

Application of MobileNetV1 and DenseNet-121 CNN Architectures for Eye-Based Gender Classification

¹Sinta Nurhalifah, ²Akmal Junaidi, ³Didik Kurniawan, and ⁴M. Iqbal Parabi

^{1,2,3,4}Department of Computer Science, University of Lampung,
Jl. Prof. Dr. Ir. Sumantri Brojonegoro No. 1, Bandar Lampung, Lampung Province, Indonesia, 35145
e-mail: ¹sinta.nurhalifah21@students.unila.ac.id, ²akmal.junaidi@fmipa.unila.ac.id,
³didik.kurniawan@fmipa.unila.ac.id, ⁴iqbal.parabi@fmipa.unila.ac.id

Abstract - Gender classification is an important field in biometric identification systems that plays a vital role in security, forensics, and human-computer interaction. Human eye images are a promising visual object for gender classification because they contain distinct anatomical features that differ between males and females. This study aims to implement and evaluate two Convolutional Neural Network (CNN) architectures, namely MobileNetV1 and DenseNet-121, for gender classification based on human eye images. The dataset used was obtained from the Kaggle platform, consisting of 11,525 eye images, with 6,323 male and 5,202 female samples. The research process involved several stages, including pre-processing, data splitting, augmentation using Affine transformations (rotation and translation), as well as model training and evaluation. The evaluation was conducted using accuracy, precision, recall, and F1-score metrics. The results showed that both architectures were capable of performing gender classification effectively, although differences in performance were observed. The best accuracy was achieved by MobileNetV1 with a rotation scenario of 92.49%, while DenseNet-121 obtained 86.84% with a combined rotation and translation scenario. This research is expected to contribute to the development of efficient and accurate eye image-based gender classification systems using deep learning approaches.

Keywords: Convolutional Neural Network; DenseNet-121; Eye Image; Gender Classification; MobileNetV1.

1. INTRODUCTION

The eye is an essential organ that stores crucial data, including gender, in addition to aiding with vision [1]. Due to the physical differences between male and female eyes, biometric research indicates that the eye can be utilized as a categorization item [2]. In the context of contemporary technology, deep learning using Convolutional Neural Networks (CNN) and artificial intelligence (AI) have demonstrated efficacy in accurately classifying visual objects [3][4]. Unique biological traits like fingerprints, face features, iris, or periocular region are frequently used in biometric systems to perform personal identification [5]. Because eye images contain periocular cues and complex textures that vary by gender, using the human eye for gender detection is seen to be promising [6]. The ability of deep learning models to capture these minor distinctions has also been highlighted in a number of publications [7][8].

CNNs have been shown to be useful at classifying ocular images in earlier research. In their comparison of CNNs and Haar Cascade Classifiers, Pradana & Wijiyanto [9] achieved a 92% accuracy rate. CNNs were able to classify gender based on eye scans with 93% accuracy, according to research by Wicaksono & Shidiq [10]. Pre-trained architectures with accuracies ranging from 80% to 93%, including VGG16, ResNet152V2, and EfficientNetV2L, were employed in other investigations [11]. This demonstrates that CNNs are superior to more conventional techniques like K-Nearest Neighbor and Naive Bayes [12]. Furthermore, deep CNNs applied to iris patterns have demonstrated exceptionally high recognition ability, with up to 98.8% accuracy [13]. Similarly, studies using periocular regions reported accuracies of 99% in gender classification [8]. These findings further justify why eye-based images are suitable for gender prediction tasks. Based on these findings, this study applies the MobileNetV1 and DenseNet-121 architectures to classify gender based on eye images. Both models were chosen due to their computational efficiency and high performance in image pattern recognition [14] [15].

2. RESEARCH METHODOLOGY

This research was conducted through several systematic stages so that the results obtained could be measured, tested, and evaluated. The research stages were arranged based on a machine learning-based research flow that is often used in digital image classification. The research flow includes dataset collection, pre-processing, data partitioning, data augmentation, model design, model training, and result evaluation [16][17]. The research stages can be seen in Figure 1.

