SHM PUBLISHER
SURYA HIJAU MANFAAT
Journal of Soft Computing Exploration
Homepage: https://shmpublisher.com/index.php/joscex
JOSCEX
p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Eye disease classification using deep learning convolutional neural networks

Eko Hari Rachmawanto[1], Christy Atika Sari[2], Andi Danang Krismawan[3], Lalang Erawan[4], Wellia Shinta Sari[5], Deddy Award Widya Laksana[6], Sumarni Adi[7], Noorayisahbe Mohd. Yaacob[8]
[1,2]Study Program in Informatics Engineering, Universitas Dian Nuswantoro, Indonesia
[3]Study Program in Animation, Universitas Dian Nuswantoro, Indonesia
[4,5]Study Program in Information Systems, Universitas Dian Nuswantoro, Indonesia
[6]Study Program in Visual Commmunicaton Design, Universitas Dian Nuswantoro, Indonesia
[7]Study Program in Information System, Universitas Amikom Yogyakarta, Indonesia
[8]Strategic Information and Software Systems Lab (SISS), Center for Software Technology and Management (SOFTAM), University Kebangsaan Malaysia, Malaysia

Article Info

Article history:

Received Nov 4, 2024
Revised Nov 23. 2024
Accepted Nov 28, 2024

Keywords:

Convolutional Neural Networks
Deep Learning
Eye Disease
Image Classification
TensorFlow

ABSTRACT

Eye diseases, if not diagnosed early, can lead to severe visual impairments, including blindness, posing significant challenges in clinical practice. Traditional diagnostic methods often face limitations in accuracy and efficiency, necessitating advanced solutions. This research aims to address these challenges by employing deep learning with Convolutional Neural Networks (CNN) enhanced by transfer learning to classify eye diseases. The study utilized a dataset of 4,217 images categorized into four classes: Normal (1,074 images), Glaucoma (1,007), Cataract (1,038), and Diabetic Retinopathy (1,098). The CNN model, implemented in TensorFlow, was trained and evaluated to achieve high accuracy. The results indicate a classification accuracy of 95%, with particularly outstanding performance for Diabetic Retinopathy, achieving 100% precision and recall. Compared to previous studies, such as Seetha et al. (2022) and Sarki et al. (2021), which reported 75% and 81.33% accuracy, respectively, this study demonstrates a significant improvement. These findings highlight the model's robustness in enhancing early detection and clinical decision-making in ophthalmology, with future work focusing on expanding the dataset and exploring more advanced deep learning architectures to improve performance further.

Corresponding Author:

Christy Atika Sari
Study Program in Informatics Engineering, Faculty of Computer Science
Universitas Dian Nuswantoro
Imam Bonjol 207, Semarang, 50131, Central Java, Indonesia
Email: christy.atika.sari@dsn.dinus.ac.id
https://doi.org/10.52465/joscex.v5i4.493

1. INTRODUCTION

Eye disease are disorders that affect the eyes and vision in temporary or permanent ways, potentially creating discomfort, impaired sight, and in some cases, complete blindness if left untreated [1]. Common eye conditions include cataracts, where the eye lens becomes cloudy; glaucoma, characterized by increased eye pressure damaging the optic nerve; and macular degeneration, which affects the retina's center, reducing sharp vision [2]. Other conditions, such as diabetic retinopathy, may be due to diabetic complications causing destruction of blood vessels in the retina [3], [4]. Early detection and treatment can prevent long-term loss of vision, and with the advancement in diagnosis methodologies, even machine learning techniques have become ever so important to diagnose and manage.

It is important not only because it is one of the causes of reduced quality of life, but also because most the eye diseases develop gradually, and their beginning is very often symptom-free [5]. Therefore, early diagnosis is an important challenge for clinical practice. For instance, glaucoma is often termed the "silent thief of sight" because many times patients may not notice that they have lost vision. Similarly, many other serious eye conditions, such as diabetic retinopathy, are often asymptomatic at their onsets but might subsequently lead to irreversible damage unless treated timely and appropriately [6]. While older adults are mainly affected by macular degeneration, it also predominantly causes central vision loss that is debilitating to essential tasks, including reading and driving. These diseases are even more insidious because they usually progress to an advanced stage before symptoms appear and diagnosis is made. Although current techniques for disease diagnosis are developing, there is often a limitation of specialists and variable human interpretations of medical images.

The problem of late and often incorrect early detection of eye diseases, in this perspective, needs a profound software solution focused on automatic deep learning models using Convolutional Neural Networks in structures such as TensorFlow. Such models, when trained with major datasets of eye images, can recognize diseases such as glaucoma, diabetic retinopathy, and macular degeneration long in advance before their symptoms manifest well [7], [8]. Clinical-diagnostic systems with integrated models, like these, therefore, raise the degree of precision in diagnosis and reduce variability that may be introduced by subjective human assessment [9]. These technologies can also be implemented in telemedicine or even on cell phones to easily screen for conditions in remote or resource-poor areas without specialist care. This will not only ensure early diagnosis but also allow for continuous improvements as new data are introduced into the models, further increasing the precision and reliability of the system over time [10]. This will enable, in turn, real-time analysis, better clinical decisions, and have a significant reduction impact on the incidence of preventable blindness and improvement in the patient's outcomes in conditions of complex eye diseases.

