LIFE: International Journal of Health and Life-Sciences
ISSN 2454-5872

Arslan & Erdaş, 2024

Volume 09, pp. 62-72

Received: 16th February 2024

Revised: 29th February 2024, 04th March 2024

Accepted: 21st February 2024

Date of Publication: 15th June 2024

DOI- https://doi.org/10.20319/lijhls.2024.10.6272

This paper can be cited as: Arslan, G., Erdaş, C. B., (2024). Efficient Detection of Multiclass Eye Diseases Using Deep Learning Models: A Comparative Study. LIFE: International Journal of Health and Life-Sciences, 9, 62-72.

EFFICIENT DETECTION OF MULTICLASS EYE DISEASES USING DEEP LEARNING MODELS: A COMPARATIVE STUDY

Gözde Arslan
MSC, Department of Computer Engineering, Başkent University, Ankara, Turkey,
arslangozde187@gmail.com

Çağatay Berke Erdaş
Ph.D., Assistant Professor, Department of Computer Engineering, Başkent University, Ankara, Turkey
berdas@baskent.edu.tr

Abstract

Eye diseases pose a significant health threat, impacting human life adversely. Conditions like cataracts, diabetic retinopathy, and glaucoma lead to irreversible and serious health issues. Age, genetics, and environmental factors play a crucial role in eye health. Accurate diagnosis is essential for effective treatment, placing a heightened responsibility on clinicians. Advanced technology and deep learning enable the detection and identification of eye diseases. This research aims to utilize prominent Convolutional Neural Network models, including DenseNet, EfficientNet, Xception, VGG, and ResNet, to detect eye diseases. Technical term abbreviations are explained, and the dataset comprises 4217 retinal fundus images, including 1038 cataracts, 1098 diabetic retinopathy, 1007 glaucoma, and 1074 healthy individuals. Model performance is evaluated through metrics like accuracy, recall, precision, F1-score, and Matthews's correlation *coefficient using 10-fold cross-validation. Among the models tested, EfficientNet demonstrates the best results with 87.84% accuracy, 92.84% recall, 94.41% precision, 93.53% F1-score, and 83.87% Matthews's correlation coefficient. Consequently, EfficientNet proves to be the most effective architecture for classifying eye diseases in this study.*

Keywords

Convolutional Neural Networks, Eye Disease, EfficientNet, Retinal Fundus

1. Introduction

Today, eye diseases, which affect many individuals, are influenced by both genetic and environmental factors. Regular eye examinations are necessary to detect any potential diseases. It is essential to address this issue promptly and effectively. Early diagnosis and proper treatment are crucial to improve eye health quickly. The impact of eye diseases on a person's quality of life is significant. With recent technological advances, the use of technological devices has increased visual impairments.

Eye diseases can arise from disorders in the membrane, lens, and nerves. Cataracts, diabetic retinopathy, and glaucoma are the primary diseases caused by these deformations. Technical terms are explained where necessary. Cataracts cause complete vision blurring which significantly impairs eyesight. Diabetic retinopathy is another type of eye disease that should be monitored. If an individual has diabetic retinopathy disease, this condition does not only affect conditions like insulin resistance and one of the places it affects is the eyes. If unchecked and the disease progresses, the result can be significant eye health problems. Another prevalent condition is glaucoma, commonly referred to as high eye pressure, whose symptoms include blurred vision and headaches in patients. Occasionally, individuals may experience ocular pressure as an indication of the presence of eye disease. Such indications may not always be immediately recognized, and it is only when the condition progresses that significant vision loss becomes apparent. Early diagnosis is imperative, as irreversible vision loss may occur if allowed to advance. According to data from the World Health Organization (WHO) in 2021, at least 2.2 billion people worldwide are affected by visual impairment and vision loss. Of this number, 1 billion individuals suffer from preventable visual impairment (Lim et al., 2023). Numerous technology-supported studies have been introduced in the literature to aid in the detection of eye diseases. The objective of artificial intelligence studies is to provide a quicker and more precise diagnosis of eye diseases.

The objective of this research is to develop a swift and precise diagnostic approach for visual ailments via CNN, which is an artificial intelligence technique. The paper is structured as follows: Section 2 entails a review of recent research on eye diseases. Section 3 elucidates the CNN methods. Section 4 presents a detailed account of the experiments, followed by Section 5, which outlines the discussion and conclusion.

2. Related Studies

The detection of eye disease today involves several technologies and associated methods. One highly effective approach to improving the accuracy of diagnosis and enhancing the range of detection options is through the use of deep learning methodology. Several deep-learning studies have been conducted, focused on specific eye diseases, and are widely documented in the literature.

