International Journal of Computer, Network Security and Information System (IJCONSIST)
Vol. 7 No. 1, September 2025

Classification of Eye Diseases Using the AlexNet Convolutional Neural Network Model Algorithm

Moch Deny Pratama
Department of Informatics Management
Universitas Negeri Surabaya
Surabaya, Indonesia
mochpratama@unesa.ac.id

Royal Fajar Sultoni
Department of Informatics, Faculty of Computer Science
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya, Indonesia
20081010175@student.upnjatim.ac.id

Adil Sandy Wardhani
Department of Informatics, Faculty of Computer Science
Universitas Pembangunan Nasional Jawa Timur
Surabaya, Indonesia
20081010021@student.upnjatim.ac.id

Maulana Hassan Sechuti
Department of Informatics, Faculty of Computer Science
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya, Indonesia
20081010183@student.upnjatim.ac.id

Yerezqy Bagus
Department of Informatics, Faculty of Computer Science
Universitas Pembangunan Nasional Jawa Timur
Surabaya, Indonesia
yerezqy.if@upnjatim.ac.id

Dina Zatusiva Haq
Department of Informatics, Faculty of Computer Science
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya, Indonesia
dinaza.if@upnjatim.ac.id

Yoga Ari Tofan
Department of Informatics, Faculty of Computer Science
Universitas Pembangunan Nasional Veteran Jawa Timur
Surabaya, Indonesia
yoga.if@upnjatim.ac.id

Abstract*—This study uses the Convolutional Neural Network (CNN) method with the AlexNet model to classify eye diseases based on medical images. The dataset includes labeled images of three types of eye diseases: cataract, glaucoma, and diabetic retinopathy. The experimental results show that the model achieved an accuracy of 75.18%, which indicates that CNN with the AlexNet architecture can classify eye diseases quite well. This research shows that deep learning can be used to help doctors or health professionals in diagnosing eye diseases through automatic image analysis. Although the accuracy still needs to be improved, this study can serve as a reference for developing an automated diagnostic system in the future. Further research is expected to increase accuracy, expand the dataset, and apply other deep learning techniques to improve the performance of eye disease detection.

Keywords;* *classification of eye diseases, convolutional neural network, alexnet, deep learning

I. INTRODUCTION

The senses, as tools in the human body, enable humans to recognize and feel the surrounding environment. The eyes, as one of the five senses, have an important role in processing visual information. Because it is often used in everyday life, maintaining eye health is crucial [1], [2]. Eye diseases or disorders such as cataract, glaucoma, and diabetic retinopathy can cause blindness and require early detection [3]. Therefore, this study aims to classify the three types of eye disorders using the Convolutional Neural Network (CNN) method with the AlexNet architecture.

Deep learning, especially CNN, has been widely used in various fields including the medical field [4]. In this research, CNN is used to analyse eye images and classify eye diseases. The datasets used include cataract, glaucoma, diabetic retinopathy, and normal eyes, with more than 1000 datasets that have been obtained from available sources. The importance of early detection of eye diseases such as cataract, glaucoma and diabetic retinopathy is the basis for the development of this method [5]. Previous research into the classification of eye diseases using image processing techniques and other algorithms has provided useful results.

Researchers have widely used various studies using image processing techniques to classify eye diseases. In previous research, conducted by Dewi (2022) regarding Classification of Eye Diseases Using Convolutional Neural Networks with the VGG-19 Architecture, the research carried out produced accuracy values of 65.29%, precision of 66.53%, recall of 65, 29%, and f1 - score of 65.40% [6]. Another research was also conducted by Ayu (2022) regarding Classification of Eye Diseases Based on Retinal Fundus Images Using Fractal Box Counting and Fuzzy K - Means Dimensions, this research obtained 76% accuracy results with classification using fuzzy k-means via fractal box counting from fundus image photos retina [7].

