


Eye Disease Classification Using Convolutional Neural Network (CNN) with Web-based MobileNetV2 Architecture

Muhammad Fahriawan¹, Heny Pratiwi^{2*}, Bartolomius Harpad³

^{1,2*,3}STMIK Widya Cipta Dharma, Samarinda, Indonesia

Article Info	ABSTRACT
Keywords: Eye disease classification, Convolutional Neural Network, MobileNetV2, Flask, K-transfer learning	The high prevalence of preventable eye diseases, such as cataracts, glaucoma, and diabetic retinopathy, emphasizes the importance of accessible and efficient diagnostic solutions. This research aims to develop a web-based eye disease classification system using a lightweight Convolutional Neural Network (CNN) architecture, MobileNetV2, to overcome computational limitations in real-time applications. CRISP-DM methodology is applied, including dataset preparation, transfer learning with MobileNetV2 and VGG16, model evaluation, and implementation using Flask. The dataset from Kaggle consisting of 4,217 eye fundus images with four classes (cataract, glaucoma, diabetic retinopathy, and normal) was divided into 80% training, 10% validation, and 10% testing. Data augmentation and normalization were performed to improve model generalization. The results showed MobileNetV2 achieved the highest accuracy (90.14%) with low computational requirements, outperforming VGG16 (89.66%) and CNN (86.78%). MobileNetV2 displays balanced precision (89-99%), recall (74-96%), and F1-score (81-99%) across all classes, especially excelling in diabetic retinopathy detection. Its efficiency on resource-constrained environments makes it ideal for web integration. The developed Flask-based application allows users to upload images for instant classification, bridging the healthcare access gap. This research proves the effectiveness of MobileNetV2 in combining high accuracy and computational efficiency, offering a scalable solution for early screening of eye diseases, especially in remote areas.
This is an open access article under the CC BY-NC license 	Corresponding Author: Heny Pratiwi STMIK Widya Cipta Dharma, Samarinda, Indonesia henypratiwi@wicida.ac.id

INTRODUCTION

Classification of eye diseases using Convolutional Neural Network (CNN) is an innovative solution to overcome the limitations of manual diagnosis that relies on experts and specialized tools [1]. The system works by analyzing fundus images of the eye such as retinal texture, blood vessel distribution, or pathological indicators (e.g. lens opacities in cataracts or increased cup-to-disc ratio in glaucoma) and then maps them into specific disease categories based on learned patterns [2]. This classification not only speeds up screening but also reduces the subjectivity of conventional diagnoses [3].

Eye health is one of the most important aspects of human life, as visual impairment can affect a person's quality of life and productivity. Based on the World Health Organization

(WHO) report in 2020 [6], around 2.2 billion people in the world experience visual impairment or blindness, and nearly 1 billion of them can actually be prevented or treated with early detection and proper care. In Indonesia, eye health issues are still a major challenge. Data from the Ministry of Health (2021) [7] shows that cataracts are the leading cause of blindness in Indonesia, with prevalence increasing every year. Factors such as lack of access to health services, limited medical personnel specializing in ophthalmology, and high examination costs are major obstacles to the prevention and treatment of eye diseases.

Eye diseases such as cataracts, glaucoma and diabetic retinopathy require early detection to prevent more serious complications. Conventional diagnosis processes that rely on manual examination by ophthalmologists often have several limitations, including subjectivity in evaluation, reliance on specialized diagnostic tools such as fundusscopes and slit lamps, and relatively long examination times[4] . In addition, in remote areas with limited health facilities, access to eye examination services is more difficult, leading to many cases of visual impairment that are not diagnosed early. Therefore, the development of artificial intelligence (AI)-based technology can be an effective solution to help the screening and diagnosis of eye diseases more quickly, accurately, and easily accessible to the wider community.