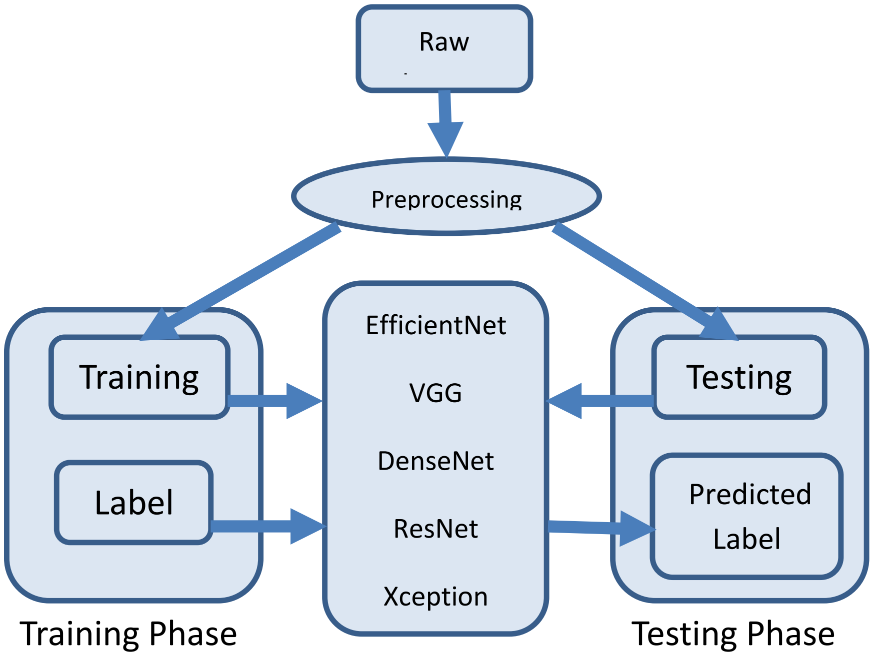
Ağalday et al. (2021) investigated an automatic diagnosis system for cataract disease utilizing color fundus images. The study involved classifying the images using CNN. A dataset comprised of 6392 color fundus images of both the right and left eyes of 5000 patients, which represented actual patient information, was used. Upon comparison of the general success criteria of CNN and Deep Residual Networks (Deep Residual Networks), it was observed that the Deep Residual Networks method achieved a higher success rate. The CNN method had an accuracy of 89%, while the Deep Residual Networks method had an accuracy rate of 95%. Yalçın et al. (2018) applied a deep learning technique to detect diseases in diabetic-retinopathy images taken from retinal imagery. The team conducted image standardization using various retinal images. The resultant CNN-based classification model displayed a performance accuracy of 98.5%. Abbas (2017) designed an automated computer-aided diagnosis (CADx) system geared toward evaluating glaucoma eye disease. The system employed a CNN architecture to extract features and distinguish between glaucoma and normal retinal fundus images. The study showed 99% accuracy performance using 1200 retinal images. Yıldırım et al. (2022) aimed to detect glaucoma in fundus images using various CNN models such as AlexNet, ResNet-18, VGG16, SqueezeNet, and GoogleNet, and assessed the models' performance criteria. A total of 1000 images, comprising 500 healthy individuals and 500 glaucoma patients from a publicly available glaucoma dataset, were utilized. After evaluating the results between the models, we determined that VGG16 performed the best in terms of recall in the test dataset with 97.96%, while GoogleNet had the highest scores for specificity, accuracy, and F1-Score with 98.97%, 97.98%, and 98%, respectively. K. Prasad et al. (2019) delved into the types of diseases in diabetic retinopathy and glaucoma. In this investigation, a deep neural network model was employed to aid in the early detection of the disease. The proposed model aims to achieve high accuracy levels, and a result of 80% was achieved by the developed model. Gunawardhana et al. (2020) examined the automatic diagnosis of diabetic retinopathy through the utilization of machine learning algorithms. The study outlined the development of an automatic system for diagnosing diabetic retinopathy using techniques from image processing and machine learning. The authors explained that a range of machine-learning algorithms were employed to process and classify retinal images. K-Nearest Neighbor Algorithm, Support Vector Machine, Naive Bayes Classifier, and Decision Tree. The results indicated that the algorithm achieved a high level of accuracy in diagnosing diabetic retinopathy. Specifically, it was discovered that the algorithm can diagnose diabetic retinopathy with 95% accuracy.

