Intelligent Methods in Engineering Sciences 2(2): 042-047, 2023

IMIENS

INTELLIGENT METHODS
IN ENGINEERING SCIENCES

https://www.imiens.org

International Open Access

June, 2023

e-ISSN 2979-9236

Research Article

https://doi.org/10.58190/imiens.2023.11

Detection Of Cataract, Diabetic Retinopathy and Glaucoma Eye Diseases with Deep Learning Approach

Gözde Arslan [a, *], ***Çağatay Berke Erdaş*** [a]

[a] *Department of Computer Engineering, Başkent University, Ankara, Turkey*

ARTICLE INFO

Article history:
Received 20 January 2023
Accepted 28 April 2023

Keywords:
CNN
Classification
Eye-disease
EfficientNet
Retinal Fundus

ABSTRACT

Eye diseases are one of the serious health problems affecting human life. Cataract, diabetic retinopathy and glaucoma eye diseases cause visual impairment and cause irreversible eye defects. Throughout human life; genetic, age and environmental factors affect people's eye health. Detection of the disease plays a critical role in order to apply the right diagnosis and therefore increase the quality of life of the patient. With the developing technology, artificial intelligence can detect eye defects and therefore whether there is a disease or not. This study aims to develop solutions for detecting an important health problem such as eye health by using deep learning models. In the related study, Convolutional Neural Networks models, one of the deep learning types are used. The data set used for disease detection includes a total of 2748 Retinal Fundus images taken from 1374 normal individuals and 1374 different disease groups. In order to compare the classification performances and to achieve better performance, a solution to the disease detection problem was sought by using a total of 5 different Convolutional Neural Networks architectures. These are DenseNet, EfficientNet, Xception, VGG, Resnet. For the validity of the approach, it was tested using the 10-fold cross-validation technique. Accuracy, Recall, Precision, F1-Score, and Matthews's coefficient correlation metrics were used as performance evaluation criteria. When the classification performances were examined, the results obtained with the EfficientNet architecture were measured as 94.88%, 94.88%, 95.02%, 94.88%, 89.89% for Accuracy, Recall, Precision, F1-Score, and Matthews's coefficient correlation metrics. In this context, the best classification performance was obtained with the EfficientNet architecture.

This is an open access article under the CC BY-SA 4.0 license.
(https://creativecommons.org/licenses/by-sa/4.0/)

1. INTRODUCTION

The eye is a vital organ that enables us to perceive the environment we live in. Many dangerous eye diseases can be prevented with early diagnosis and without vision loss. Today, eye diseases are experienced by many individuals based on genetic and environmental factors. Eye diseases primarily occur due to disorders in the eye membrane, eye lens, and nerves. According to the World Health Organization (WHO), vision loss and visual impairment affect at least 2.2 billion people globally, 1 billion of which have preventable visual impairment. [1] Visual impairment, which significantly affects quality of life, adversely affects all areas where people live. In recent years, the increased use of technological devices has proportionally increased visual impairments. Regarding the age of visual disease onset, diseases have begun to be seen at very young ages. According to WHO, at least 450 million children globally have a vision problem requiring diagnosis [2]. Early diagnosis is crucial so that the visual impairment does not progress to an advanced level. Along with early diagnosis, using the right diagnosis method is also critical, as the use of incorrect diagnosis methods increases the rate of visual impairment. Individuals should undergo regular evaluations of vision level and eye health using the right diagnostic methods. Many studies have been conducted to determine the right diagnostic methods. With the development of technologies in the field of health, deep learning (DL) and patient diagnosis-oriented applications are increasing. DL aims to provide people with faster and better-quality methods for the diagnosis of eye diseases. The World Health Assembly approved the 2014–2019 action plan for universal access to eye health. This plan, which is a roadmap for member states, WHO secretariat, and international partners, aimed to reduce visual impairment by 25% and increase vision screenings by 2019 [1].

In this context, many recent studies have investigated early diagnosis and diagnosis of eye diseases. To detect eye diseases more easily and ensure accurate results, models using artificial DL technologies have been investigated. This study aims to use the convolution neural network (CNN) model, a DL technology, to identify an

*** Corresponding Author:** berdas@baskent.edu.tr

accurate and fast diagnostic method for visual diseases. Chapter II presents recent studies of eye disease. Chapter III explains the CNN method. Chapter IV describes the experimental details. Chapter V presents the discussion and conclusion.