Recent advancement in the area, as related to the detection of a variety of eye ailments, has been by deep CNN [11], whose role it is to scan medical images for disease-related patterns that may be impossible for human observation. In application to the diagnosis of eye diseases, these deep CNN excel in processing images of the retina while autonomously learning about important features such as those changing in the optic nerve head structure due to glaucoma or the development of hemorrhages in diabetic retinopathy [12]. What is very special about ConvNet is that it can inherently learn the features most relevant visually without necessarily being specified explicitly by making use of convolutional layers, pooling layers, and fully connected layers [13]–[15]. That inherently suggests that CNN can be applied to most ocular diseases and a variety of image modalities, which thus allows generalizing seamlessly across different populations of patients and image sources. Further, CNN reduce dependence on human interpretation because, with the offering of consistent and reproducible analysis, huge in a clinical environment, such skills are hardly available. Its use optimizes the whole diagnosis process in diagnostic systems and finally creates an opportunity for better treatment by making decisions more appropriately on time.

In study by Seetha et al. (2022) [16]proposed a holistic approach in tackling the challenges involved in the diagnosis of diabetic retinopathy, a retinal disease wherein blood vessels in the tissues of the retina are impaired. They noted in their work the serious effect of the disease should it not be identified and treated, thus the need for comprehensive eye check-ups by ophthalmologists. While most of the earlier methods involve several classification methodologies and machine learning algorithms, most of these techniques require very intensive time consumption for both training and testing, and most of them are not validated on diverse data sets. Authors focusing on filling these gaps concentrated on the development of an effective and robust ensemble CNN classification model identifying the presence of diabetic retinopathy from fundus images. They proposed a plain CNN and ensemble of CNN for classification and presented their respective performances in terms of accuracy. The CNN ensembles performed better than the simple CNN with 75% accuracy, which could be attributed to the extensive hyper-parameter tuning. Therefore, the ensemble CNN proved to be a better classification approach for the diagnosis of diabetic retinopathy, hence enhancing early intervention in the clinical practice of the condition.

Other related research by Sarki et al. (2021) [17] proposed a new dimension in diagnosing DED through the development of a computer-aided diagnosis framework with deep learning techniques. They were fully aware that only early treatment has the potential to optimize benefits while minimizing incidence related to irrecoverable deterioration in vision. Thus, they have directed the focus of the study toward the use of retinal fundus images as a major diagnostic technique for DED and other eye diseases. They thus noticed that such image detection was laborsome and time-consuming; hence, the development of more efficient methods would be required. Further motivated by the significant progress demonstrated by deep learning in clinical applications, the authors tried to address challenges about multiclass classification of the retinal eye diseases-a domain of active research. The various studies conducted by the authors were implemented with a novel CNN model in turn tested on a wide variety of retinal fundus images, sourced from openly available datasets and annotated by ophthalmologists. In fact, these results proved that this model reached its best performance with an accuracy of 81.33%, besides sensitivity and specificity of 100%. Thus, this model is promising for the improvement in the early-detection process of the different DED types, enhancing decision-making in ophthalmology.

On the other hand, Chakraborty et al. (2020) [18]proposed a new automatic diseased region detection system using the Convolutional Neural Network in order to assist the doctors in the diagnosis with scan and X-ray images. This shall improve decision-making with high precision in disease detection and hence leverage the benefits of CNN, a subset of deep learning, which is a major domain under Artificial Intelligence. CNN inherently require very little preprocessing compared to other deep learning algorithms, and therefore they have become very suitable for application in medical image analysis. The authors considered two different kinds of medical image data, namely, OCT images and chest X-ray images taken among children within the age bracket of 1 to 5 years, as input for the classification process. The model CNN was designed for the treatment and classification of medical images, where several performance metrics like accuracy, loss, and training time have been measured with much care. After the implementation of this system in hardware, some testing using already-trained models achieved a significant improvement of about 90% validation accuracy on the eye dataset and approximately 63% on the lung dataset. The proposed system will help improve diagnostic accuracies for the medical fraternity while simultaneously helping reduce the infant mortality rate due to pneumonia and helping early detection of eye diseases for better patient outcomes.

Based on relevant research carried out above, in this paper, TensorFlow will be employed in the implementation of a Deep Learning Convolutional Neural Network for classification into a total of four classes, namely, Cataract, Diabetic Retinopathy, Glaucoma, and Normal. CNN is a type of deep learning model specifically designed for image processing, leveraging convolutional layers to automatically extract features from input images. This enables the detection of intricate patterns critical for disease classification. TensorFlow, a robust open-source framework, provides tools for building and training CNNs efficiently. Drawing inspiration from the work of Seetha et al. (2022) [5], who realized better classification accuracy using an ensemble CNN. Based on inspirations drawn from this research study, the application of some of the techniques in improving diagnostic precision will be attempted. Additionally, the findings of Sarki et al. (2021) [17]had good performance regarding identifying multiclass classification using deep learning methods, it has motivated the approach followed in this work. Furthermore, Chakraborty et al. (2020) [18] presented CNNs, which do not require much preprocessing and thus were considered suitable for different imaging modalities. This is made real in the study through the adoption of the TensorFlow-based approach, which actually gave a higher classifying accuracy than what was realized in previous studies. Thus, integrating these insights into this research has been done to present an enhanced classification framework that may not only enhance the early detection of eye-related pathologies but also support decision-making processes in clinical practice for improved outcomes in ophthalmology.