3. Materials & Methods

3.1. General Overview

The detection of eye diseases can be approached as a supervised classification task, wherein retinal fundus images are classified by a neural network. To facilitate its processing in our convolutional neural network, all retinal fundus images are initially resized to 224x224 pixels. Following pre-processing steps, images and corresponding labels are segregated for training and testing phases. During the training phase, the suggested deep-learning models are trained with both the images and their respective labels. In the testing phase, the models operate on images they have not encountered previously in an attempt to determine their class. The model's classification performance is assessed by comparing the predicted labels with the actual labels. Figure 1 provides a visual depiction of the suggested approach's overall structure.

Figure 1: *General Overview*



(*Source: Self/Authors' Own Illustration*)

3.2. Convolutional Neural Networks

The main characteristic of EfficientNet is that it uses a scaling coefficient to optimize three key dimensions: width, depth, and resolution. This scaling factor is utilized to establish the width, depth, and resolution of the network. Essentially, a sole parameter can be specified to scale the mesh, resulting in the network becoming either larger or smaller. Increasing network size leads to amplified parameters and computational power, which typically equates to superior performance. However, excessively large networks may prove to be disadvantageous in terms of both training and running expenses. EfficientNet prioritizes efficiency in network design by employing a scaling factor to achieve this equilibrium. EfficientNet has been designed to produce smaller and more computationally efficient networks which, in turn, have demonstrated high performance on numerous tasks (Nayak et al., 2022).

DenseNet represents a significant advancement in the architecture of Convolutional Neural Networks. By incorporating dense connections, this deep learning model promotes information flow, whereby the output of each layer relies on the outputs of all preceding layers. This feature facilitates deeper and more effective learning. Furthermore, BottleNeck layers decrease the computational burden of the network by utilizing fewer parameters. Transition layers adapt the network's depth and width, improving its flexibility. Global Average Pooling (GAP) is commonly applied in the final layer to smooth the feature maps and produce the network's output (Huang et al., 2019).

The VGG architecture is recognized for its remarkably straightforward and repetitive structures. It utilizes consecutive blocks, typically comprised of compact cores measuring 3x3 and maximum pooling layers of 2x2 dimensions. Fully connected layers follow these blocks. The VGG depth comprises either 16 or 19 layers, often in two different types. This profound structure increases network complexity and frequently offers efficient performance in computer vision assignments (Sengupta et al., 2019).

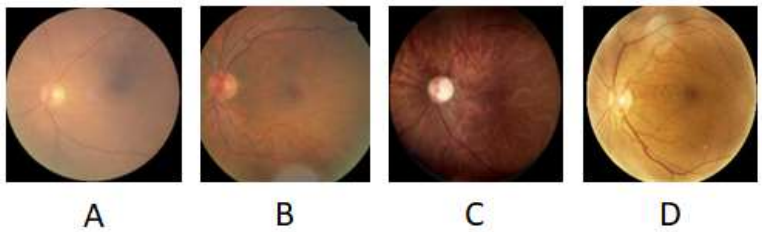
Xception revolutionizes the architecture of convolutional neural networks by implementing a technique known as "depthwise separable convolution." This method separates traditional convolution operations, resulting in more efficient performance, lower computational costs, and boosted learning capabilities. Xception builds upon Inception's "factorized convolutions" strategy while being heavily inspired by the said architecture. Xception is frequently acknowledged for its lower parameter count and computational cost, whilst exhibiting exceptional performance in transfer learning applications and tasks including image classification (Chollet, 2017). The key characteristic of ResNet is the utilization of a system termed the "residual block". The residual block acquires "residuals" by adding together the inputs and outputs of a layer, and leverages these residuals to aid the network's learning process. This enables the network to become deeper by circumventing the "vanishing gradient" issue, a recurrent problem in traditional deep networks, which usually experience a deterioration in performance as their depth increases (Koonce, 2021).

4. Results

4.1. Dataset

In this study, retinal fundus images of different classes are analyzed in detail. The selected images were sourced from various datasets, including Mendeley and IDRID (Yoo, 2021), Oculus (Linchundan, 2019), and the Shantou International Eye Centre (JSIEC) (Doddi, 2022). These datasets provided a wealth of data, categorized into four different classes: cataract, diabetic retinopathy, glaucoma, and healthy. The dataset comprises 4217 images in total, depicting 1038 cases of cataracts, 1098 cases of diabetic retinopathy, 1007 cases of glaucoma, and 1074 healthy individuals. Example data from each class is shown in Figure 2, with (A) representing cataracts, (B) diabetic retinopathy, (C) glaucoma, and (D) healthy individuals. This provides an overview of the different classes.