2. Related Work

Many different studies have investigated the detection of eye diseases. Malik et al. [3] assessed artificial intelligence techniques for rapid diagnosis in medical healthcare systems. They used Decision Tree, Random Forest, Naïve Bayes, and Neural Network machine learning algorithms and included variables related to patient data, age, disease history, and clinical observations. The random forest and decision tree algorithms had an accuracy (ACC) of over 90% compared to the other algorithms. Gelder et al. [4] discussed human clinical trials and cell types for Retinal Pigment Epithelium (RPE) transplantation. In the study, they examined developments in retinal regenerative medicine over the last decade. They discussed human clinical trials for RPE transplantation. The article provides a comprehensive overview of regenerative and restorative medicine for eye disease and highlights the potential for continued progress in developing new diagnoses for eye diseases. Abbas [5] developed automated computer-aided diagnostic (CADx) systems to evaluate glaucoma eye disease, using CNN architecture to extract the features. It aimed to distinguish between glaucoma and non-glaucoma retinal fundus images and used 1200 retinal images. On average, 99% ACC was achieved. Jain et al. [6] aimed to identify normal individuals with retinal problems and normal individuals without retinal problems via feature extraction using images. The model was trained using images taken from the hospital studied on patient and normal individual's retinal fundus images data sets. ACC results were between 96.5–99.7%. Metin et al. [7] studied retinal diseases, which negatively affect the lives of individuals and may occur with ageing, emphasising the importance of early diagnosis. Machine learning and DL methods have been used in studies on the optical coherence tomography (OCT) imaging method. In the study of classification of retinal diseases, CNN based ResNet50 and MobileNetV2 models were used, which obtained average macro ACC values of approximately 81–94%. When considered together, the ResNet50 and MobileNetV2 models had an average F1 score of 0.75 for Myopic Choroidal Neovascularization (MKV-CNV), 0.86 for Drusen, 0.90 for Diabetic Macular Edema (DME) and 0.96 for normal retinas.

Rarki et al. [8] discussed diabetic eye disease (DED) using the CNN method for the detection of retinal eye disease. They used multiple classes of DED, and the model was tested on various retinal fundus images collected from a public dataset and described by an ophthalmologist. Overall, 81.33% ACC, 100% sensitivity, and 100% specificity were achieved for multiclass classification. Umer et al. [9] assessed the use of OCT for retinal eye disease detection and classification. Ophthalmologists currently manually detect retinal eye disease with the aid of OCT images, which may be inaccurate and subjective. They presented different methods to automate the detection of retinal eye disease using a public four-class retinal eye diseases dataset and modified Alexnet and ResNet-50 models for deep feature vector extraction. The proposed retinal eye disease detection method achieved an overall average ACC index of over 99.95%. Gargeya et. al. [10] discussed the automatic diagnosis of diabetic retinopathy (DR). An automatic system for DR diagnosis was developed using image processing techniques. CNN model is used for the processing and classification of retinal images. Total of 75 137 publicly available retinal fundus images from diabetic patients was used. Area under the receiver operating characteristic curve (AUC) as a metric to measure model was used. 5-fold cross validation was used. The study presented the results using different metrics to measure the performance of the model and observed it was able to diagnose DR with 97% AUC, %94 sensitivity, %98 specificity.

3. Methods

3.1. General Overview

CNN, a DL model, is typically used for image processing. It is considered an ideal approach for the evaluation of eye images as the model captures and classifies features in different images. After scaling retinal fundus images to determine the presence of eye disease, the images are classified with models developed based on Resnet, DenseNet, EfficientNet, Xception and VGG architectures. A general overview of this study is presented in Figure 1.