2. METHOD

The process begins with a dataset containing four classes of retinal images, representing different eye conditions. The dataset is split into training and validation subsets, with 80% of the data allocated for training and 20% for validation. The training data is fed into the tensorflow model, where parameters are initialized to optimize the model's learning capabilities. The trained model is then evaluated using the validation data to assess its accuracy and effectiveness in classifying the images correctly. After training, the trained model undergoes the classification phase, where it predicts the class of each input image. The model’s performance is further analyzed by a confusion matrix, which provides a detailed breakdown of classification accuracy across the four classes. This workflow ensures a structured approach to developing a reliable model for eye disease classification based on retinal images. Figure 1 illustrates the research methodology flow for the classification of eye diseases using a tensorflow-based deep learning model.

.

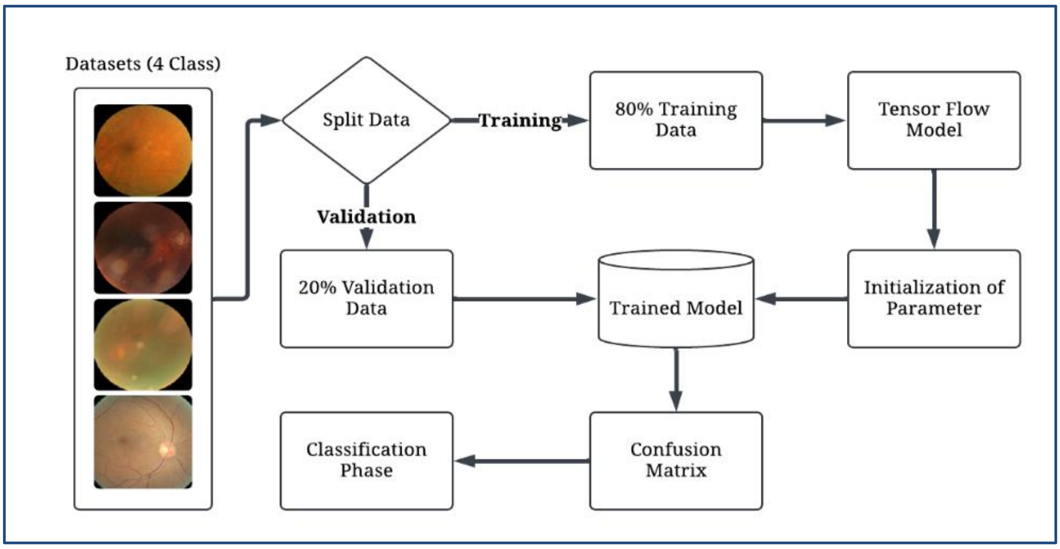


Figure 1. Research Flow Methodology

2.1 Datasets Information

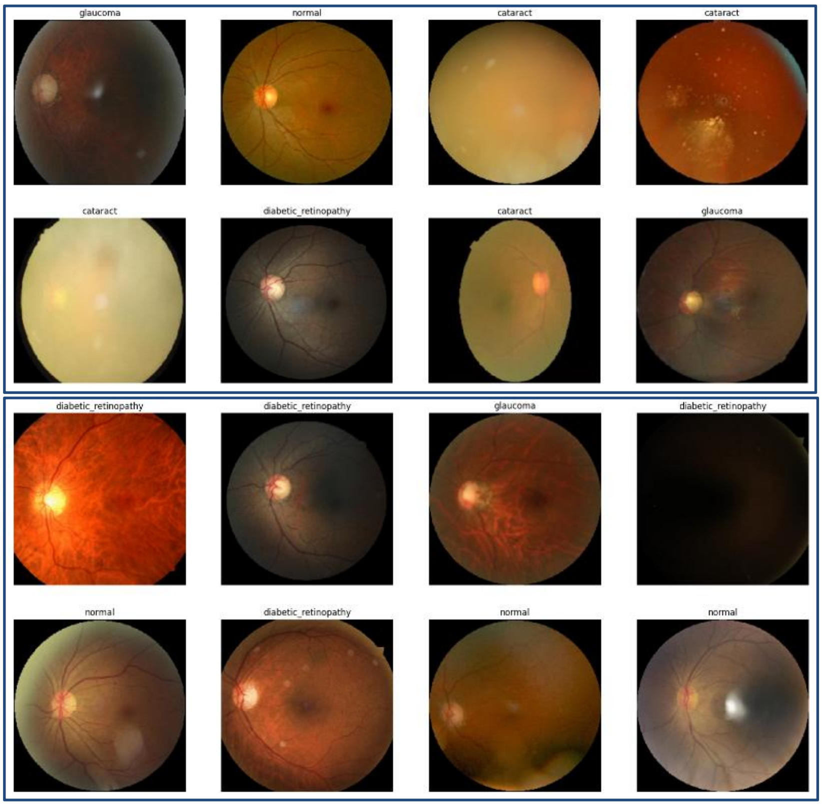
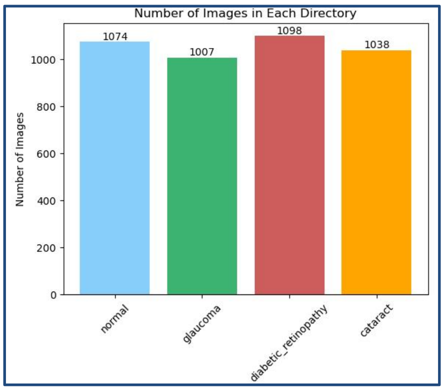
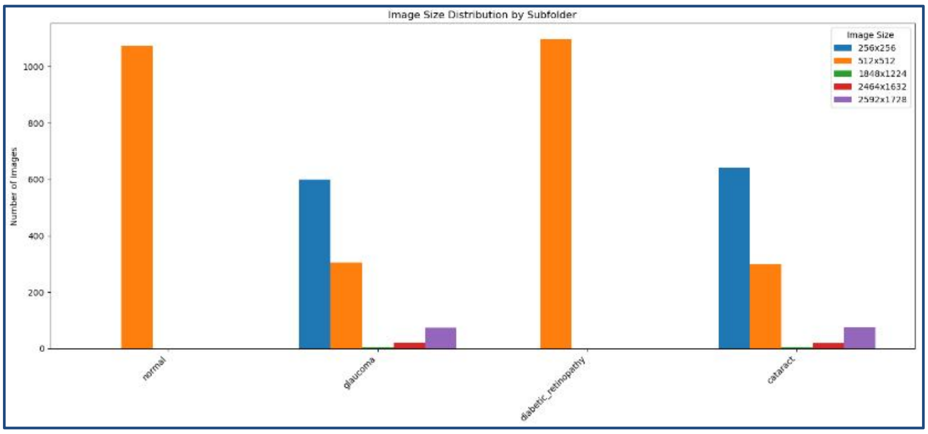