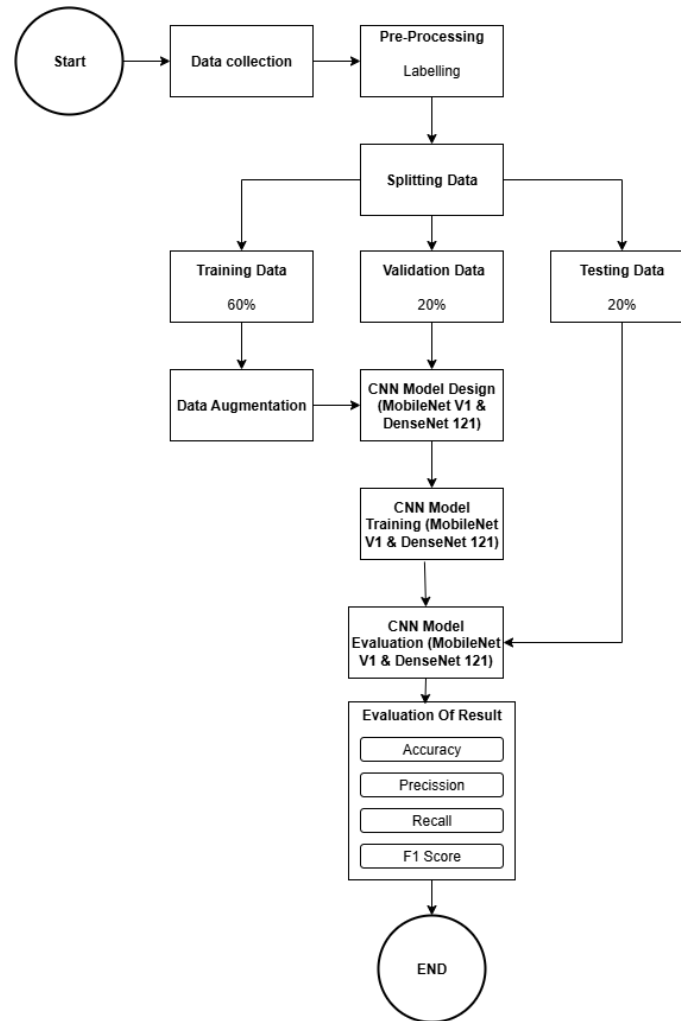


Figure 1. Research workflow.

This research began with the collection and division of the eye image dataset into training and testing subsets. This division was important to ensure objective evaluation and avoid any indication of overfitting [6]. Data augmentation was then performed on the training data. Next, the CNN model was trained using MobileNetV1 and DenseNet-121, architectures known for their efficiency and high recognition accuracy [14][15]. MobileNetV1 offers a lightweight design and strong performance even on limited devices [8]. DenseNet-121, with its dense layer connections, improves feature reuse and captures detailed eye patterns [13]. After model training was complete, the model was evaluated using accuracy, precision, recall, and F1-score metrics [7]. These metrics ensure performance assessment beyond mere accuracy.

2.1. Research Data

The dataset used in this study was taken from Kaggle [18] with a total of 11,525 eye images, consisting of 6,323 male eye images and 5,202 female eye images. The data was then divided into three parts: 60% for training, 20% for validation, and 20% for testing. This division was intended to ensure optimal model training,

with the validation data preventing any indication of overfitting and the test data measuring performance objectively. Details of the data division for each class are presented in Table 1.

Table 1. Dataset distribution.

Category	Number of Image
Man	6.323
Woman	5.202
Total	11.525

Table 1 shows the dataset used in this study, consisting of 6,323 images of male eyes and 5,202 images of female eyes, for a total of 11,525 images. This large dataset provides an important basis for model training because it provides a wider variety of visual patterns, both in terms of anatomy and image characteristics and patterns. To provide a clearer picture of the data used in this study, sample images of eyes from each category, namely men and women, are displayed. This visualization aims to help readers understand the anatomical differences and visual variations that form the basis of classification, while also showing how the raw data collected from the dataset is used in the model training stage. The following are sample images of eyes in each category, which can be seen in Figure 2.

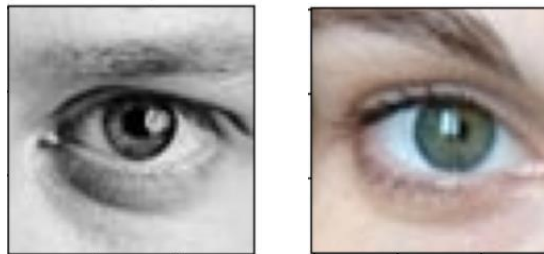


Figure 2. Example of male and female Eye Dataset (Pavelbiz/Eyes-rtte/Kaggle).

Figure 2 shows an example dataset of eye images for the male and female eye categories used. This visualization not only shows the differences in image representation between the two classes, but also provides insight into the variations in eye shape, lighting, and anatomical details that form the basis of the model's visual pattern learning. In addition, this visualization also helps illustrate the diversity of the data samples, which is very important for improving model generalization and reducing bias in classification results.

2.2. Data Preprocessing

During the first stage of this project, images are labeled as femaleeyes and maleeyes. These labels designate the corresponding images for females and for males. Then, the images are resized to 224×224 pixels to match the input of the CNN architecture used. Following this, pixel values are normalized to a specific range. As a result, the training process becomes faster, and the model's performance becomes more stable [19]. This step is crucial to avoid any discrepancies in new data and model training, especially for images of different dimensions or with varying pixel values. Uniform formatting ensures consistent extraction of visual features, mitigating classification errors.

Standard normalization and resizing of images help to achieve stability and convergence of the model during training. Uniform preprocessing enhances comparability across features, which is especially vital in biometric systems. CNN models may misinterpret the variations as gender-related patterns due to lighting or resolution in the images. Hence, consistent preprocessing becomes crucial in eye-based gender recognition for attaining precise classification outcomes.

2.3. Data Splitting

The division of the data into 75% training data and 25% test data ensures the ability of the model to generalize. This split is useful in the sense that the model is capable of recalling the patterns learned during training and

can also be tested on data that has not been previously encountered. With validation data, the training process can be controlled to avoid overfitting [17] and the test data can be used to assess the model's performance in an objective manner [20]. This approach to data division is frequently encountered in research, especially those that use machine learning, as it captures the effectiveness of its application in real-life situations, thus providing a more accurate and dependable estimation of the model's performance [16]. The number of images used in the described dataset is illustrated in Table 2.