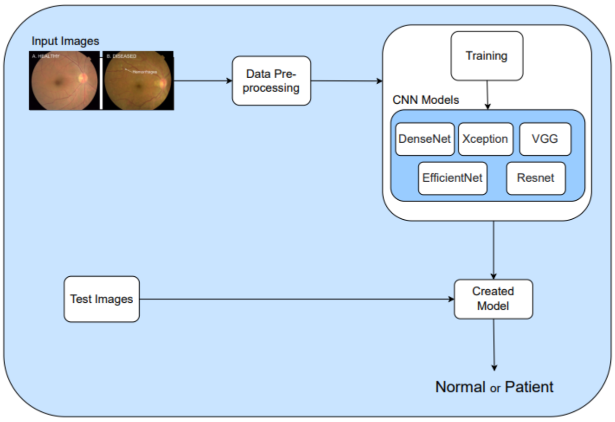

Figure 1. General overview

3.2. Convolutional Neural Networks

CNN is a class of DL neural networks used in many areas, such as image recognition, problem detection, and data analysis [11]. CNN is one of the most effective architecture classes as it can detect patterns well and has a high level of ACC regarding image recognition. In addition to being widely used for image analysis, CNN is also used in various fields to obtain accurate results, including in healthcare for the diagnosis and interpretation of medical images. It can learn data-driven, highly representative, hierarchical image properties from adequate training data [12-13]. CNN architecture captures features in images from different operations and categorises them and consists of convolutional, pooling, and fully connected layers [14]. The convolutional layer is the first place where the image is taken, and it is filtered by entering the input. After filtering, the values and attributes are formed. The pooling layer creates smaller features by ignoring unnecessary features and reducing processing power. In the fully connected layer, the image is turned into a flat vector.

EfficientNet is a family of image classification models which achieve state-of-the-art ACC while being an order of magnitude smaller and faster than previous models [15]. EfficientNet architecture develops models with much better efficiency by scaling each dimension equally rather than randomly scaling the width, depth, or resolution of input data. DenseNet is a network architecture in which each layer directly connects to all other layers in a feed-forward manner (within each dense block). In DenseNet architecture, each layer is interconnected with all other layers. Each layer can access all previous layers. This prevents the image feature from being lost in the transition between layers [16]. In Visual Geometry Group (VGG) architecture, 224*224 size images are used in the input layer. VGG architecture consists of 11, 13, 16, and 19 convolution layers across 6 different models [17]. Each convolution operation is a model that aims to create a feature map that is smaller but deeper than the image. Resnet has demonstrated state-of-the-art performance on various image classification benchmarks, such as ImageNet, CIFAR-10, and CIFAR-100, and has also been applied in other computer vision tasks, such as object detection and segmentation [18]. Resnet architecture skips several layers, using a technique called skip connection, and connects directly to the output. This architecture can be used when too much depth occurs in a more complex structure. Xception means 'extreme start'. It applies filters to each of the depth maps. Xception architecture has the same number of parameters as Inception V3; however, the performance gains seen in Xception are not due to increased capacity but rather due to more efficient use of model parameters [19].

4. Results

4.1. Dataset

In this study, normal and diseased data from eye-related images were used. Images were obtained from different sources. Combining the Mendeley dataset [20], data from the Joint Shantou International Eye Center (JSIEC), Shantou City, Guangdong province, China [21], and data from IDRID, Oculus [22], 1374 normal and 1374 diseased images (a total of 2748) were used. There were 2 different classes in this dataset: normal and diseased. Figure 2 shows an example of normal (B) and diseased (A) eyes.

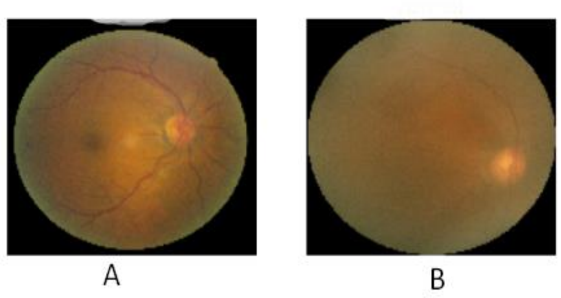

Figure 2. Eye images

4.2. Exprerimental Setup

The results of the DenseNet, EfficientNet, Xception, VGG, and Resnet CNN architectures were discussed. ReLu and softmax functions are used throughout the architectures. Figure 2 presents the architecture of the general CNN model. An image is first filtered in the convolution layer, and feature maps are created. The ReLu function then resets negative values. The goal is to make the resulting features easier to understand. Finally, the final vector for classification is created. This final vector calculates the probability that the classes belong to a specific image. These probabilities are calculated in the last layer of the block using the softmax function.