Figure 2. Sample Datasets

For this dataset, it will contain normal and three major classes of eye diseases, namely glaucoma, cataract, and diabetic retinopathy images, in total 4217. It includes 1074 samples as Normal, 1007 samples as Glaucoma, 1038 samples as Cataract, and 1098 samples as Diabetic Retinopathy. Figure 3 (b) represents that each class is unequal, which corresponds to the natural occurrence of this disease in the dataset. Figure 3 (c) gives an idea of the total samples available in each class. Figure 2 shows sample images from the dataset, and there is great variability in the appearance and characteristics of each of the eye diseases. A dataset like this is used for training and evaluation in the classification model. Each image adds up and forms a discriminative hyper-plane for classifying any new observation amongst these four classes.



(a) Total of Sample Each Class



(b) Image Size Each Class

Figure 3. Information of Datasets

From the information on the datasets, the dataset applied in this research is quite balanced, as the total samples for all the four classes- Normal, Glaucoma, Cataract, and Diabetic Retinopathy-are quite comparable in number. The closeness of the distribution reduces the chances of class imbalance, one of the problems that are usually seen in medical image datasets, because some conditions may be represented with fewer images. Due to this fact, no extra over-sampling should be done through SMOTE - Synthetic Minority Over-sampling Technique - or any balancing technique, as in this way the model will not be biased regarding class imbalance. For the most part, none of these extra operations concerning balancing is really time-consuming; hence, the raw data will be allowed to feed the deep learning model directly, keeping the pipeline pretty simple and straightforward. This dataset is available on Kaggle: https://www.kaggle.com/code/mojtabaameri/eye-diseases-classification-by-tensorflow-94-9-ac#eye-diseases-classification, It really builds a perfect foundation to train a quite strong classification model, as it exactly represents each class without extensive adjustment, hence improving model training efficiency and reducing unnecessary preprocessing interventions.

2.2 Initialization of Model Tensor Flow and Transfer Learning

A transfer learning model was thus developed, using TensorFlow, to classify eye disease images based on the four classes of images [19]. Our model architecture has each convolution block consisting of convolutional layers following max-pooling layers with batch normalization layers that stabilize and accelerate the training process [20], [21]. First Convolution block consists of two convolutional layers having 128 filters, kernel size (3, 3) followed by max pooling and batch normalization. Each subsequent block increases the number of filters further to 256 to enable the model to grasp higher and richer features layer by layer. In this model, instead of flattening the feature maps, Global Average Pooling is used which reduces the overfitting along with parameters [22], [23]. After that, add a fully connected layer, followed by a dropout

.

with the rate set to 0.5, and finally the softmax output layer, which classifies an input image among the four classes.

To further enhance model performance and prevent overfitting, two key callbacks were implemented: EarlyStopping and ModelCheckpoint. EarlyStopping monitors validation accuracy, halting training if no improvement is observed over 12 consecutive epochs and restoring the best model weights. This approach prevents excessive training and potential overfitting. ModelCheckpoint saves the best model based on validation loss, ensuring the best-performing model is retained. This model is then compiled with the best optimizer, Adam, together with a sparse categorical cross-entropy loss function that will be appropriate for multiclass classification. Train on up to 75 epochs, leveraging the callbacks provided in the code that will help in optimizing the performance of this model by saving the best performing version of the model. The model developed is shown in Figure 4.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 256, 256, 128)	3,584
conv2d_1 (Conv2D)	(None, 256, 256, 128)	147,584
max_pooling2d (MaxPooling2D)	(None, 128, 128, 128)	0
batch_normalization (BatchNormalization)	(None, 128, 128, 128)	512
conv2d_2 (Conv2D)	(None, 126, 126, 256)	295,168
conv2d_3 (Conv2D)	(None, 124, 124, 256)	590,080
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 256)	0
batch_normalization_1 (BatchNormalization)	(None, 62, 62, 256)	1,024
conv2d_4 (Conv2D)	(None, 60, 60, 256)	590,080
conv2d_5 (Conv2D)	(None, 58, 58, 256)	590,080
max_pooling2d_2 (MaxPooling2D)	(None, 29, 29, 256)	0
batch_normalization_2 (BatchNormalization)	(None, 29, 29, 256)	1,024
conv2d_6 (Conv2D)	(None, 27, 27, 256)	590,080
conv2d_7 (Conv2D)	(None, 25, 25, 256)	590,080
max_pooling2d_3 (MaxPooling2D)	(None, 12, 12, 256)	0
batch_normalization_3 (BatchNormalization)	(None, 12, 12, 256)	1,024
global_average_pooling2d (GlobalAveragePooling2D)	(None, 256)	0
dense (Dense)	(None, 512)	131,584
dropout (Dropout)	(None, 512)	0
dense_1 (Dense)	(None, 4)	2,052