By involving a large dataset and involving the latest CNN architecture, this research aims to contribute to the development of early eye disease detection technology. It is hoped that the conclusions of this study will provide a better understanding of the use of CNN methods in the context of eye health, with a focus on the classification of significant diseases.

Through this research, it is hoped that a solution can be found to detect eye disorders earlier, enabling quicker medical action to be taken. The CNN method with the AlexNet architecture was chosen due to its effectiveness in image recognition tasks [8].

II. RESEARCH METHOD

This research uses deep learning methods with Convolutional Neural Network (CNN) algorithm using AlexNet model to classify three types of eye diseases namely cataract, glaucoma, and diabetic retinopathy. The method that will be used in this research is described with a flowchart as in figure 1.

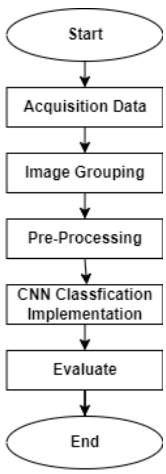

Fig. 1. Eye Disease Classification Methodology with CNN

A. Acquisition Data

In this research we use data sourced from the Kaggle website. The data that will be used in this study is an image of the retinal image of the eye. On the Kaggle website there are human retinal images with various diseases and what will be used in this study is the retinal image of cataract eye disease, glaucoma, diabetic retinopathy, and normal eyes [9]. The eye retinal data that will be used in the Convolutional Neural Network method process is composed of 1038 cataract eye disorder retinal data, 1098 diabetic retinopathy eye disorder retinal data, 1007 glaucoma eye disorder retinal data, and 1074 normal eye retinal data.

Eye Disease Retinal Image Data		
No	Name	Amount of Data
1	Cataract	1038
2	Diabetic Retinoplaty	1007
3	Glaucoma	1074

Fig. 2. Eye Disease Retinal Image Data

B. Image Grouping

The next process is grouping datasets containing retinal images. This grouping is performed to facilitate research in conducting training and testing that will be carried out on human retinal image data. The grouping performed will be divided based on cataract diseases, glaucoma, diabetic retinopathy, and normal eye retina. the data will be grouped and divided into separate folders based on the type of eye disorder that will be stored on the g drive as shown in figure 3.

Fig. 3. Division of Eye Disease Disorders Folder

C. Pre-Processing

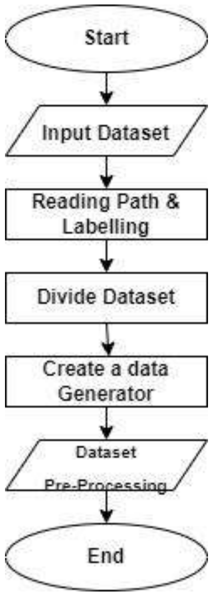

Fig. 4. Pre-Processing Datasets

The next stage is repairing or removing noise on the eye disease retinal image dataset. At this stage, an analysis of the dataset is carried out before being processed to the classification stage. In this study, only a few pre-processing stages were carried out including:

- Reading Path & Labelling

  The dataset located in Google Drive will have its file path and label taken. After that, the file path and label will be merged into one data frame. This data frame creation is used to make it easier for researchers in the next dataset sharing process.

- Divide Dataset

  After obtaining paths and labels, data separation will be carried out by dividing the dataset into three parts, there are training data, validation data, and test data. Researchers divided the training data by 0.8 or 80% with the remaining data of 0.2 or 20%. Followed by the division of validation data and test data which is divided based in the remaining data which is divided into validation data by 50% and test data by 50%. The division is done to facilitate data analysis in training, validation, and testing data.

- Create Data Generator

  At this stage researchers create data generators for training, validation, and testing. This stage is carried out so that the eye disease retinal image dataset can be entered into the AlexNet model for the training process. To adjust the AlexNet model, the size and color channel are determined. Determination of the size of each image is set to the expected size of (224,224) for adjustment to the model. Furthermore, the image input channel is initiated with the number three because the input that will be used in an RGB image. After determining the size and color channel, it is continued by determining the batch size for test data of 35. The last stage of this function is to create train, valid, and test generators.