The Adam Optimizer was used during the training process. The transaction was terminated after 100 epochs. The batch size used was 64. The image size used was 128*128. The learning rate was set at 0.001. Before the epoch, the training data was shuffled.

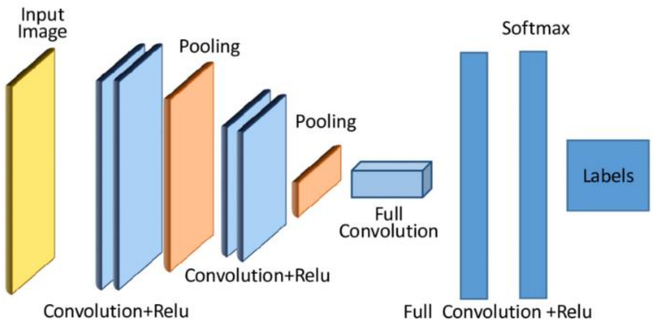

Figure 2. A generic CNN Architecture [23]

4.3. Evaluation

K-fold cross validation is a method used to evaluate the performance of machine learning models [24]. K-fold cross-validation works by dividing the dataset into k parts. One of these parts is used as the test set, and the rest are used as the training set. This process is done k times, and each time a different piece is used as the test set, meaning each piece is used as both a training and test set When the process is completed, the performance evaluation criteria of the model are averaged for each test and an overall performance score is generated [25]. The illustration of the mentioned process is illustrated in Figure 3.

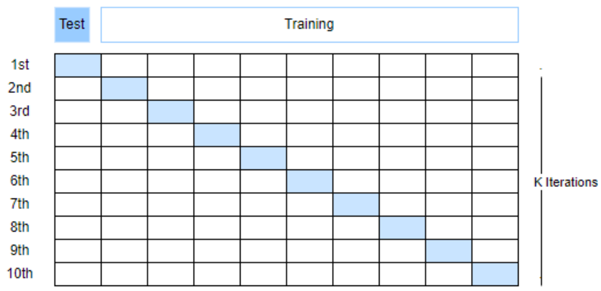

Figure 3. K- Folds

Experiments in this study were performed using 10-fold cross-validation. To measure the performance of the experiment, the ACC (1), sensitivity (recall or true positive rate (TPR)) (2), precision (positive predictive value (PPV)) (3), F1-score (4), and Matthew's correlation coefficient (MCC) were calculated using the MCC (5) metrics. The results are comprised of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) metrics.

$$Accuracy\ (ACC) = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FN} + \mathrm{TN} + \mathrm{FP}} \qquad (1)$$

$$Sensitivity\ (TPR) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \qquad (2)$$

$$Precision\ (PPV) = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \qquad (3)$$

$$F1 - Score = 2 * \frac{\mathrm{Precision} * \mathrm{Sensitivity}}{\mathrm{Precision} + \mathrm{Sensitivity}} \qquad (4)$$

$$Matthews's\ coefficient\ correlation\ (MCC) = \frac{\mathrm{TP} * \mathrm{TN} - \mathrm{FP} * \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP})(\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN})}} \qquad (5)$$

4.4. Empirical Results

In this study is designed to detect eye disease using binary classification. Table 1 presents the results of the 10-fold cross-validation.

Table 1. Results from the methods

	ACC %	TPR %	PPV %	F1-Score %	MCC %
EfficientNet	94.88	94.88	95.02	94.88	89.89
DenseNet	85.35	85.35	85.67	85.31	71.02
Xception	68.28	68.78	69.58	68.45	38.34
VGG	83.08	83.08	84.36	82.92	67.42
Resnet	65.87	65.87	75.65	62.25	40.34

In the statistical calculations shown in Table 1, the EfficientNet architecture demonstrated the highest performance in ACC, recall, precision, F1-score and MCC. Resnet lagged behind the other architectures in terms of ACC, recall, precision, F1-Score and MCC performance.