Figure 4. Developed Model

Figure 4 shows the architecture of the CNN, including the layer types, output shapes, and parameter counts. The convolutional model starts with convolutional layers in the first block with 128 filters, kernel size 3x3, and stride 1, reducing spatial dimensions of the feature maps. Further, a max-pooling layer reduces the dimensions by half, and a batch normalization layer for stabilizing the training. Successive convolutional blocks increased the filters to 256, increasing feature extraction, and were followed by similar max-pooling and batch normalization layers. Instead of flattening, Global Average Pooling was used, which drastically reduced the number of parameters while preventing overfitting as well. A dropout of 0.5 is also added for further regularization. The fully connected layer provided the probabilities for classification using the softmax function-classifying images into Cataract, Diabetic Retinopathy, Glaucoma, and Normal. The architecture performs balancing between accuracy and computational efficiency; hence, the parameters increase progressively across layers to enhance feature learning.

2.3 Confusion Matrix Evaluation

A further analysis has been made in model performance using a confusion matrix, which will actually give the breakdown concerning the model's predictions across each class [24], [25]. This confusion matrix here is showing the actual numbers for each category: true positives, false positives, true negatives, and false negatives. It has given us an idea as far as the strength of this model in distinguishing between these four classes of eye diseases, namely: Normal, Diabetic Retinopathy, Glaucoma, and Cataract. The matrix helps to single out any particular classes that are more vulnerable to misclassifications so adjustments can be made to increase accuracy. An ideal confusion matrix is that which is well distributed and maintains high values on the diagonal-a sign that the model predicts most samples in a class correctly, therefore performing excellent classification and minimizing errors. The evaluation of confusion matrix can be seen in Eq (1) – (4).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

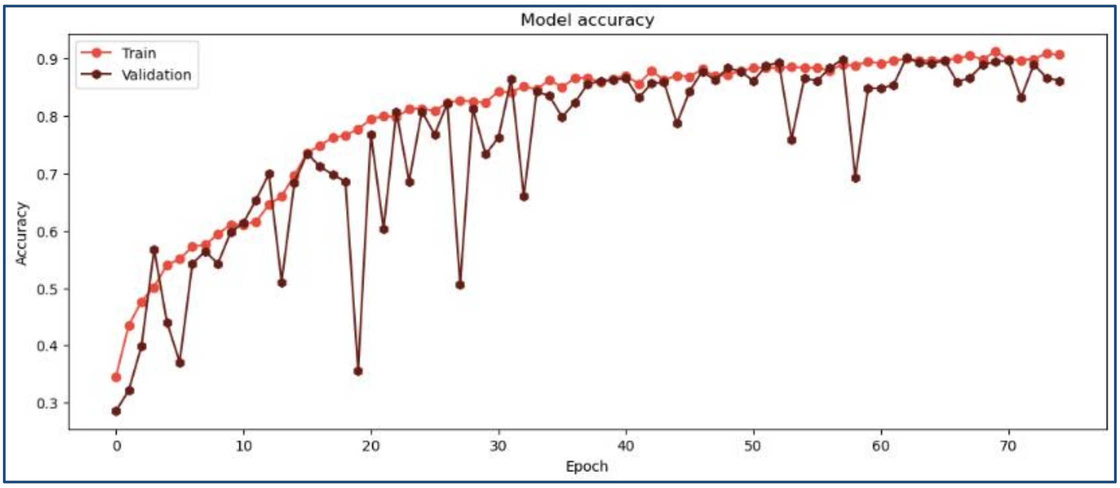
$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - Score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \tag{4}$$

In the Confucian Matrix, TP means that the model rightly predicted the sample belonging to a certain class-say, such as correctly diagnosing the image as Diabetic Retinopathy. TN would mean the model has rightly predicted the sample not being part of the particular class-like an image predicted not to be cataract but it actually is Glaucoma. FP will be where the model classifies a sample as belonging to one class when, in reality, it is not part of that class-for example, when it predicts Glaucoma on an image of Normal ones. FN is the situation where the model will fail to catch an image from a particular class and will result in misclassification; for example, Diabetic Retinopathy classified as Normal. These values will be useful in calculating the precision, recall, and F1-score and will provide insight regarding the model's effectiveness for different classes.