In recent years, Convolutional Neural Network (CNN) technology has developed rapidly and is widely used in various medical image processing applications, including eye disease detection and classification [5]. CNNs have the ability to analyze visual features of eye fundus images, such as retinal texture, blood vessel distribution, and other pathological indicators, in order to classify eye disease types based on patterns learned from training data. The study conducted by Agmalaro et al. (2023)[8] showed that CNN models can identify images of patients with keratitis using datasets obtained from various sources, including images from slit lamps and smart devices, with an accuracy rate of 90%. In addition, research by Austrin et al. (2024) [9] also revealed that CNN-based models can be applied for eye disease classification on Optical Coherence Tomography (OCT) data, with results showing high effectiveness in distinguishing various types of eye diseases.

However, although CNNs have been proven to perform well in medical image classification tasks, their implementation in web-based platforms still faces several challenges. One of the main obstacles is the high computational requirements, especially in models with complex architectures such as VGG-16 [10], which have a large number of parameters and require significant processing power. These models are not always ideal to be applied to web applications that require fast response and minimal resource usage. Therefore, a CNN architecture that is more lightweight but still able to maintain high accuracy in eye disease classification is required.

MobileNetV2 is one of the CNN architectures specifically designed for computational efficiency, with fewer parameters than models such as VGG-16[8] , making it more suitable for web-based applications and resource-constrained devices. By using depthwise separable convolutions, MobileNetV2 is able to reduce the number of computations required without compromising classification performance[11] . Several studies have shown that MobileNetV2 can be effectively used in medical image classification tasks, including eye disease detection.

A study conducted by Pratama and Utaminingrum (2024)[10] developed a cataract severity detection system using the MobileNetV2-based CNN method, and the results showed an increase in computational efficiency without a significant decrease in accuracy.

In this study, a web-based eye disease classification system using MobileNetV2 architecture was developed to overcome the challenges involved in implementing CNN in real-time applications. This research aims to evaluate the performance of MobileNetV2 in classifying various types of eye diseases by considering factors such as accuracy, computational efficiency, and inference speed on a web-based platform [12]. With this system, it is hoped that the eye disease screening process can be carried out more quickly and easily accessible to the wider community, especially in areas with limited health facilities. In addition, the integration of lighter CNN technology with web-based applications can be an innovative step in bridging the gap between technology and health services, thus helping to reduce the number of blindness that can actually be prevented through more effective early detection.

METHODS

In this research, the CRISP-DM data mining methodology is used. This methodology consists of six stages, namely, Business Understanding, Data Understanding, Data Preparation, Modeling, Evaluation and Deployment. The stages can be seen in Figure 2.1

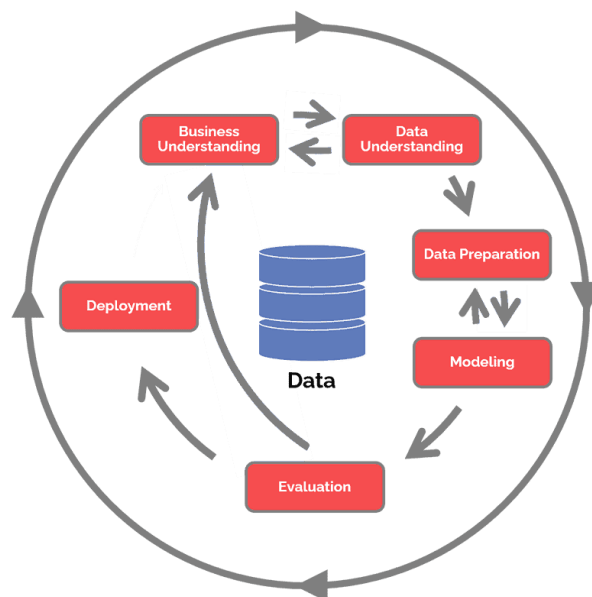


Figure 1. CRISP-DM methodology

a. Business Understanding

The first step is to understand the needs, define the specific objectives of the data analysis to be conducted, and determine the criteria that will be used to evaluate the success of the project.

b. Data Understanding

The second stage involves collecting data, describing the data, and evaluating the data.

c. Data Preparation

The third stage is data preprocessing or usually called data preparation that will be processed. The data preparation includes resizing the image according to the input size in the Convolutional Neural Network (CNN) model, then dividing the image dataset into 80% training, 10% validation and 10% testing. Then data augmentation is done to increase the diversity of data available for the training model, without the need to add new data.