5. CONCLUSIONS

Vision defects are one of the important health issues affecting human life. Eye diseases may not show any symptoms until they progress. Therefore, early and accurate diagnosis is critical for preventing vision loss. The eyes, which are one of the most vital organs, greatly affect individuals' quality of life. Consequently, it is crucial that eye disease diagnoses satisfy and meet expectations by being efficient and accurate. With technological advancements, many diagnosis methods have emerged. The diagnosis methods used should be successful and safe. Artificial intelligence, which is one such advancement, has had a great impact in this context. The primary purpose of artificial intelligence studies, such as this one, is to identify the fastest and most reliable method for targeted diagnosis. The DenseNet, EfficientNet, Xception, VGG, and Resnet models are CNN-based DL architectures distinct from other diagnosis methods that have been developed to determine whether an eye is normal or diseased based on binary classification. In this study, a confusion matrix was used to calculate the results. TP refers to the number of accurate predictions of normal individuals. FP refers to the inaccurate prediction of eye disease. TN refers to the correctly estimated number of individuals with eye disease. FN refers to the number of false predictions of normal individuals [26]. Among the models, EfficientNet demonstrated the best results, with 94.88% ACC, 94.88% recall, 95.02% precision, an F1 score of 94.88%, and an MCC of 89.89%. Therefore, this model can be used to detect and distinguish eye diseases. Although the results obtained are promising, future studies are required to obtain precise and accurate results by using more retinal fundus images and working with different models.

6. DISCUSSION

This paper assessed eye disease detection problems using cataract, DR, and glaucoma images obtained from individuals. To solve the related problems, a binary classification-based approach was developed. This approach was based on the principle of extracting features by taking various data from a dataset of different eye diseases. A new dataset was obtained by combining various datasets to assess the ability of different DL architectures to detect and classify eye disease. CNN architectures were trained on the existing dataset, and the results between the architectures were compared. The EfficientNet architecture was found to be more effective than the other models. EfficientNet has a network structure predominantly comprised of CNN layers and can achieve high ACC rates with less computational cost than deeper and wider networks due to its ability to scale depth and width [27]. As EfficientNet can achieve high ACC rates using limited data, it may be more useful than other models for the early diagnosis of eye diseases. EfficientNet is more effective than other DL models for diagnosing eye diseases due to its high performance, scalability, and ability to transfer learning across large and complex datasets. EfficientNet's success relies on combining the outputs of a pre-trained model to achieve high ACC rates on smaller datasets. Therefore, when working with smaller datasets, the overall performance of the model increases, and the risk of over-learning is reduced [27]. The proposed system is a preliminary step in eye disease detection that can be further explored in future studies. The current results were obtained by examining retinal fundus images using DL architectures.

Previous studies on eye disease detection and diagnosis were reviewed. Bulut et al. [28] conducted a study based on the EfficientNET-B6 CNN model to diagnose various eye diseases. In their study, a test dataset was used to identify 'referable' retinal disorders, achieving a sensitivity of 0.9439, specificity of 0.8604, and an ACC of 0.86. Paradisa et al. [29] investigated a DL approach with a fusion model for the classification of three classes of fundus images: no DR, non-proliferative DR, and proliferative DR. The model architectures used were DenseNet121 and Inception-ResNetV2. The proposed method achieved a success rate of 91% using a combination of both architectures. Ağca et al. [30] proposed a hybrid model using ESA and machine learning for the detection and grading of DR. The Aptos 2019 dataset was used, while the SMOTE technique was leveraged to address the class imbalance. The study used the 10-fold cross-validation method for training and testing. The study achieved a 93% ACC and 93% F1 score. Compared to previous studies, the results obtained for EfficientNet in the current study, in particular its 94.88% ACC, indicate that it performs better.

Author's Note

Abstract version of this paper was presented at 10th International Conference on Advanced Technologies (ICAT'22), 25-27 November 2022, Van, Turkey.

References

[1] World Health Organization (2019), https://www.who.int/news-room/newsletters, (Last Access: 02.04.2023)

[2] https://www.iapb.org/learn/vision-atlas/magnitude-and projections/child-eye-health/ (Last Access: 01.02.2023)

[3] Malik, S., Kanwal, N., Asghar, M. N., Sadiq, M. A., Karamat, I., & Fleury, M. (2019). Data driven approach for eye disease classification with Machine Learning. *Applied Sciences*, *9*(14), 2789. doi:10.3390/app9142789.