D. CNN Classification Implementation

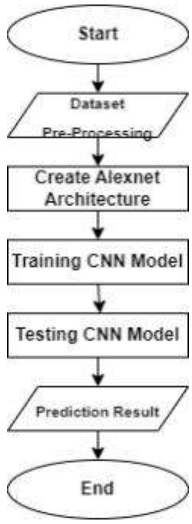

Fig. 5. Classification CNN

- Create AlexNet Architecture

  After the eye disease retinal image dataset is pre-processed, it will be continued by entering the eye disease retinal image data into the AlexNet model. In this research, AlexNet architecture is used for the process of classifying retinal images of eye disorders. Based on the architecture, there are more than 6 main types of layers in CNN, but only 3 types of layers are applied in this research, including :

  a) Convolution Layer

     In this research, 5 convolutional layers will be applied to the AlexNet model with different sizes including [10]:

     - Convolutional layer with size 11 x 11 using relu activation
     - Convolutional layer with size 5 x 5 using relu activation
     - Convolutional layer with size 3 x 3 using relu activation
     - Convolutional layer with size 3 x 3 using relu activation
     - Convolutional layer with size 3 x 3 using relu activation
     - Convolutional layer with size 3 x 3 using relu activation

  b) Max-Pooling Layer

In this research, there are 3 max-pooling layers with the same size as the AlexNet model, which is 3 x 3 with strides (2,2) [11].

c) Fully Connected Layer
In this research, 3 fully connected layers will be applied to the AlexNet model including [12]:
- fully connected layer with 4096 units and using relu activation
- fully connected layer with 4096 units and using relu activation
- fully connected layer with 4 units and using relu softmax

d) Training CNN Model
The next process after making the AlexNet model is to train the eye disease retinal image dataset on the AlexNet model. Training data that has been pre-processed will be continued with testing on the AlexNet model to obtain its accuracy.

e) Testing CNN Model
At this stage, the classification process will be carried out with new test data of 20% or using weights from the results of the training process to determine the prediction results and accuracy of the implemented CNN classification. The test data will be tested on the alexnet model that has been created. At this stage, researchers will get accuracy in processing eye disease retinal datasets using the CNN method with the alexnet model.

E. Evaluation

The evaluation process is carried out to measure the performance level of the classification model [13]. In this study, the confusion matrix was used to evaluate the accuracy of the CNN–AlexNet model in classifying eye diseases. The confusion matrix has a size of $n \times n$, where n represents the number of disease classes in the dataset [14]. This matrix provides detailed information on the model's predictive results through four categories: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). By analyzing these values, the overall performance of the classification system can be assessed more comprehensively, allowing identification of the model's strengths and weaknesses in detecting specific types of eye diseases [15].

III. RESULTS AND DISCUSSION

A. Dataset Analysis

This section presents a series of information related to the data analysis process, model implementation, training, and experimental results. Firstly, a comprehensive description of the dataset used in this research will be discussed. A good understanding of the dataset characteristics is a crucial step to validate the results and interpretations. The dataset consists of images depicting three types of eye diseases and normal eyes. The eye images are obtained using Funduscopy/Ophthalmoscopy techniques. The dataset, sourced from Kaggle, comprises a total of 4,217 images, with 1007 images for Glaucoma, 1098 for Diabetic Retinopathy, 1038 for Cataracts, and 1074 for normal eyes.