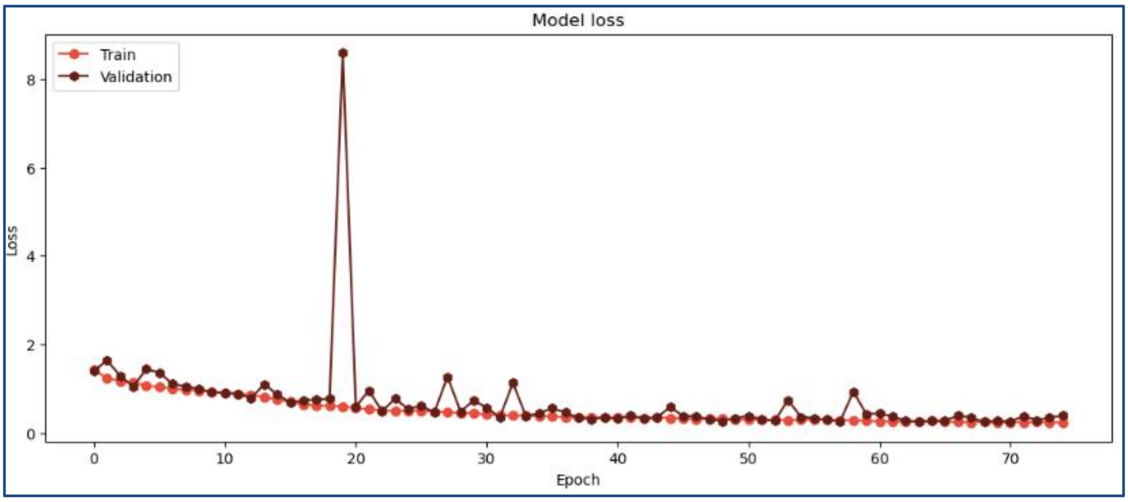
3. RESULTS AND DISCUSSIONS

Following initialization described under Methods, the model was run using Python software, with the help of deep learning libraries, specifically TensorFlow, best suited for the creation and training of CNN models. Segregation of the dataset into training and validation sets by utilizing the proposed transfer learning CNN architecture for this study formed the first step in model training. The model is then optimized using the Adam algorithm and set with relevant parameters, such as EarlyStopping and ModelCheckpoint; this allows validation accuracy to keep track and save only the best weights of the model in the process of training. Training was stopped after the model showed the best performance on improved accuracy and reduced loss. Lastly, the results of training were portrayed in graphs that represent accuracy and loss versus time. Figure 5: The training accuracy graph shown in Figure 5(a) depicts the performance of the model concerning both training and validation data with respect to each epoch. This includes training of the model for recognition of the pattern in data. Figure 5(b) represents a training loss plot presenting a drop in the loss, which clearly denotes how well the model managed to minimize the error in its prediction. These graphs will allow us to decide whether there is overfitting or underfitting and if this model has achieved the best performance on the data it was fitted to.



(a) Model Accuracy

.



(b) Model Loss
Figure 5. Graph Model

After completing the training phase, the trained model was evaluated using a confusion matrix to assess its classification performance. This evaluation reveals how the model's predictions compare to the actual labels, providing a clear insight into the accuracy and classification errors. As illustrated in the confusion matrix displayed in Figure 6.

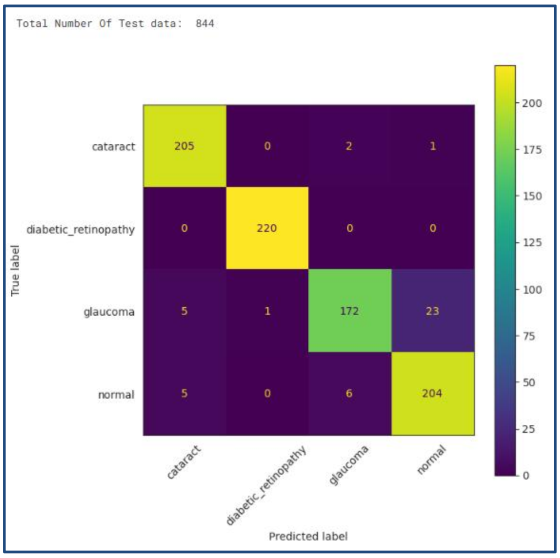


Figure 6. Table of Confusion Matrix Using 20% Testing Data

The different model performance metrics computed are Accuracy, Precision, Recall, and F1-score; these are summarized in the table 1. Overall, these measures of effectiveness describe the classification tasks that the model has conducted: The accuracy will describe the total correctness of its predictions, while precision will basically refer to the proportion of true positives out of all the positives that have been predicted. On the other hand, recall reflects a model's performance in identifying all instances that are relevant. The F1-score is the harmonic mean of precision and recall; it therefore provides a single metric that balances both aspects. The values in Table 1 therefore highlight strengths and areas of improvement concerning the predictive performance of the model.

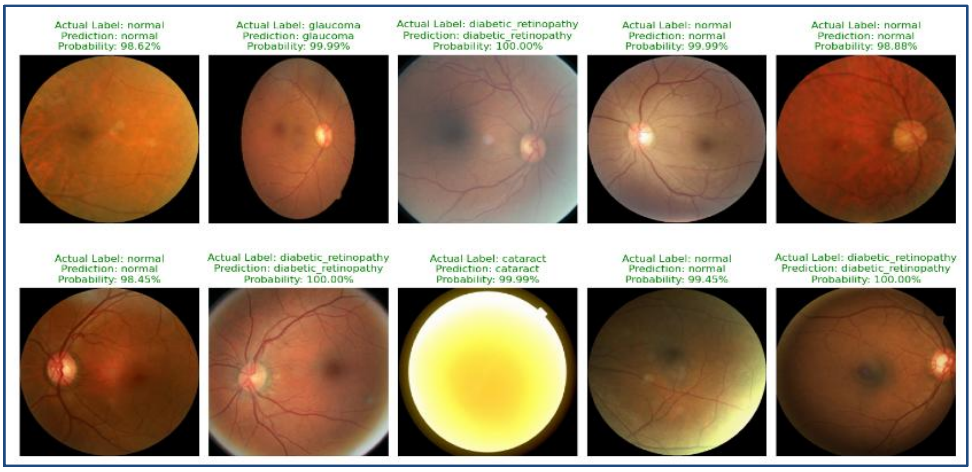


Figure 7. Classification Phase

The final step of this research involved conducting prediction tests to evaluate the model's performance in real-world scenarios. These tests aimed to assess the model's ability to generalize its learning to unseen data, thereby validating its predictive capabilities. The outcomes of these prediction tests are illustrated in Figure 7, which provides a visual representation of the model's performance across various test instances. The results showcase the model's accuracy and reliability in making predictions, further reinforcing the effectiveness of the methodologies employed throughout the study.