Table 2. Dataset distribution.

Data Types	Male	Female	Total
Training Data	3794	3121	6915
Validation Data	1264	1041	1265
Testing Data	1040	2305	2305

In Table 2, the dataset is divided into 60% training, 20% validation, and 20% testing. Augmentation is applied only to the training set to prevent variation from affecting the validation or testing data [17]. This strategy ensures that the evaluation results reflect the model's true generalization ability [20].

2.4. Data Augmentation

To increase data variation, affine transformations in the form of 20° rotation, 10-pixel translation, and combinations thereof were applied [16][17]. The data distribution is summarized in Table 3.

Table 3. Number of data before and after augmentation.

Data Type	Amount of Data
Training Data (Before Augmentation)	6.915
Data Validation	2.305
Data Testing	2.305
Training Data (After Augmentation)	13.830

Table 3 shows the distribution of the amount of data used at each stage of the model training process. Before augmentation, the amount of data at the training stage was 6,915 images. Next, the dataset was divided into 2,305 images for the validation stage and 2,305 images for the testing stage. Meanwhile, examples of image results after the augmentation process are shown in Figure 3.

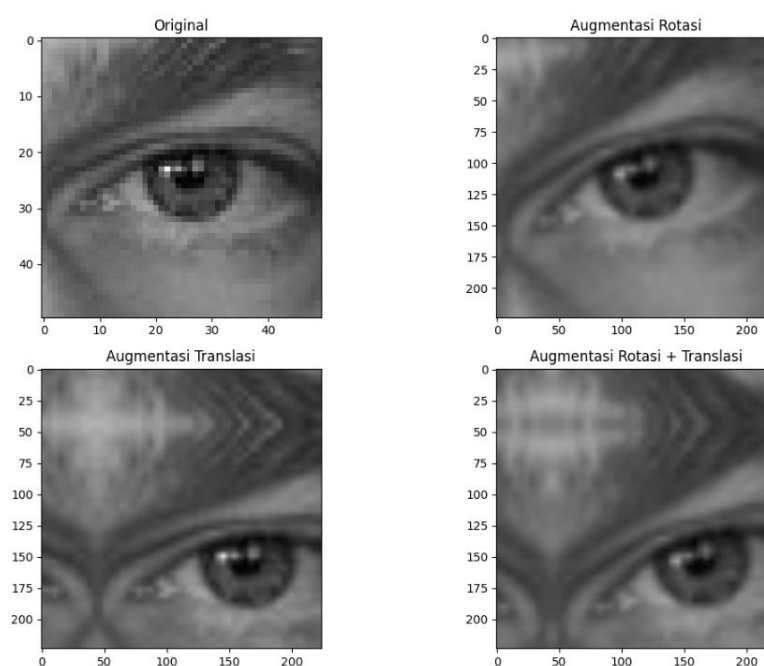


Figure 3. Augmentation example (Pavelbiz/Eyes-rtte/Kaggle).

Figure 3 shows the results of image augmentation using affine transformation with three types of scenarios, namely rotation, translation, and a combination of rotation and translation. The image on the top left is the original image used before augmentation. The top right shows the augmentation results with a 20-degree rotation. Meanwhile, the bottom left shows the augmentation results with a 10-pixel translation to the right and a 10-pixel translation down, while the bottom right shows the results of the combination of rotation and translation augmentation.

2.5. Model Design

The architectures used are MobileNetV1, which uses separable convolutions based on depth for computational efficiency [14], and DenseNet-121, which uses dense connections between layers to strengthen feature propagation [15]. MobileNetV1 is a lightweight model that can reduce parameters and training time without significant loss of performance. Conversely, DenseNet-121 utilizes dense connections to improve feature reuse, reduce gradient vanishing, and enhance representation quality. These two models provide a comprehensive comparison between lightweight and deeper architectures for eye image classification, emphasizing the importance of choosing an approach that balances efficiency and accuracy in biometric applications.

2.5.1 MobileNetV1

MobileNetV1 uses separable convolutions to speed up the process and lower the number of parameters [14]. The main benefit of MobileNetV1 is its ability to cut down on computational complexity while still retaining effective feature extraction. Thanks to its efficient design, this architecture delivers solid classification performance. You can see a full visualization of the MobileNetV1 architecture used in this study in Figure 4.

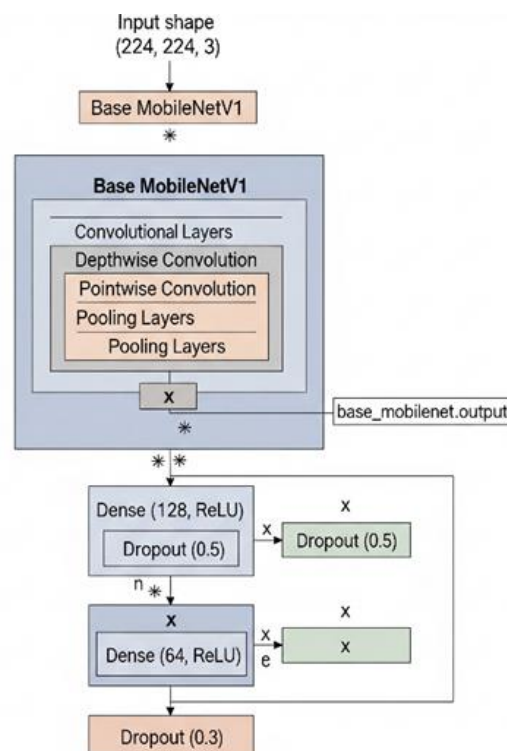


Figure 4. MobileNet-V1 architecture.