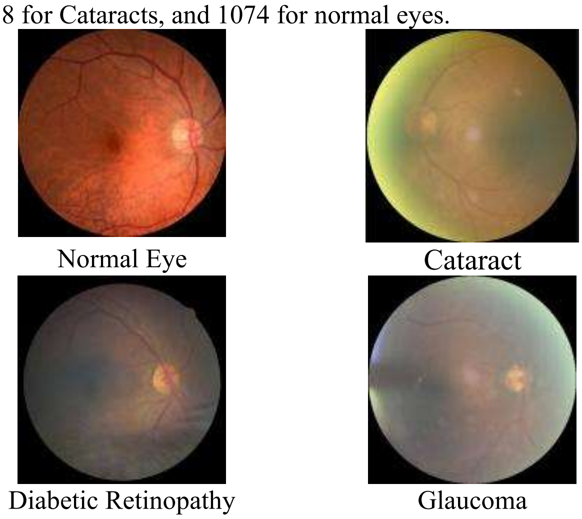

Fig. 6. Example image in each class

B. The Results of Codes Implementation

In the implementation process related to the experimental testing of eye disease classification using the AlexNet architecture, it was carried out in the Google Colab environment using the Python language. The following is an elaboration of the different sections of the code that has been developed.

- The Process of Splitting The Dataframe into Training, Validation, and Testing Subsets.

Fig. 7. Code of Splitting Process

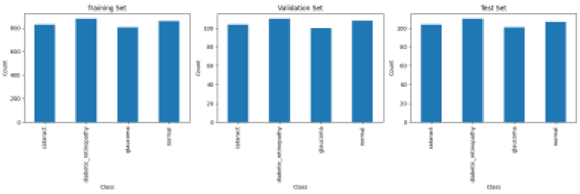

Fig. 8. Results of Splitting Process

The code above is designed to split an eye image dataset into training, validation, and testing subsets. First, the define_paths function is used to collect file paths and labels. Subsequently, a DataFrame is created by calling the define_df function. The separation process begins by stratified splitting of the dataset to ensure a balanced class distribution in the training subset (train_df). After that, the testing subset (dummy_df) is further divided into validation (valid_df) and testing (test_df) subsets with appropriate proportions. This function returns the three DataFrames for use in processing and training deep learning models.

- Defining and Implementing AlexNet Model

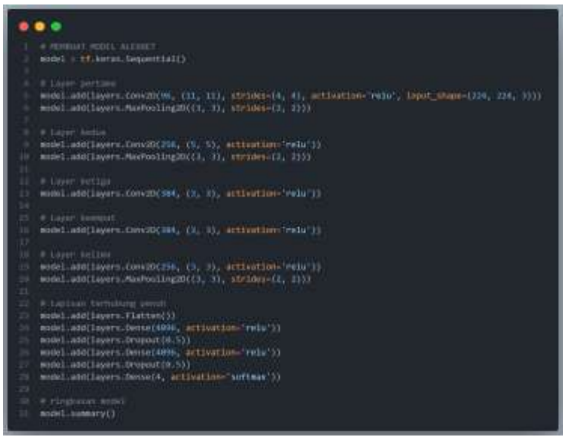

Fig. 9. Code of AlexNet's Model

The code above represents the implementation process of the AlexNet model architecture using TensorFlow and Keras for image recognition. This model consists of five convolutional layers and three fully-connected layers. The first convolutional layer has 96 filters with a (11, 11) kernel, followed by a MaxPooling layer. This process is repeated for the next two convolutional layers, each with an increasing number of filters. After the fifth convolutional layer, the output is flattened and passed through two Dense layers with dropout to reduce overfitting. The output layer has four neurons corresponding to the number of classes, with softmax activation. The model summary displays the parameters and dimensions of each layer, providing evidence that the AlexNet architecture has proven to be effective in image classification.

Below is the output summary of the aforementioned AlexNet architecture that has been executed. The summary illustrates the architecture of the AlexNet model that has been executed. This model consists of convolutional and fully-connected layers. The first convolutional layer uses 96 filters with a (11, 11) kernel, followed by a MaxPooling layer. This process continues for the next two convolutional layers, each with an increased number of filters. After the fifth convolutional layer, the output is flattened and passed through two fully-connected layers with dropout to reduce overfitting. The output layer has four neurons, corresponding to the number of classes, and utilizes softmax activation. The total model parameters are 21,597,572, encompassing trainable and non-trainable parameters.