d. Modeling

The fourth stage is to determine the data mining technique. This research utilizes a Convolutional Neural Network (CNN) model to perform image classification. In addition, the transfer learning method is applied by utilizing two popular pre-trained model architectures, namely MobileNetV2 and VGG16 to compare the accuracy between them. Transfer learning allows the use of pre-trained models on large datasets, such as ImageNet, which can save time and resources in training new models and improve model performance on smaller datasets. By comparing the accuracy of the two models, this project aims to find the architecture that gives the best results for this specific classification task.

e. Evaluation

In the fifth stage, the evaluation of each model includes accuracy, precision, recall, F1 score, and loss. Model evaluation is important to ensure that the model not only works well on training data but can also generalize well to unknown data. Accuracy measures how often the model's predictions match the correct labels. Precision measures how many of the positive predictions were actually positive, while recall measures how many of the overall positive cases were successfully detected by the model. F1 score is a harmonized metric between precision and recall, providing a single number that reflects the balance between the two.

f. Deployment

In the last stage, the model is saved in .h5 format, so that it can be implemented on the website. Then the website is created using the flask framework.

RESULTS AND DISCUSSION

Business Understanding

The application of data mining in this study relates to the classification of eye diseases using several popular Convolutional Neural Network (CNN) architectures, namely, MobileNetV2 and VGG16 with transfer learning and then the model is implemented into the website using the flask framework. The results of eye disease classification testing are in the form of disease class display, accuracy, and image visualization. This research aims to improve the accuracy of medical diagnoses, provide higher consistency than manual examination by humans, and build a clinical decision support system that can assist ophthalmologists in making faster and more precise diagnoses.

Data Understanding

This research uses an eye disease dataset sourced from Kaggle created by Guna Venkat Doddi. This dataset consists of various eye images with 4 classes namely cataract, glaucoma, diabetic retinopathy, and normal with a total of 4217 images that can be used to train and test eye disease classification models.

Data Preparation

At this stage, split the dataset by dividing the dataset with a split ratio of 80:10:10 into three subsets: 80% for training, 10% for validation, and 10% for testing. This division allows the model to be trained, validated, and tested effectively. The explanation of the 80:10:10 data split is as follows:

1. Training (80%): Most of the data is used to train the model, ensuring the model has enough information to learn the patterns in the data.
2. Validation (10%): The validation subset is used to test the model during training, assisting in hyperparameter tuning and preventing overfitting by providing useful feedback on model performance on data not seen during training.
3. Testing (10%): The testing subset is used for final evaluation of the model after training is complete, providing an estimate of the model's performance on completely new data.

At this stage, data augmentation is also carried out to balance the image data between classes, it is necessary to add image data or data augmentation. In this research, data augmentation is carried out using several techniques to ensure that the trained model can generalize well and not overfitting, namely :

1. Resizing
Resize the images to 256x256 to ensure that all images in the dataset have the same dimensions. This is important because deep learning models require fixed-size inputs.
2. Data Augmentation and Normalization
Normalization changes the range of image pixel values to 0 to 1 or -1 to 1. This helps the model converge faster during training by ensuring the data has a consistent scale, which speeds up learning and reduces bias. Meanwhile, horizontal flip flips the image horizontally, adding variety to the training data and improving the model's ability to recognize objects from different orientations. By applying these two techniques, the image data is optimally prepared to improve the accuracy and performance of the model.

Modeling

Table 1. Modeling

Hyperparameters	Type
Architecture	MobileNetV2, VGG16
Transfer learning database	ImageNet
Optimizer	Adam
Loss function	Categorical crossentropy
Number of epochs	50

At this stage, the hyperparameter configuration for the Convolutional Neural Network (CNN) model is done manually to achieve optimal performance. Table 3.1 lists the configurations used, namely MobileNetV2 and VGG16 architectures. MobileNetV2 was chosen for its efficiency, while VGG16 is known for its ability to detect complex features. The transfer learning database used is ImageNet to utilize existing knowledge, allowing the model to start with previously learned features. The optimizer applied is Adam, by adaptively adjusting the learning rate for each parameter. The loss function used is Categorical Crossentropy, for multi-class classification and assists in measuring the prediction error of the model. The number of epochs set was 50, ensuring the model was trained effectively without overfitting. This configuration was designed to compare the effectiveness of MobileNetV2 and VGG16 in detecting eye disease classification. The process of training results of the 3 models:

Table 2. Training Model

Model	Epoch	Accuracy	Val_accuracy	Loss	Val_loss
CNN	50	0.8678	0.8846	0.1877	0.2944
MobileNetV2	50	0.9014	0.9111	0.2076	0.4579
VGG16	50	0.8966	0.9207	0.1939	0.2201

Evaluation

Model evaluation

a) Accuracy and Confusion Matrix Convolutional Neural Network (CNN)

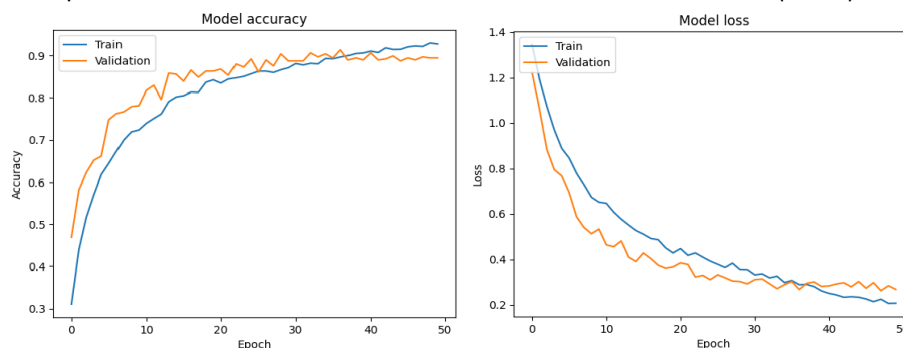


Figure 2. CNN Accuracy

Figure 2 shows the performance of the machine learning model over 50 epochs, with accuracy and loss for training and validation data. The training and validation accuracies increase rapidly at first and then stabilize, with the validation accuracy slightly higher and stabilizing faster than the training accuracy. Loss decreased dramatically in the early epochs and continued to decrease with little difference between training and validation, indicating a good fit. The log data shows a significant increase in accuracy from 0.5413 at epoch 1 to 0.9272 at epoch 50 for training, and from 0.6394 to 0.8846 for validation, as well as a decrease in loss reflecting the improved performance of the model. This indicates an effective learning model with a good balance between bias and variance, showing good performance on new data

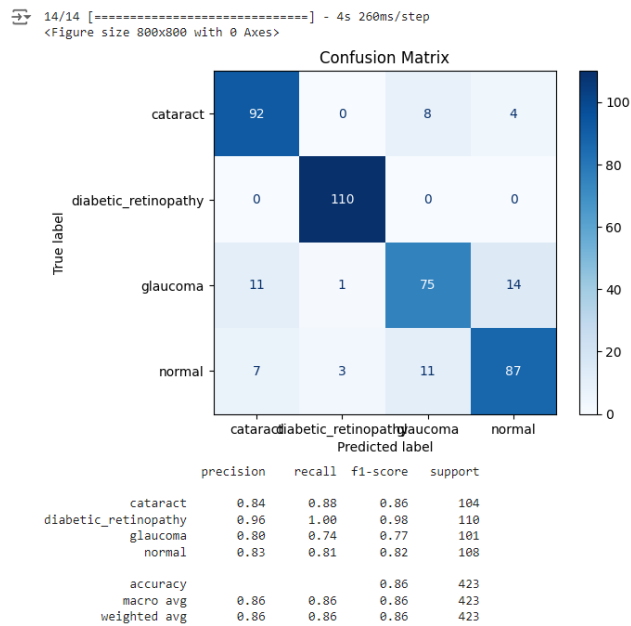


Figure 3. CNN Confusion Matrix

As for the CNN confusion matrix, Figure 3 shows that the classification model performed well, with an overall accuracy of 86%. The model is highly accurate in detecting diabetic retinopathy (precision and recall 0.96 and 1.00), while cataract and normal conditions also perform well (precision 0.84 and 0.83 recall 0.88 and 0.81). However, there was some misclassification between glaucoma and other categories, with a precision of 0.80 and recall of 0.74. Overall, the model showed a balanced and efficient performance in classifying the four eye conditions, according to the precision, recall, and F1-score metrics, which averaged 0.86.