Figure 4 shows the MobileNetV1 architecture with deep separable convolutions as the main feature extractor. The extracted features are processed through two dense layers of 128 and 64 neurons with ReLU activation, combined with dropout layers of 0.5 and 0.3 to reduce overfitting. The integration of efficient feature extraction and structured classification layers is designed to achieve accurate gender recognition from eye images while maintaining stability on new data.

2.5.2 DenseNet-121

The DenseNet-121 architecture uses dense connections between layers to maximize feature reuse and improve gradient flow [15]. This design improves feature propagation, reduces gradient vanishing issues, and maintains parameter efficiency while producing richer feature representations [15][21]. With dense blocks and transition layers, the network achieves better classification performance even with fewer parameters compared to other deep architectures [14][15]. In this study, DenseNet-121 is expected to capture fine details in eye images and improve training stability. The DenseNet-121 architecture used in this study can be seen in Figure 5.

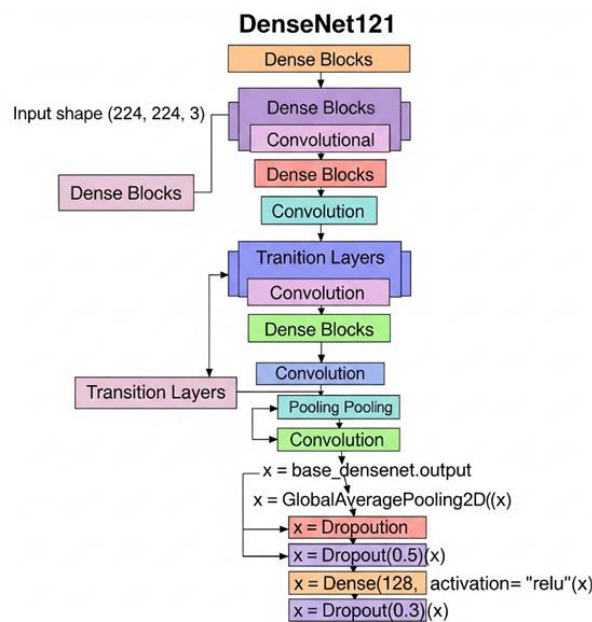


Figure 5. DenseNet-121 architecture.

Figure 5 shows the architecture of DenseNet-121, which consists of 121 layers arranged in dense blocks and transition layers to perform gradual and comprehensive feature extraction. Each dense block allows direct information propagation between layers, resulting in richer and more efficient feature representations. After the feature extraction process, the network output is processed using Global Average Pooling, which reduces the data dimension without losing important information. Next, a dropout layer is added as a regulation strategy to prevent overfitting. The final stage is a dense layer with 128 neurons activated by ReLU, which plays a role in strengthening generalization and improving classification performance.

2.6. Model Training

Before evaluation, model training was conducted. The training process was carried out until convergence by monitoring the accuracy, precision, recall, and F1-score metrics [21]. With this design, it is expected that the model will be able to achieve a balance between high accuracy and good generalization capabilities, so that the eye image classification results can be more reliable and consistent on previously unseen test data. The parameters used in the CNN model training process are shown in Table 4.

Table 4. CNN Model Training Parameters.

Parameter	Description
Optimizer	Adam
Learning rate	0.0001
Epoch	10
Loss	Catgorical crossentropy

Table 4 shows that model training used the Adam optimizer with a learning rate of 0.0001. The training took place over 10 epochs, which let the model gradually improve its parameters and reduce overfitting. The team

applied the categorical cross-entropy loss function because it works well for multi-class classification tasks. This ensures good probability estimates of class labels. These hyperparameters were selected to keep training stable while helping the model effectively extract important features from eye images. This setup offers a balanced approach to performance and efficiency, aiding accurate gender classification in this study.

2.7. Model Evaluation

Model evaluation was done using testing data. Prediction results were examined with a confusion matrix, accuracy, precision, recall, and F1-score [20]. To assess the performance of the classification model created with the Convolutional Neural Network (CNN) architecture in this study, a detailed analysis was conducted using the confusion matrix method. This approach helps identify patterns in prediction errors, like the model's tendency to misclassify specific classes. Additionally, using several evaluation metrics ensures that the model is judged not just on overall accuracy, but also on its ability to consistently differentiate between male and female eyes. Figures 6 shows the results of testing the MobileNet and DenseNet models in classifying gender based on eye images under the best scenario results. These confusion matrices demonstrates how effectively each model can correctly classify male and female eye images. Confusion matrix is intended to provide an overview that can be assessed or improved in the future.