[4] R. N. Van Gelder, M. F. Chiang, M. A. Dyer, T. N. Greenwell, L. A. Levin, R. O. Wong, and C. N. Svendsen, "Regenerative and restorative medicine for eye disease," *Nature Medicine*, vol. 28, no. 6, pp. 1149–1156, 2022.

[5] Q. Abbas, "Glaucoma-deep: Detection of glaucoma eye disease on retinal fundus images using Deep Learning," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, 2017.

[6] L. Jain, H. V. Murthy, C. Patel, and D. Bansal, "Retinal eye disease detection using Deep Learning," *2018 Fourteenth International Conference on Information Processing (ICINPRO)*, 2018.

[7] Batuhan Metin, Bahadır Karasulu, 'Derin Öğrenme Modellerini Kullanarak İnsan Retinasının Optik Koherans Tomografi Görüntülerinden Hastalık Tespiti', 2022, 9 – 19.

[8] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, and K. Wang, "Convolutional neural network for multi-class classification of Diabetic Eye Disease," *ICST Transactions on Scalable Information Systems*, p. 172436, 2018.

[9] M. J. Umer, M. Sharif, M. Raza, and S. Kadry, "A deep feature fusion and selection-based retinal eye disease detection from oct images," *Expert Systems*, 2023, doi: https://doi.org/10.1111/exsy.13232.

[10] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using Deep Learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017, doi:doi.org/10.1016/j.ophtha.2017.02.008.

[11] N. Jmour, S. Zayen, A. Abdelkrim, "Convolutional neural networks for image classification", Conference: 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET).

[12] D. Olcer and C. B. Erdas, "Detection of visual impairment from retinal fundus images with deep learning," *2022 Medical Technologies Congress (TIPTEKNO)*, 2022.

[13] H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN Architectures, dataset characteristics and transfer learning," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, 2016.

[14] Ö. İnik ve E. Ülker , "Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri", Gaziosmanpaşa Bilimsel Araştırma Dergisi, 6, 3, ss. 85-104, 2017.

[15] E. J. Humphrey and J. P. Bello, "Rethinking automatic chord recognition with Convolutional Neural Networks," *2012 11th International Conference on Machine Learning and Applications*, 2012.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected Convolutional Networks," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[17] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, 64 ICLR 2015 - Conference Track Proceedings, 1–14.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[19] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[20] https://data.mendeley.com/datasets/4k64fwnp4k/3, (Last Access: 07.08.2022)

[21] https://www.kaggle.com/datasets/linchundan/fundusimage100

0, (Last Access: 07.08.2022)
[22] https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification, (Last Access: 07.08.2022)
[23] https://www.researchgate.net/figure/A-generic-CNN-Architecture_fig1_344294512 (Last Access: September 2020)
[24] Brownlee, J. (2021). How to Use K-Fold Cross-Validation for Model Evaluation. Machine Learning Mastery. https://machinelearningmastery.com/k-fold-cross-validation/ (Last Access: 07.04.2023)
[25] Balaban, M. E., & Kartal, E. (2015). Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları. Çağlayan Kitabevi, İstanbul.
[26] Kohavi, R., & Provost, F. (2016). Glossary of Terms. November. doi:10.1023/A:1017181826899.
[27] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In International Conference on Machine Learning (pp. 6105-6114).
[28] Batuhan Bulut, Volkan Kalın, Burcu Bektaş Güneş, Rim Khazhin, ‘Classification of Eye Disease from Fundus Images Using EfficientNet’, APRIL 3, 2022.
[29] R. H. Paradisa, A. Bustamam, W. Mangunwardoyo, A. A. Victor, A. R. Yudantha, and P. Anki, “Deep feature vectors concatenation for eye disease detection using fundus image,” *Electronics*, vol. 11, no. 1, p. 23, 2021. doi: https://doi.org/10.3390/electronics11010023.
[30] K. AĞCA and H. TAKCI, “Hibrit Bir Model Oluşturarak Diyabetik Retinopati Sınıflandırılması,” *European Journal of Science and Technology*, 2022. doi: 10.31590/ejosat.1112980