b) Accuracy and Confusion Matrix MobileNetV2

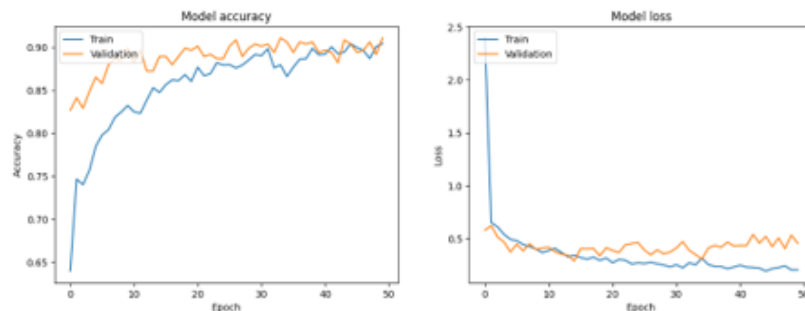


Figure 4. Accuracy of MobileNetV2

Figure 4 shows the training progression of the machine learning model over 50 epochs. Initially, the model starts with a fairly high loss and decent accuracy, but as the training continues, both loss and val_loss decrease, while accuracy and val_accuracy increase significantly. By the 50th epoch, the model achieved a loss of 0.2076 an accuracy of 0.9045 and val_accuracy of 0.9111, demonstrating substantial learning and generalization capabilities. The consistent decrease in loss and increase in val_accuracy indicate that the

model effectively learns from the training data without overfitting, as shown by the close tracking of training and validation accuracy.

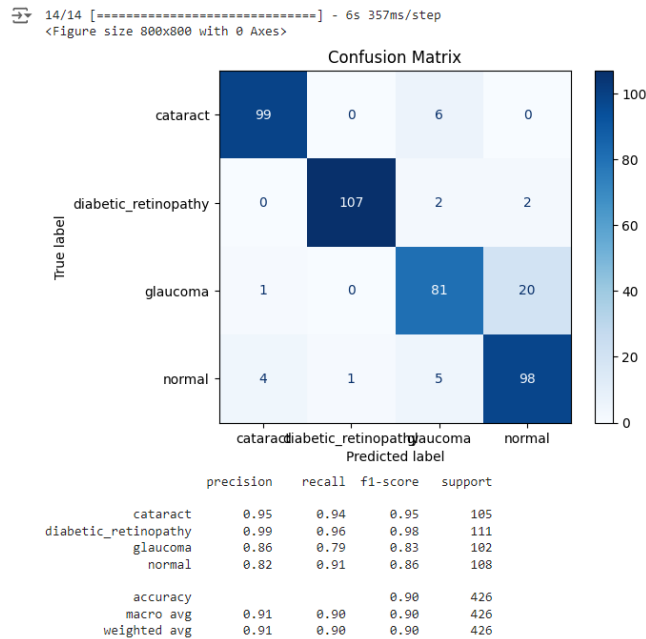


Figure 3. MobileNetV2 Confusion Matrix

The Confusion Matrix above for MobileNetV2 shows its performance in classifying eye diseases. The 'cataract' classification is very reliable with a precision of 0.95 and recall of 0.94, showing few false positives and a high true positive rate. 'diabetic retinopathy' has an almost perfect precision and recall of 0.99 and 0.96 respectively, indicating accurate identification with minimal misclassification. 'glaucoma' and 'normal' had lower but balanced precision and recall scores, indicating moderate reliability. Overall, the model achieved an accuracy rate of 90%, with macro average and weighted average precision, recall and f1 scores of around 89-90%, reflecting consistent performance across all categories.