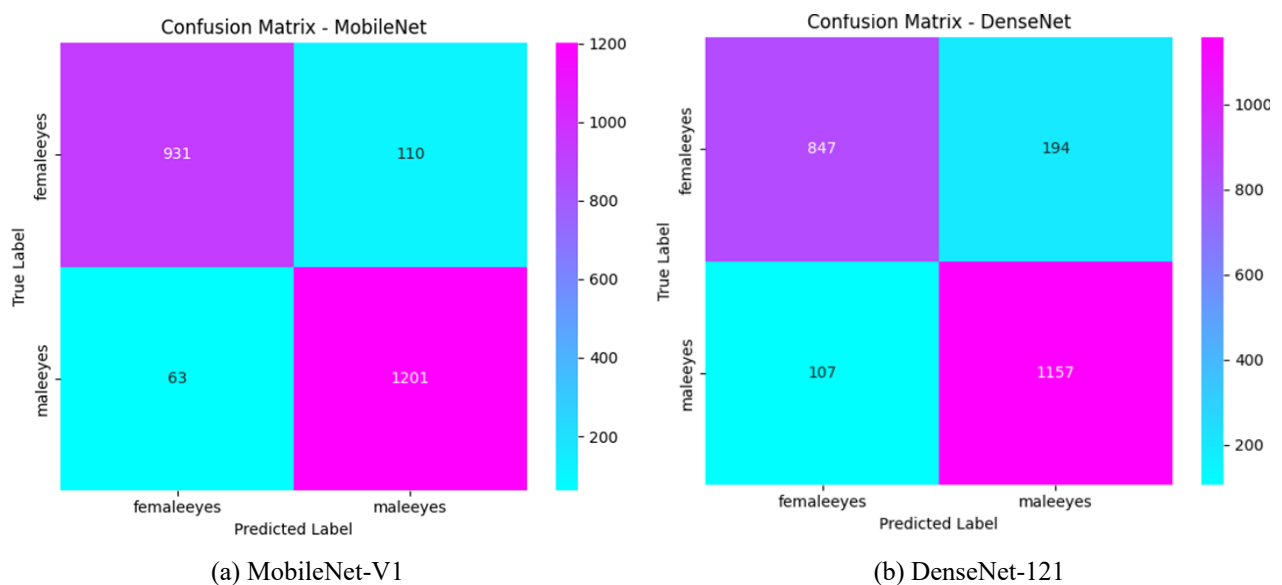


Figure 6. (a) Confusion matrix MobileNet V1 and (b) Confusion matrix DenseNet-121.

Figure 6a shows the confusion matrix of the MobileNetV1 model in the best-case scenario, which uses rotation augmentation. This model can classify images of male and female eyes quite well. The high number of correct predictions on the main diagonal reflects this success, with 931 for the female class and 1201 for the male class. These values indicate that most images are recognized correctly according to their class. However, some prediction errors exist in certain classes, pointing to the model's limitations in distinguishing images with similar visual features. This finding implies that more preprocessing or augmentation techniques may be needed to lower the classification error rate.

Figure 6b shows the confusion matrix for the DenseNet-121 model in the best-case scenario, which combines rotation and translation. This model performs well, achieving 847 correct predictions for the female class and 1,157 for the male class. However, the number of prediction errors in the female class is higher than in the MobileNetV1 model. This suggests that while DenseNet-121 captures more complex feature representations through its dense interlayer connections, it still struggles to differentiate between female eye images that look similar to those of the male class. This finding confirms that unbalanced data distribution can impact the model's sensitivity to minority classes.

3. RESULT AND DISCUSSION

3.1. Experimental Results

The experiment was conducted on two CNN architectures, MobileNetV1 and DenseNet-121. Both were trained using affine augmentation scenarios, namely rotation, translation, and their combinations, to increase data variation and generalization [22]. Performance evaluation was conducted on test data using accuracy, precision, recall, and F1-score [23] to ensure a balanced assessment beyond overall accuracy.

The results showed that MobileNetV1 achieved the highest accuracy of 92.49% in the rotation scenario, demonstrating strong adaptability to changes in orientation. DenseNet-121 performed best with an accuracy of 86.84% in the combined rotation and translation augmentation scenario, demonstrating better utilization of complex variations but generally having lower accuracy than MobileNetV1. These findings highlight that MobileNetV1 excels in efficiency and rotation augmentation, while DenseNet-121 is more adaptable to combined rotation and translation augmentation transformations.

Table 6. CNN model evaluation results.

Architecture	Scenario Augmentation	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
MobileNetV1	Without Augmentation	91.02	0.91	0.91	0.91
MobileNetV1	Rotation	92.49	0.87	0.87	0.87
MobileNetV1	Translation	90.76	0.93	0.92	0.92
MobileNetV1	Rotation + Translation	85.86	0.87	0.85	0.85
DenseNet-121	Without Augmentation	86.59	0.91	0.91	0.93
DenseNet-121	Without Augmentation	85.47	0.87	0.86	0.86
DenseNet-121	Without Augmentation	85.86	0.87	0.86	0.86
DenseNet-121	Without Augmentation	86.84	0.88	0.87	0.87

Table 6 presents the results of the CNN model evaluation on two architectures, MobileNetV1 and DenseNet-121, with various augmentation scenarios performed. On the MobileNetV1 architecture, the best results were obtained in the rotation augmentation scenario with an accuracy of 92.49%, although the precision, recall, and F1-score were lower (0.87) than in the translation scenario. Meanwhile, in the translation scenario, MobileNetV1 showed better metric balance, marked by a precision of 0.93, a recall of 0.92, and an F1-score of 0.92. This indicates that MobileNetV1 is stronger at translation in maintaining detection consistency between classes. Conversely, in the combined rotation and translation scenario, MobileNetV1's performance declined with an accuracy of 85.86% and other evaluation metrics relatively lower, indicating that significant geometric distortion can affect model stability.