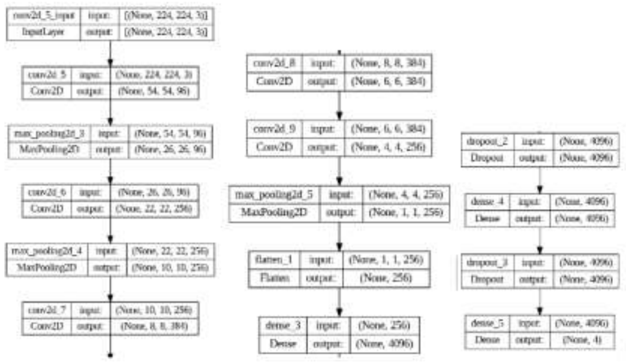

Fig. 10. Summary of AlexNet's Structure

The model is trained using the "fit" method with training data from "train_gen" and validation data from "valid_gen". The number of steps for each epoch during training and validation is determined by "steps_per_epoch" and "validation_steps". The training process is carried out for a total of 10 epochs. In the training results, there are metrics such as "loss" (the loss value on the training data), "accuracy" (model accuracy on the training data), "val_loss" (the loss value on the validation data), and "val_accuracy" (model accuracy on the validation data). For example, in the first epoch, the training data loss is 0.6195, with a training accuracy of 72.95%. The validation data loss is 0.5717, with a validation accuracy of 71.19%.

- Trainng Process with FeedForward and Backpropagation

Fig. 11. Code of AlexNet's Model

- Evaluating Trained Model

The evaluation process yields two main metrics: loss and accuracy. In the printed evaluation results, it is observed that for the test data, the loss value is 0.8362, indicating how well the model makes predictions. The accuracy on the test data is 75.18%, representing the percentage of successful classifications by the model on unseen data. This provides an overview of the model's generalization ability to data it has never seen during training. With an accuracy of 75.18%, the model can make accurate predictions on a majority of the test data. However, this evaluation can also serve as a guide for improving or optimizing the model if necessary.

Fig. 12. Model'sTraining Results

Fig. 13. In previous research using the MobileNetV2 architecture, a comparison can be drawn between the results of models using the AlexNet and MobileNetV2 architectures. Some important comparisons can be made. The AlexNet model achieves a test data accuracy of 75.18%, slightly higher than the accuracy of MobileNetV2 at 72%. Furthermore, the loss value on the test data for AlexNet is higher, indicating a higher error rate. Although MobileNetV2 has consistent evaluation metrics with accuracy, precision, recall, and F1-score all at 72%, indicating a good balance between precision and recall. It's worth noting that the running time or execution time for MobileNetV2 is 417.59 seconds, while the execution time information for AlexNet is not provided. The decision to choose a particular model should consider the trade-off between accuracy, performance, and computational needs, with specific considerations for the application at hand. While MobileNetV2 provides good results, the final decision will depend on specific priorities and application requirements.

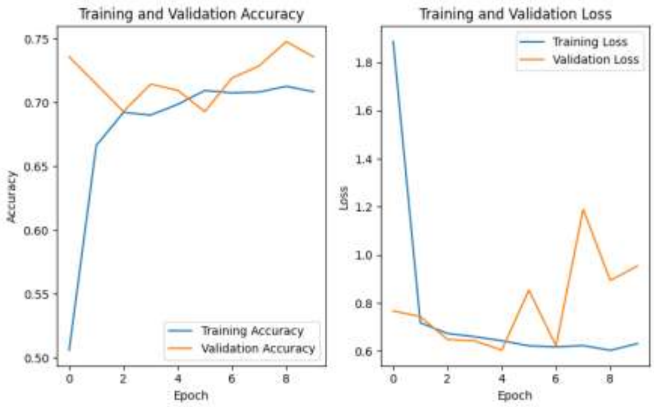

Fig. 14. Model'sTraining Results in Graph

IV. CONCLUSION

The model with the AlexNet architecture achieved accuracy on test data of 75.18%, slightly higher than MobileNetV2's accuracy of 72%. In addition, the loss function value on the AlexNet test data is higher, which can be interpreted as a larger error rate. Although MobileNetV2 has consistent evaluation metrics with accuracy, precision, recall, and F1-score of 72%, showing a good balance between precision and recall. It should be noted that the running time or execution time of MobileNetV2 reached 417.59 seconds, while AlexNet execution time information was not included. The decision to select a particular model must consider the trade-offs between accuracy, performance, and computational requirements, with particular consideration to the application in question. Although MobileNetV2 provides good results, the final decision will depend on the specific priorities and requirements of the individual application.