c) Accuracy and Confusion Matrix VGG16

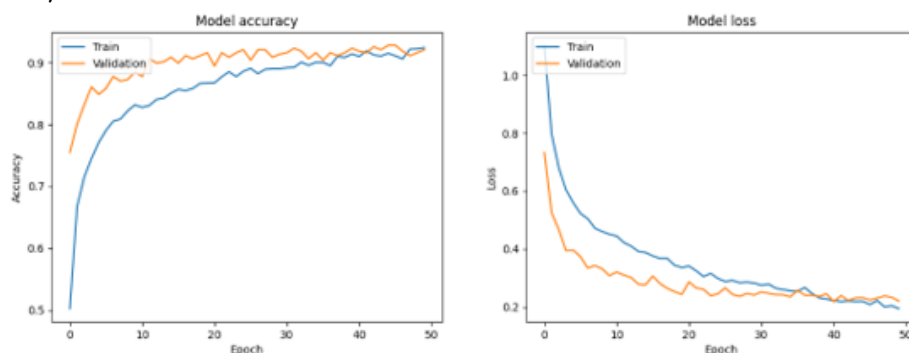


Figure 5. Accuracy of VGG16

In Figure 5 the training process shows the model improving its performance over time. Initially, at epoch 1, the model had a training accuracy of 50.21% and a validation accuracy of 75.48%. As the training progressed to epoch 26, both accuracies improved significantly with the training accuracy being 89.04% and the validation accuracy being 90.38%. This shows

that the model not only learns effectively from the training data but is also able to generalize well to new unseen data, as evidenced by the consistent increase in validation accuracy. The downward trend in loss values for both the training and validation sets further supports this improvement, indicating that the model becomes more precise in its predictions as it learns.

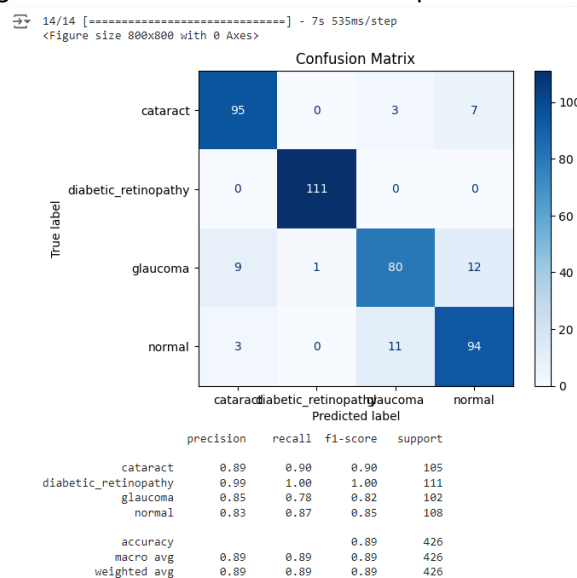


Figure 6 Confusion Matrix VGG16

In Figure 6 the confusion matrix VGG16 model used to classify eye conditions performed quite well with an overall accuracy of 89%. The model shows high precision, especially for cataract with a score of 0.89 and diabetic retinopathy with a perfect score of 0.99, meaning there are no false positives for these conditions. Recall was also high, with scores above or close to 0.78 for all conditions; diabetic retinopathy stood out with a recall of 1.00, indicating almost all cases were actually correctly identified by the model. The F1 score (weighted average of precision and recall) shows a balanced performance between precision and recall with scores ranging from 0.81 to 0.99. The support shows how many cases are available for each condition during testing, indicating a fairly even distribution of data among the different classes.

Evaluation of results

Based on the performance evaluation of the three models trained by Convolutional Neural Network (CNN), MobileNetV2, and VGG16, it can be concluded that MobileNetV2 showed the best performance with an overall accuracy of 90%. The model not only achieved consistently high precision, recall, and F1 score values, but also showed strong generalization ability on new data, with no significant signs of overfitting. MobileNetV2 excelled in classifying eye conditions with near-perfect precision and recall for diabetic retinopathy, as well as reliable performance for cataract, glaucoma, and normal conditions. Therefore, based on comprehensive evaluation data, MobileNetV2 was selected as the most suitable model for this study.