Tabel 1. Results of Confusion Matrix Evaluation

Class	Accuracy	Precision	Recall	F1-Score
Normal	95%	89%	95%	92%
Cataract		95%	99%	97%
Glaucoma		96%	86%	90%
Diabetic Retinopathy		100%	100%	100%

Previous works have identified various classification accuracies on the detection of eye diseases using CNN models. For example, Seetha et al. (2022) [16] illustrated an accuracy of 75% using an ensemble CNN for diabetic retinopathy, while Sarki et al. (2021) [17] put forward an accuracy of 81.33% in using a new CNN model for the multiclass classification of retinal diseases. Chakraborty et al. (2020) [18] achieved a 90% validation accuracy of eye datasets but were focused principally on general medical imaging. In this aspect, the current paper outperforms previous works with its high accuracy of 95%, using a CNN based on TensorFlow. Its robustness and reliability are checked for four categories of classification relating to eye diseases: Cataract, Diabetic Retinopathy, Glaucoma, and Normal.

4. CONCLUSION

This study focused on the classification of eye diseases using Deep Learning Convolutional Neural Networks to serve as proof that Tensorflow contributes immensely to enhancing the diagnostic capability of ocular conditions in the medical field. The methodologies adopted include the training of the CNN on the dataset containing comprehensive eye images, followed by the performance evaluation of the model using the confusion matrix. These results are really very high across different classes: 89% for normal, 95% for cataract, 96% for glaucoma, and for diabetic retinopathy, it was a full 100%. The recall rates, however, were equally impressive: 95% for normal, 99% for cataract, 86% for glaucoma, and 100% for diabetic retinopathy. The F1-scores in this regard were also very good, standing at 92% for normal, 97% for cataract, 90% for glaucoma, and again 100% for diabetic retinopathy. The overall observed accuracy by the model was 95%, hence proving to be efficient in the classification task of eye disease diagnosis. The study really outshines the contribution of CNN in improving diagnostic accuracy to achieve appropriate timely interventions and better outcomes in ophthalmology. In the future, a number of directions may be pursued in a search for more robust and clinically applicable models. This would involve the use of a dataset comprising a wide range of ocular pathologies with variability in subject demographics, which would result in better generalizability and performance across populations. Besides that, the research of some advanced architectures-such as the mechanism of attention or transfer learning-might probably bring even more accuracy and efficiency.

REFERENCES

[1] M. A. Stepp dan A. S. Menko, "Immune responses to injury and their links to eye disease," *Transl. Res.*, vol. 236, hal. 52–71, Okt 2021, doi: 10.1016/j.trsl.2021.05.005.

[2] K. Naderi, J. Gormley, dan D. O'Brart, "Cataract surgery and dry eye disease: A review," *Eur. J. Ophthalmol.*, vol. 30, no. 5, hal. 840–855, Sep 2020, doi: 10.1177/1120672120929958.

[3] E. B. Papas, "The global prevalence of dry eye disease: A Bayesian view," *Ophthalmic Physiol. Opt.*, vol. 41, no. 6, hal. 1254–1266, Nov 2021, doi: 10.1111/opo.12888.

[4] P. Agarwal, J. P. Craig, dan I. D. Rupenthal, "Formulation Considerations for the Management of Dry Eye Disease," *Pharmaceutics*, vol. 13, no. 2, hal. 207, Feb 2021, doi: 10.3390/pharmaceutics13020207.

[5] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, J. Ma, dan K. Wang, "Image Preprocessing in Classification and Identification of Diabetic Eye Diseases," *Data Sci. Eng.*, vol. 6, no. 4, hal. 455–471, Des 2021, doi: 10.1007/s41019-021-00167-z.

[6] B. K. Betzler, S. M. Young, dan G. Sundar, "Intraocular Pressure and Glaucoma in Thyroid Eye Disease," *Ophthalmic Plast. Reconstr. Surg.*, vol. 38, no. 3, hal. 219–225, Mei 2022, doi: 10.1097/IOP.0000000000002049.

[7] N. Papandrianos, E. Papageorgiou, A. Anagnostis, dan K. Papageorgiou, "Efficient Bone Metastasis Diagnosis in Bone Scintigraphy Using a Fast Convolutional Neural Network Architecture," *Diagnostics*, vol. 10, no. 8, hal. 532, Jul 2020, doi: 10.3390/diagnostics10080532.

[8] Y. E. Almalki *et al.*, "Isolated Convolutional-Neural-Network-Based Deep-Feature Extraction for Brain Tumor Classification Using Shallow Classifier," *Diagnostics*, vol. 12, no. 8, hal. 1793, Jul 2022, doi: 10.3390/diagnostics12081793.