In the DenseNet-121 architecture, the best performance was achieved in the scenario without augmentation, with an accuracy of 86.59% and an F1-score of 0.93, indicating that the model is quite effective on raw data without additional transformations. However, when given rotation or translation augmentation, the accuracy dropped to around 85%, with precision, recall, and F1-score ranging from 0.86 to 0.87. This indicates that DenseNet-121 is less adaptive to data variations generated from simple augmentation. Interestingly, in the combined rotation and translation scenario, DenseNet-121 performs slightly better than in the single augmentation scenario with an accuracy of 86.84%, although the F1-score is only 0.87.

Overall, these results show that MobileNetV1 excels in terms of accuracy, especially in the rotation augmentation scenario, while DenseNet-121 performs more consistently on original data without augmentation. This difference indicates that the selection of CNN architecture must consider the type of augmentation and data characteristics to achieve the best balance between accuracy and generalization. The training performance of the MobileNetV1 model with rotation augmentation is visualized through the accuracy and loss graphs in Figure 7.

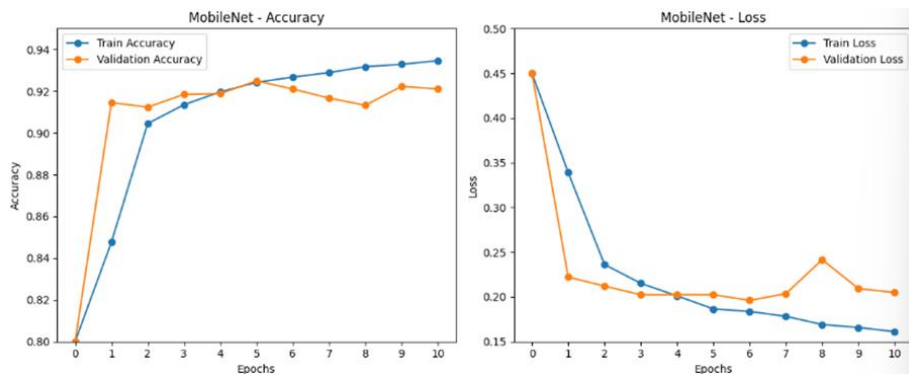


Figure 7. Graph of accuracy and loss of rotation augmentation of MobileNetV1.

Figure 8 shows that training and validation accuracy increase with the number of epochs, while loss decreases consistently, indicating stable training without overfitting. This proves that the model effectively learns patterns while maintaining generalization, supported by appropriate parameters such as learning rate and epoch [14][21]. The accuracy and loss graphs for the DenseNet-121 architecture can be seen in Figure 8.

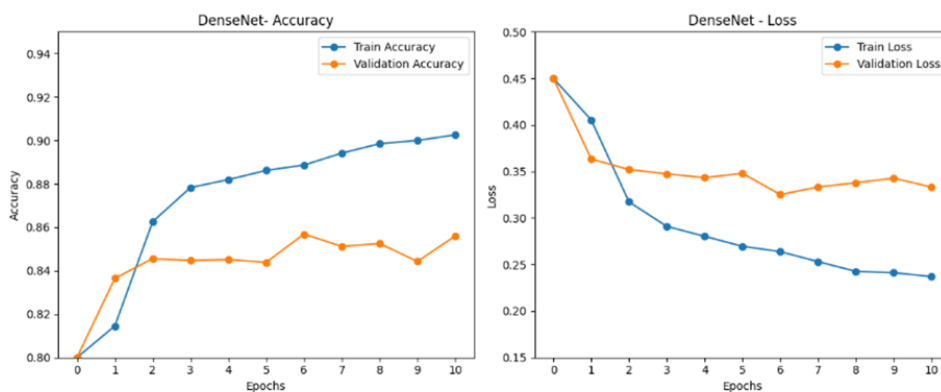


Figure 8. Graph of accuracy and loss of rotation and translation augmentation of DenseNet-121.

Figure 8 demonstrates that training accuracy continuously improves, that means this DenseNet-121 model can sufficiently learn training samples' patterns. But val acc seems to fluctuate more and stay around 86-87% So it looks like a gap between training and val acc. This difference indicates that the model performs well on the training set, but generalization on an unseen dataset is still not enough. Training loss decreases consistently, while validation loss tends to stagnate and sometimes increases, reflecting sensitivity to data variation. This trend highlights the need for fine-tuning and augmentation strategies to improve the stability and generalization ability of the model.

3.2. MobileNetV1 Performance Analysis

MobileNetV1 has proven to be superior in parameter efficiency and training speed. Deep separable convolutions enable effective pattern recognition despite their relatively small size [24]. MobileNetV1 has shown that it is suitable for devices with limited resources while maintaining high accuracy [14]. In addition, MobileNetV1 performs well under varying conditions, reinforcing its reliability for modern biometric systems [13].

