dhruba, A. R., Khan, M. M., Albraikan, A. A., & Almalki, F. A. (2022). Deep learning for ocular disease recognition: an inner-class balance. *Computational Intelligence and Neuroscience*, *2022*.

Kumar, B. R., & Shimna, M. P. (2017). Recent approaches for automatic cataract detection analysis using image processing. *Journal of Network Communications and Emerging Technologies (JNCET)*, *7*(10).

Kumar, Y., & Gupta, S. (2022). Deep transfer learning approaches to predict glaucoma, cataract, choroidal neovascularization, diabetic macular edema, drusen and healthy eyes: an experimental review. *Archives of Computational Methods in Engineering*. https://doi.org/10.1007/s11831-022-09807-7.

Lee, A., Taylor, P., Kalpathy-Cramer, J., & Tufail, A. (2017). Machine learning has arrived! *Ophthalmology*, *124*(12), 1726–1728. https://doi.org/10.1016/j.ophtha.2017.08.046.

Liu, X., Jiang, J., Zhang, K., Long, E., Cui, J., Zhu, M., An, Y., Zhang, J., Liu, Z., & Lin, Z. (2017). Localization and diagnosis framework for pediatric cataracts based on slit-lamp images using deep features of a convolutional neural network. *PloS One*, *12*(3), e0168606.

Manchalwar, M., & Warhade, K. (2017). Detection of cataract and conjunctivitis disease using histogram of oriented gradient. *International Journal of Engineering and Technology*, *9*, 2400–2406. https://doi.org/10.21817/ijet/2017/v9i3/1709030214.

Mercioni, M. A., & Holban, S. (2020). P-Swish: activation function with learnable parameters based on swish activation function in deep learning. *2020 International Symposium on Electronics and Telecommunications (ISETC)* (pp. 1–4), Romania. https://doi.org/10.1109/ISETC50328.2020.9301059.

Ocular Disease Recognition. (2021). Senior data scientist at hospital israelita Albert Einstein São Paulo, State of São Paulo, Brazil.

Organization, W. H. (1998). *The World health report: 1998: life in the 21st century: a vision for all: executive summary*. World Health Organization.

Pascolini, D., & Mariotti, S. P. (2012). Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, *96*(5), 614–618.

Qian, S., Ning, C., & Hu, Y. (2021). MobileNetV3 for image classification. *2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)* (pp. 490–497), China. https://doi.org/10.1109/ICBAIE52039.2021.9389905.

Qiao, Z., Zhang, Q., Dong, Y., & Yang, J. (2017). Application of SVM based on genetic algorithm in classification of cataract fundus images. *2017 IEEE International Conference on Imaging Systems and Techniques (IST)* (pp. 1–5), China. https://doi.org/10.1109/IST.2017.8261541.

Raju, B., Raju, N. S. D., Akkara, J. D., & Pathengay, A. (2016). Do it yourself smartphone fundus camera–DIYretCAM. *Indian Journal of Ophthalmology*, *64*(9), 663.

Rana, J., & Galib, S. M. (2017). Cataract detection using smartphone. *2017 3rd International Conference on Electrical Information and Communication Technology (EICT)* (pp. 1–4), Bangladesh. https://doi.org/10.1109/EICT.2017.8275136.

Roy, A. M., Bose, R., & Bhaduri, J. (2022). A fast accurate fine-grain object detection model based on YOLOv4 deep neural network. *Neural Computing and Applications*, *34*(5), 3895–3921. https://doi.org/10.1007/s00521-021-06651-x.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: inverted residuals and linear bottlenecks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4510–4520), USA.

Triyadi, A. B., Bustamam, A., & Anki, P. (2022). Deep learning in image classification using VGG-19 and residual networks for cataract detection. *2022 2nd International Conference on Information Technology and Education (ICIT&E)* (pp. 293–297), Indonesia.

Wang, Y., Tang, C., Wang, J., Sang, Y., & Lv, J. (2021). Cataract detection based on ocular B-ultrasound images by collaborative monitoring deep learning. *Knowledge-Based Systems*, *231*, 107442. https://doi.org/10.1016/j.knosys.2021.107442.

Xiong, L., Li, H., & Xu, L. (2017). An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis. *Journal of Healthcare Engineering*, 1–16. https://doi.org/10.1155/2017/5645498.

Yang, J.-J., Li, J., Shen, R., Zeng, Y., He, J., Bi, J., Li, Y., Zhang, Q., Peng, L., & Wang, Q. (2016). Exploiting ensemble learning for automatic cataract detection and grading. *Computer Methods and Programs in Biomedicine*, *124*, 45–57.

Zhang, L., Li, J., Zhang, i, Han, H., Liu, B., Yang, J., & Wang, Q. (2017). Automatic cataract detection and grading using deep convolutional neural network. *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)* (pp. 60–65), Italy. https://doi.org/10.1109/ICNSC.2017.8000068.