Figure 2: *Examples of Retinal Fundus Images Obtained from Individuals with Cataract (A), Diabetic Retinopathy (B), Glaucoma (C), and Healthy (D)*



(Source: Self/Authors' Own Illustration)

4.2. Experimental Setup & Performance Evaluation

Cross-validation is a method used to evaluate model performance. K-fold cross-validation involves splitting the dataset into k segments, with each segment serving as a test set in sequence, while the remainder is used for training. This process is repeated k times, with a different segment used for testing in each iteration. Once the process is completed, the evaluation criteria for each test are averaged to give an overall performance score for the model (Erdaş & Sümer, 2023). Experiments in this study were conducted using 10-fold cross-validation. The dataset was partitioned into 10 equal parts resulting in 10 unique outcomes from each iteration. To evaluate the experiment performance, accuracy, recall, precision, F1-Score, and Matthews's coefficient correlation (MCC) metrics were used.

The training process utilized the Adam Optimizer at an epoch value of 100 and an image size of 224x224. The learning rate was set to 0.001, and the training data was randomized before each epoch to increase variability. Batch normalization was implemented to expedite training, boost the learning rate, and enhance stability. ReLU and softmax functions were applied to all architectures.

4.3. Empirical Results and Findings

Based on Table 1, the EfficientNet architecture demonstrates superior performance across several critical metrics. With an accuracy rate of 87.84%, recall of 92.84%, precision of 94.41%, F1-Score of 93.53%, and a Matthews coefficient correlation of 83.87%, the model is well-equipped for accurate classification and adeptly spotting positive instances, while simultaneously maintaining excellent overall performance balance. In addition to EfficientNet, the DenseNet, VGG, and ResNet architectures yield significant achievements. DenseNet achieves outstanding precision of 92.43% and F1-Score of 91.67%, displaying its proficiency in accurately detecting positive instances. VGG, on the other hand, exhibits a high accuracy of 84.92% and a precision of 92.35%. The ResNet architecture is impressive with elevated accuracy of 87.98% and F1-Score of 89.26%. In comparison to its counterparts, the Xception architecture displays inferior performance. Its accuracy stands at 67.16%, recall at 70.85%, precision at 82.75%, F1-Score at 76.16%, and MCC at 56.59%, indicating suboptimal performance, especially in properly identifying certain classes. This thorough examination highlights the subtle advantages and disadvantages of each design, stressing the significance of meticulous evaluation when determining the optimal framework for a particular task.

Table 1: *Performance Comparison of Different CNN Architectures for Multi-Class Detection on Retinal Fundus Images*

ARCHİTECTURE	ACCURACY	RECALL	PRECİSİON	F1-SCORE	MCC
EFFİCİENTNET	87.84	92.84	94.41	93.53	83.87
VGG	84.92	92.35	91.38	91.8	79.92
DENSENET	83.26	92.43	90.96	91.67	77.69
RESNET	79.2	90.67	87.98	89.26	72.34
XCEPTİON	67.16	70.85	82.75	76.16	56.59

(Source: Self/Authors' Own Illustration)

5. Conclusion & Discussion

The eyes, one of the most important vital organs of the human body, have a major impact on a person's quality of life. Visual impairments may initially manifest as mild symptoms but can progress to severely compromise the quality of life. Detecting eye problems early with thorough examinations is critical before they cause significant harm to an individual's well-being. Unfortunately, individuals often experience ocular health issues but fail to seek appropriate care. Therefore, eye disorders progress, and treatment becomes more challenging. There are various methods to diagnose various eye diseases. Technology, being an integral part of daily life, plays a crucial role in diagnosing eye diseases. New and accurate diagnostic methods have been created due to advancements in technology, which are the two primary components. The most crucial technology for diagnosing purposes is the deep learning method. In this study, our objective is to achieve the most efficient and precise targeted treatment method using deep learning architectures.

It was proposed the use of DenseNet, EfficientNet, Xception, VGG, and Resnet models, which are CNN-based deep learning architectures that set them apart from other treatment methods.

An approach has been devised to ascertain whether an eye is normal or afflicted with cataracts, diabetic retinopathy, or glaucoma disease through categorical classification.

Based on the obtained results, the EfficientNet model demonstrates considerable superiority in performance metrics for multi-classification detection tasks. EfficientNet's high coefficient indicates consistent performance across different classes. On the other hand, the Xception architecture scores lower on these performance measures. This may suggest difficulties in accurately recognizing or distinguishing particular classes, potentially leading to unstable performance. This technical analysis highlights the significance of opting for EfficientNet, particularly for classification tasks, and stresses the need for meticulous model selection in the given task's context.