3.3. DenseNet-121 Performance Analysis

DenseNet's performance is lower than MobileNet's, but it still demonstrates the model's effectiveness in classifying images because DenseNet is a convolutional network that connects each layer to all previous layers through dense connections, so that each layer receives input in the form of features from all previous layers.

This approach strengthens information and gradient propagation, encourages feature reuse, and significantly reduces the number of parameters required compared to traditional CNN architectures [15].

3.4. Comparison with Previous Research

The findings of this study were compared with several other studies that also utilised the same dataset. This decision was made to illustrate the level of differences in performance (of both approach and numerical value), rather than to suggest any of these algorithm be false; they are in fact not. This comparison highlights the advantages of this study and how its results stack up compared with previous reports. Furthermore, in this work we review some of the items that affect the model performance (CNN architecture, the use of preprocessing techniques, and data augmentation). As a final note, for an overview comparison of the results of this study and earlier studies that provide information about method, data and accuracy level used, see Table 7.

Table 7. Comparison of research results with previous research.

Researchers	Scenario Data Augmentation	Method	Accuracy (%)
Patil & Hangarge (2023)	-	Custom CNN	93.08
Wicaksono & Shidiq (2024)	-	KNN	81.00
Pradana & Wijiyanto (2024)	-	CNN + Haar Cascade	92.00
Research conducted	Without Augmentation	MobileNetV1	91.02
Research conducted	Rotation	MobileNetV1	92.49
Research conducted	Translation	MobileNetV1	90.76
Research conducted	Rotation + Translation	MobileNetV1	85.86
Research conducted	Without Augmentation	DenseNet-121	86.59
Research conducted	Rotation	DenseNet-121	85.47
Research conducted	Translation	DenseNet-121	85.86
Research conducted	Rotation + Translation	DenseNet-121	86.84

Based on Table 7, the accuracy of this study using MobileNetV1 with the best accuracy using the rotation augmentation scenario is 92.49%, slightly lower than the study by Patil & Hangarge, which reached 93.08% [11]. This difference can be explained by the fact that the Custom CNN architecture used by Patil & Hangarge was designed with a more specific number of filters and layers, enabling it to extract features with higher complexity [4][3]. Meanwhile, MobileNetV1 uses depthwise separable convolution, which in principle reduces the number of parameters and computational operations, making it more efficient but potentially losing some detailed information during the feature extraction process [14].

On the other hand, this study obtained higher accuracy compared to the study by Wicaksono & Shidiq, which only reached 81.00% with the KNN method [10]. This can be explained because KNN only relies on proximity in feature space for classification, without a deep representation learning process [12]. In contrast, MobileNetV1 has the ability to learn features through multilevel convolution layers, making it more effective in recognizing complex patterns in eye images [9].

When compared to the research by Pradana & Wijiyanto, which obtained 92.00% using the CNN + Haar Cascade method [9], the results of this study remain competitive. This small difference may occur because the use of Haar Cascade helps improve the initial detection stage of objects (region of interest), so that the input received by CNN is more focused [1][2]. However, the advantage of MobileNetV1 lies in its architectural efficiency, making it a more suitable choice for implementation on devices with computational limitations [14][17].

3.5. Missed Classification Analysis

Missed classifications were mainly found in images with low lighting conditions and images of eyes with similar anatomical features between genders, a phenomenon also observed in periocular biometric research, where recognition performance can decline in poor lighting conditions and ambiguous visual cues [25]. For example, male eyes with thin eyebrows are sometimes classified as female eyes, while female eyes with strong eyelid contours are more often detected as male eyes. To support the analysis of classification errors, several examples of misclassified images are shown in Figure 9.



Figure 9. Example of missed classification image (Pavelbiz/Eyes-rtte/Kaggle).

Classification errors often occur in images with poor lighting conditions or non-standard eye orientation. This is consistent with the findings of Pradana & Wijiyanto, who also noted errors in images of women that were incorrectly predicted as men. Another factor contributing to classification errors is the model's limitation in extracting subtle features around the eyes, especially when images are noisy or low-resolution [1][2]. In addition, data imbalance also affects classification results, where the number of male images (6,323) is greater than the number of female images (5,202). This condition has the potential to cause bias in the model, so that the model tends to be better at recognizing the male class than the female class. As a result, classification errors occur more frequently in images of women because they have less data representation. These findings show that the quality and balance of the input data play an important role in improving the accuracy of eye image-based biometric classification systems [17][19].

4. CONCLUSION

The Convolutional Neural Network (CNN) method was successfully applied to classify gender through images of human eyes. Two CNN architectures were used: MobileNetV1 and DenseNet-121, both implemented using a transfer learning approach tailored to binary classification needs. The CNN model was able to produce classifications with a high level of accuracy. The MobileNetV1 architecture with rotation and translation augmentation provided the best performance, with an accuracy of 92.49% using rotation augmentation, surpassing the results of the DenseNet-121 architecture, which achieved a maximum accuracy of 86.84% with a combination of rotation and translation augmentation.

LITERATURE

- [1] I. P. E. Sutariawan, G. R. Dantes, and K. Y. E. Aryanto, "Segmentasi mata katarak pada citra medis menggunakan metode operasi morfologi," *Jurnal Ilmu Komputer Indonesia*, vol. 3, no. 1, pp. 23–31, 2018.
- [2] A. Triwijoyo, "Pemrosesan citra untuk pengenalan pola biometrik," *Jurnal Ilmiah Informatika*, vol. 4, no. 2, pp. 77–85, 2019.