Deployment

After the evaluation stage is completed, the model files are saved in .h5 format to facilitate integration in the development of web-based machine learning applications. The .h5 format allows complete storage of the model structure and weights, making it easy to call and use the trained model. In this way, the saved model can be directly loaded and used to make classifications without the need to retrain. After that, the MobileNetV2 model was implemented into the website. This website was created using the flask framework. Inside the website there is a button to upload an image or image and below is the appearance of a website.

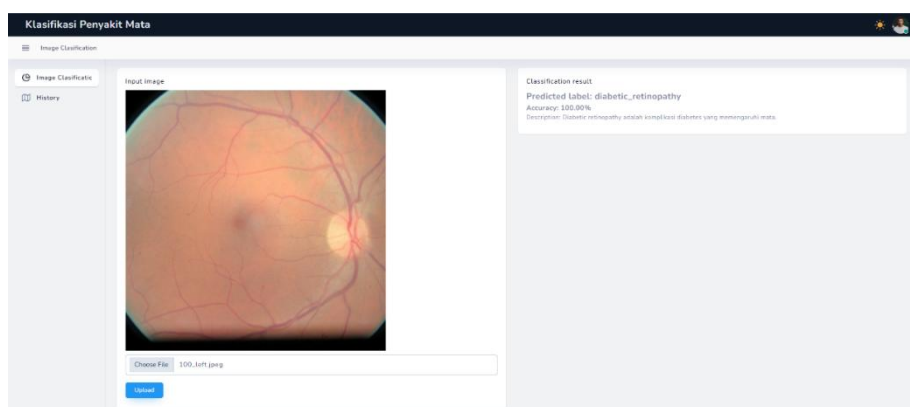


Figure 7. Deployment

CONCLUSION

Based on the analysis and discussion of eye disease classification, it can be concluded that the MobileNetV2 model achieves the highest accuracy of 0.9014 with a relatively low loss value of 0.2806. This model shows a better balance between training and validation accuracy and has advantages in computational efficiency. MobileNetV2 is also able to run processes quickly and efficiently, even in environments with limited computing resources, so it is suitable for applications that require high performance but limited computing power. With its solid performance and high efficiency, MobileNetV2 proved to be the most optimal choice in this study. After the training process is complete, the model is saved in .h5 format, which is the standard for storing deep learning models in Keras. Furthermore, this model will be implemented into the website with the flask framework.

REFERENCE

- [1] H. Pratiwi, *Buku Ajar Kecerdasan Buatan: Disertai Praktik Baik Pemanfaatannya*. Asadel Liamsindo Teknologi, 2024.
- [2] A. Refani, H. Pratiwi, dan Azahari, "Penerapan Metode K-Nearest Neighbor dalam Klasifikasi Penerima Bantuan Dana Desa pada Kecamatan Muara Ancalong," Skripsi, STMIK Widya Cipta Dharma, 2024.
- [3] A. A. Khair, H. Pratiwi, dan N. J. Saputra, "Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Penerima Beasiswa pada STMIK Widya Cipta Dharma," 2024.