REFERENCES

[1] M. Ptito, M. Bleau, and J. Bouskila, "The retina: a window into the brain," *Cells*, vol. 10, no. 12. MDPI, p. 3269, 2021.

[2] K. Kansal and H. Khan, "Environmental factors and eye health: Protecting your vision in a changing world," *Int. J. Ophthalmol. Optom.*, vol. 5, pp. 33–35, 2023.

[3] G. Arslan and Ç. B. Erdaş, "Detection of cataract, diabetic retinopathy and glaucoma eye diseases with deep learning approach," *Intell. Methods Eng. Sci.*, vol. 2, no. 2, pp. 42–47, 2023.

[4] D. C. R. Novitasari, P. Wulandari, and D. Z. Haq, "Cervical Cancer Diagnosis System using Convolutional Neural Network ResidualNet," *Int. J. Comput.*, vol. 21, no. 1 SE-, pp. 61–68, Mar. 2022, doi: 10.47839/ijc.21.1.2518.

[5] J. Sanghavi and M. Kurhekar, "Ocular disease detection systems based on fundus images: a survey," *Multimed. Tools Appl.*, vol. 83, no. 7, pp. 21471–

21496, 2024.
[6] D. Marcella, Y. Yohannes, and S. Devella, “Klasifikasi penyakit mata menggunakan Convolutional Neural Network dengan arsitektur VGG-19,” *J. Algoritm.*, vol. 3, no. 1, pp. 60–70, 2022.
[7] A. E. Suwanda and D. Juniati, “Klasifikasi Penyakit Mata Berdasarkan Citra Fundus Retina Menggunakan Dimensi Fraktal Box Counting Dan Fuzzy K-Means,” 2022.
[8] J. Qiu, X. Lu, X. Wang, C. Chen, Y. Chen, and Y. Yang, “Research on image recognition of tomato leaf diseases based on improved AlexNet model,” *Heliyon*, vol. 10, no. 13, 2024.
[9] G. V. Doddi, “Eye Disease Retinal Images,” *Kaggle*, 2022. https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification.
[10] H. Eldem, E. Ülker, and O. Y. Işıklı, “AlexNet architecture variations with transfer learning for classification of wound images,” *Eng. Sci. Technol. an Int. J.*, vol. 45, p. 101490, 2023.
[11] T. Sivakumari and R. Vani, “Implementation of alexnet for classification of knee osteoarthritis,” in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 2022, pp. 1405–1409.
[12] S. Kavitha, B. Dhanapriya, G. N. Vignesh, and K. R. Baskaran, “Neural style transfer using vgg19 and alexnet,” in *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, 2021, pp. 1–6.
[13] G. Naidu, T. Zuva, and E. M. Sibanda, “A review of evaluation metrics in machine learning algorithms,” in *Computer science on-line conference*, 2023, pp. 15–25.
[14] S. Sathyanarayanan and B. R. Tantri, “Confusion matrix-based performance evaluation metrics,” *African J. Biomed. Res.*, vol. 27, no. 4S, pp. 4023–4031, 2024.
[15] D. Z. Haq and C. Fatichah, “Ultrasound Image Synthetic Generating Using Deep Convolution Generative Adversarial Network For Breast Cancer Identification,” *IPTEK J. Sci. Technol.*, vol. 34, no. 1, 2023, doi: http://dx.doi.org/10.12962/j20882033.v34i1.14968.