- [3] S. F. Handono, F. T. Anggraeny, and B. Rahmat, "Implementasi convolutional neural networks (CNN) untuk deteksi retinopati diabetik," *Jurnal Informatika dan Sistem Informasi (JIFoSI)*, vol. 1, no. 1, pp. 669–678, 2020.
- [4] M. R. Alwanda, R. P. K. Ramadhan, and D. Alamsyah, "Implementasi metode convolutional neural network menggunakan arsitektur LeNet-5 untuk pengenalan doodle," *Jurnal Algoritme*, vol. 1, no. 1, pp. 45–56, 2020.
- [5] A. Kuehlkamp and K. W. Bowyer, "Predicting gender from iris texture may be harder than it seems," in *Proc. IEEE Winter Conf. Applications of Computer Vision (WACV)*, Waikoloa Village, HI, USA, pp. 904–912, Jan. 2019.
- [6] D. Bobeldyk and A. Ross, "Analyzing covariate influence on gender and race prediction from near-infrared ocular images," *IEEE Access*, vol. 7, no. 1, pp. 7905–7919, 2018.
- [7] A. Rehman, M. N. Sajid, M. R. U. R. Talib, M. S. A. Malik, and A. Ullah, "Authentication through gender classification from iris images using support vector machine," *Expert Systems*, vol. 38, no. 5, pp. e12710, 2021.
- [8] B. M. S. Hasan and R. J. Mstafa, "Exploring the feasibility of deep learning techniques for accurate gender classification from eye images," *Science Journal of University of Zakho*, vol. 10, no. 4, pp. 222–234, 2022.
- [9] R. Pradana and D. Wijiyanto, "Identifikasi jenis kelamin otomatis berdasarkan mata menggunakan CNN dan Haar cascade classifier," *Jurnal Informatika dan Sistem Cerdas*, vol. 7, no. 1, pp. 44–51, 2024.
- [10] R. D. Wicaksono and G. F. Shidiq, "Comparison of KNN and CNN algorithms for gender classification based on eye images," *Scientific Journal of Informatics*, vol. 11, no. 4, pp. 915–924, 2024.
- [11] B. Patil and M. Hangarge, "Pre-trained convolutional neural networks for gender classification," *International Journal of Computer Vision and Signal Processing*, vol. 13, no. 1, pp. 33–40, 2023.
- [12] D. Kurniawan and A. Irsyad, "Perbandingan metode K-nearest neighbor dan naive Bayes untuk klasifikasi gender berdasarkan mata," *Jurnal Sains Komputer dan Informatika*, vol. 6, no. 2, pp. 155–162, 2022.
- [13] J. Tapia, C. Aravena, and C. Busch, "Deep iris: Deep learning for gender classification through iris patterns," *Sensors*, vol. 19, no. 19, pp. 4061, Sep. 2019.
- [14] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition Workshops (CVPRW)*, Honolulu, HI, USA, pp. 684–692, 2017.
- [15] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 4700–4708, 2017.
- [16] A. Gani, "Transformasi affine pada bidang," *Jurnal Matematika Unisba*, vol. 1, no. 1, pp. 1–8, 2018.
- [17] B. Purnomo, R. Santosa, and D. Widodo, "Augmentasi citra dengan transformasi affine untuk peningkatan robustness CNN," *Jurnal Teknologi dan Sistem Informasi*, vol. 10, no. 2, pp. 101–109, 2022.
- [18] Pavelbiz, "Eyes-RTTE: Female and male eyes dataset," Kaggle, Dataset, 2021. [Online]. Available: <https://www.kaggle.com>. [Accessed: Dec. 2024].
- [19] I. Wulandari, H. Yasin, and T. Widiari, "Klasifikasi citra digital bumbu dan rempah dengan algoritma convolutional neural network (CNN)," *Jurnal Gaussian*, vol. 9, no. 3, pp. 273–282, 2020.
- [20] A. Rahman, A. P. Windarto, and A. P. U. Siahaan, "Evaluasi metode klasifikasi menggunakan confusion matrix," *Jurnal Sistem Informasi*, vol. 13, no. 2, pp. 45–52, 2017.

- [21] M. Resa, A. Yudianto, and H. Al Fatta, “Analisis pengaruh tingkat akurasi klasifikasi citra wayang dengan algoritma convolutional neural network,” *Jurnal Teknologi Informasi*, vol. 4, no. 2, pp. 182–190, 2020.
- [22] T. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 60, pp. 1–48, 2019.
- [23] S. Lakra, P. Singh, and A. Kumar, “Gender classification using convolutional neural networks for biometric applications,” *Procedia Computer Science*, vol. 167, pp. 2530–2537, 2020.
- [24] D. Darmatasia and A. M. Syafar, “Implementasi convolutional neural network untuk klasifikasi tanaman rimpang secara virtual,” *Jurnal INSTEK*, vol. 8, no. 1, pp. 122–131, 2023.
- [25] P. Kumari and K. R. Seeja, “A novel periocular biometrics solution for authentication during COVID-19 pandemic situation,” *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 11, pp. 8427–8440, 2021.