[9] H. A. Craze, N. Pillay, F. Joubert, dan D. K. Berger, "Deep Learning Diagnostics of Gray Leaf Spot in Maize under Mixed Disease Field Conditions," *Plants*, vol. 11, no. 15, hal. 1942, Jul 2022, doi: 10.3390/plants11151942.

[10] K. Ahmed, M. K. Dubey, dan S. Dubey, "Guarding Maize: Vigilance Against Pathogens Early Identification, Detection, and Prevention," 2024, hal. 301–318. doi: 10.1007/978-981-99-9621-6_19.

[11] I. U. W. Mulyono, E. H. Rachmawanto, C. A. Sari, dan M. K. Sarker, "A high accuracy of deep learning based CNN architecture: classic, VGGNet, and RestNet50 for Covid-19 image classification," *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 22, no. 5, hal. 1187, Jul 2024, doi: 10.12928/telkomnika.v22i5.26017.

[12] N. R. D. Cahyo, C. A. Sari, E. H. Rachmawanto, C. Jatmoko, R. R. A. Al-Jawry, dan M. A. Alkhafaji, "A Comparison of Multi Class Support Vector Machine vs Deep Convolutional Neural Network for Brain Tumor Classification," in *2023 International Seminar on Application for Technology of Information and Communication (iSemantic)*, IEEE, Sep 2023, hal. 358–363. doi: 10.1109/iSemantic59612.2023.10295336.

[13] M. Ahmad, S. Abbas, A. Fatima, G. F. Issa, T. M. Ghazal, dan M. A. Khan, "Deep Transfer Learning-Based Animal Face Identification Model Empowered with Vision-Based Hybrid Approach," *Appl. Sci.*, vol. 13, no. 2, hal. 1178, Jan 2023, doi: 10.3390/app13021178.

[14] E. A. Aldakheel, M. Zakariah, G. A. Gashgari, F. A. Almarshad, dan A. I. A. Alzahrani, "A Deep Learning-Based Innovative Technique for Phishing Detection in Modern Security with Uniform Resource Locators," *Sensors*, vol. 23, no. 9, hal. 4403, Apr 2023, doi: 10.3390/s23094403.

[15] G. Lassalle, M. P. Ferreira, L. E. C. La Rosa, dan C. R. de Souza Filho, "Deep learning-based individual tree crown delineation in mangrove forests using very-high-resolution satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 189, hal. 220–235, Jul 2022, doi: 10.1016/j.isprsjprs.2022.05.002.

[16] M. Seetha, N. Kalyani, dan Y. Sravani Devi, "An Ensemble CNN Model for Identification of Diabetic Retinopathy Eye Disease," 2022, hal. 191–198. doi: 10.1007/978-981-16-9705-0_19.

[17] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, dan K. Wang, "Convolutional Neural Network for Multi-class Classification of Diabetic Eye Disease," *ICST Trans. Scalable Inf. Syst.*, hal. 172436, Jul 2018, doi: 10.4108/eai.16-12-2021.172436.

[18] P. Chakraborty dan C. Tharini, "Pneumonia and Eye Disease Detection using Convolutional Neural Networks," *Eng. Technol. Appl. Sci. Res.*, vol. 10, no. 3, hal. 5769–5774, Jun 2020, doi: 10.48084/etasr.3503.

[19] J. Gupta, S. Pathak, dan G. Kumar, "Deep Learning (CNN) and Transfer Learning: A Review," *J. Phys. Conf. Ser.*, vol. 2273, no. 1, hal. 012029, Mei 2022, doi: 10.1088/1742-6596/2273/1/012029.

[20] K.-Y. Chen, J. Shin, M. A. M. Hasan, J.-J. Liaw, O. Yuichi, dan Y. Tomioka, "Fitness Movement Types and Completeness Detection Using a Transfer-Learning-Based Deep Neural Network," *Sensors*, vol. 22, no. 15, hal. 5700, Jul 2022, doi: 10.3390/s22155700.

[21] S. Showkat dan S. Qureshi, "Efficacy of Transfer Learning-based ResNet models in Chest X-ray image classification for detecting COVID-19 Pneumonia," *Chemom. Intell. Lab. Syst.*, vol. 224, hal. 104534, Mei 2022, doi: 10.1016/j.chemolab.2022.104534.

[22] Y. Zheng *et al.*, "Application of transfer learning and ensemble learning in image-level classification for breast histopathology," *Intell. Med.*, vol. 3, no. 2, hal. 115–128, Mei 2023, doi: 10.1016/j.imed.2022.05.004.

[23] D. Agarwal, G. Marques, I. de la Torre-Díez, M. A. Franco Martin, B. García Zapiraín, dan F. Martín Rodríguez, "Transfer Learning for Alzheimer's Disease through Neuroimaging Biomarkers: A Systematic Review," *Sensors*, vol. 21, no. 21, hal. 7259, Okt 2021, doi: 10.3390/s21217259.

[24] S.-C. Kim dan Y.-S. Cho, "Predictive System Implementation to Improve the Accuracy of Urine Self-Diagnosis with Smartphones: Application of a Confusion Matrix-Based Learning Model through RGB Semiquantitative Analysis," *Sensors*, vol. 22, no. 14, hal. 5445, Jul 2022, doi: 10.3390/s22145445.

[25] I. Markoulidakis dan G. Markoulidakis, "Probabilistic Confusion Matrix: A Novel Method for Machine Learning Algorithm Generalized Performance Analysis," *Technologies*, vol. 12, no. 7, hal. 113, Jul 2024, doi: 10.3390/technologies12070113.