- [4] B. Harpad, S. Salmon, dan R. M. Saputra, "Sistem Monitoring Kualitas Udara di Kawasan Industri dengan NodeMCU ESP32 Berbasis IoT," *Jurnal Informatika Wicida*, vol. 12, no. 2, hal. 39-47, 2022.
- [5] B. Harpad, A. Yusnita, dan H. Priscila, "Sistem Pendukung Keputusan Penentuan Ketua Koperasi pada Koperasi Leka Mandiri Menggunakan Metode Smarter Berbasis Web," *Jurnal Informatika Wicida*, vol. 13, no. 1, hal. 1-7, 2023.
- [6] World Health Organization, "World report on vision," 2020. [Online]. Available: <https://www.who.int/publications/i/item/9789241516570>. [Accessed: Apr. 2, 2025].
- [7] Kementerian Kesehatan Republik Indonesia, "Katarak Penyebab Terbanyak Gangguan Penglihatan di Indonesia," 2021. [Online]. Available: <https://sehatnegeriku.kemkes.go.id/baca/umum/20211012/5738714/katarak-penyebab-terbanyak-gangguan-penglihatan-di-indonesia/>. [Accessed: Apr. 2, 2025].
- [8] M. A. Agmalaro et al., "Implementasi Pendekatan Algoritma Deep Learning CNN untuk Identifikasi Citra Pasien Keratitis," *Jurnal Ilmu Komputer dan Agri-Informatika*, vol. 10, no. 2, hal. 164–175, 2023.
- [9] F. V. Austrin, J. Danil, R. I. Iman, M. R. Putri, dan P. Rosyani, "Klasifikasi Penyakit Mata Pada Data OCT Menggunakan Convolutional Neural Network (CNN)," *OKTAL: Jurnal Ilmu Komputer dan Sains*, vol. 1, no. 1, hal. 1–7, 2024.
- [10] A. A. Pratama dan F. Utaminigrum, "Sistem Pendeteksi Tingkat Keparahan Katarak Berdasarkan Citra Digital Menggunakan Metode U-Net dan CNN," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 3, 2024.
- [11] R. Firdaus, J. Satria, dan Baidarus, "Klasifikasi Jenis Kelamin Berdasarkan Gambar Mata Menggunakan Algoritma Convolutional Neural Network (CNN)," *Jurnal CoSciTech (Computer Science and Information Technology)*, vol. 3, no. 3, hal. 267–273, 2022.
- [12] Y. Rong et al., "Surrogate-assisted retinal OCT image classification based on convolutional neural networks," *Biomedical Signal Processing and Control*, vol. 71, hal. 103107, 2022.
- [13] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, dan D. Won, "Eye Disease Classification Using Deep Learning Techniques," *arXiv preprint arXiv:2307.10501*, 2023.
- [14] D. Hananta Firdaus, B. Imran, L. Darmawan Bakti, dan E. Suryadi, "Klasifikasi Penyakit Katarak pada Mata menggunakan Metode Convolutional Neural Network (CNN) berbasis Web," *Jurnal Kecerdasan Buatan Dan Teknologi Informasi (JKBTI)*, vol. 1, no. 3, hal. 18–26, 2022.
- [15] A. R. Prasetyo, Sussi, dan B. Aditya, "Analisis Perbandingan Algoritma Support Vector Machine (SVM) dan Convolutional Neural Network (CNN) untuk Sistem Deteksi Katarak," *Jurnal Ilmiah Teknik Mesin, Elektro Dan Komputer*, vol. 3, no. 1, hal. 1–10, 2023.
- [16] F. S. Putra, Kusriani, dan M. P. Kurniawan, "Deteksi Otomatis Jerawat Wajah Menggunakan Metode Convolutional Neural Network (CNN)," *JIFoTech (Journal of Information Technology)*, vol. 1, no. 2, 2021.

- [17] B. D. Satoto, M. I. Utoyo, R. Rulaningtyas, dan E. B. Koendhori, "Custom Convolutional Neural Network untuk Deteksi Penyakit Mata pada Citra Fundus Retina," *Jurnal Teknologi dan Sistem Komputer*, vol. 8, no. 2, hal. 93–100, 2020.
- [18] A. I. Pradana, R. W. Abdullah, dan Harsanto, "Deteksi Ketepatan Penggunaan Masker Wajah Dengan Algoritma CNN Dan Haar Cascade," *Jurnal Teknik Informatika Dan Sistem Informasi*, vol. 3, 2022.
- [19] F. N. Cahya, N. Hardi, D. Riana, dan S. Hadiani, "Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network (CNN)," *SISTEMASI: Jurnal Sistem Informasi*, vol. 10, no. 3, hal. 618–626, 2021.
- [20] D. Marcella, Y. Yohannes, dan S. Devella, "Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network Dengan Arsitektur VGG-19," *Jurnal Algoritme*, vol. 3, no. 1, hal. 60–70, 2022.
- [21] Kaggle, "Eye Diseases Classification Dataset," [Online]. Available: <https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification/data>. [Accessed: Apr. 2, 2025].