The EfficientNet design yields superior outcomes attributed to the optimized implementation of depth and width scaling. This enhances the model's learning capacity while augmenting computational efficiency and accuracy rates. Conversely, the Xception design offers a less extensive model learning capability compared to other Convolutional Neural Network architectures, which may lead to underwhelming performance outcomes by decreasing the model's ability to learn. DenseNet and VGG architectures possess the capability to acquire complex visual characteristics. Consequently, they produced results almost equivalent to the EfficientNet architecture in terms of recall, precision, and F1-Score performance metrics. Retinal fundus images were analyzed through the use of deep learning Convolutional Neural Network architectures, resulting in the detection of ocular diseases. While the results obtained from analyzing these detections show potential, further research will aim to achieve greater precision and accuracy for the related problem through the use of additional retinal fundus images and alternative models.

REFERENCES

Lim, Z. W., Chee, M.-L., Soh, Z. D., Majithia, S., Sahil, T., Tan, S. T., Sabanayagam, C., Wong, T. Y., Cheng, C.-Y., & Tham, Y.-C. (2023). Six-year incidence of visual impairment in a multiethnic Asian population. *Ophthalmology Science, 3*(4), 100392. https://doi.org/10.1016/j.xops.2023.100392

AĞALDAY, F., & ÇINAR, A. (2021). Derin öğrenme Mimarilerini kullanarak katarakt tespiti. *European Journal of Science and Technology*. https://doi.org/10.31590/ejosat.1012694

Yalcin, N., Alver, S., & Uluhatun, N. (2018). Classification of retinal images with deep learning for early detection of diabetic retinopathy disease. *2018 26th Signal Processing and Communications Applications Conference (SIU)*. https://doi.org/10.1109/siu.2018.8404369

Abbas, Q. (2017). Glaucoma-deep: Detection of glaucoma eye disease on retinal fundus images using Deep Learning. *International Journal of Advanced Computer Science and Applications*, *8*(6). https://doi.org/10.14569/ijacsa.2017.080606

YILDIRIM, Ö., & ALTUNBEY ÖZBAY, F. (2022). Fundus Görüntülerinden Derin öğrenme teknikleri Ile Glokom Hastalığının tespiti. *European Journal of Science and Technology*. https://doi.org/10.31590/ejosat.1216404

Prasad, K., Sajith, P. S., Neema, M., Madhu, L., & Priya, P. N. (2019). Multiple eye disease detection using Deep Neural Network. *TENCON 2019 - 2019 IEEE Region 10 Conference (TENCON)*. https://doi.org/10.1109/tencon.2019.8929666

Gunawardhana, P. L., Jayathilake, R., Withanage, Y., & Ganegoda, G. U. (2020). Automatic diagnosis of diabetic retinopathy using Machine Learning: A Review. *2020 5th International Conference on Information Technology Research (ICITR)*. https://doi.org/10.1109/icitr51448.2020.9310818

Nayak, D. R., Padhy, N., Mallick, P. K., Zymbler, M., & Kumar, S. (2022). Brain tumor classification using dense efficient-net. *Axioms*, *11*(1), 34. https://doi.org/10.3390/axioms11010034

Huang, W., Peng, G., & Tang, X. (2019). *A Limit of Densely Connected Convolutional Networks V1*. https://doi.org/10.17504/protocols.io.8j6hure

Sengupta, A., Ye, Y., Wang, R., Liu, C., & Roy, K. (2019). Going deeper in spiking neural networks: VGG and residual architectures. *Frontiers in Neuroscience*, *13*. https://doi.org/10.3389/fnins.2019.00095

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2017.195

Koonce, B. (2021). ResNet 50. *Convolutional Neural Networks with Swift for Tensorflow*, 63–72. https://doi.org/10.1007/978-1-4842-6168-2_6

Yoo, T. (2021, December 22). *Deep-learning-based segmentation of fundus photographs to detect central serous chorioretinopathy*. Mendeley Data. https://data.mendeley.com/datasets/4k64fwnp4k/3

Linchundan. (2019, June 18). *1000 fundus images with 39 categories*. Kaggle. https://www.kaggle.com/datasets/linchundan/fundusimage1000

Doddi, G. V. (2022, August 28). *Eye_diseases_classification*. Kaggle. https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification

Erdaş, Ç. B., & Sümer, E. (2023). A fully automated approach involving neuroimaging and deep learning for parkinson's disease detection and severity prediction. *PeerJ Computer Science*, *9*. https://doi.org/10.7717/peerj-cs.1485