JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 413 – 424 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

THE EFFECT OF FACIAL ACCESSORY AUGMENTATION ON THE ACCURACY OF DEEP LEARNING-BASED FACIAL RECOGNITION SYSTEMS

Ahmad Nur Hidayat[1], Nanik Suciati [1*], Ahmad Saikhu [1]
[1]Department of Informatics Engineering, institut Teknologi Sepuluh Nopember
email: 6025222012@student.its.ac.id, nanik@if.its.ac.id, saikhu@if.its.ac.id

Abstract: Face recognition based on deep learning has become an important technology in many areas. However, these systems often face challenges in real-world conditions, such as when the face is partially covered by accessories such as masks or glasses. This study aims to evaluate the effect of data augmentation by adding facial accessories (masks, glasses, and a combination of both) and geometric augmentation on the accuracy of face recognition systems. There are three types of datasets used in this method: the original dataset (category 1), the dataset with facial accessories augmentation (category 2), and the dataset with geometric augmentation (category 3). Data augmentation was performed on the training dataset to increase diversity, followed by the face detection process using SCRFD and feature extraction with ArcFace. The model was then trained using Multi-Layer Perceptron (MLP). Based on the results, adding face accessories (category 2) made the model a lot more accurate, hitting 99% accuracy. In category 3, adding geometric features improved accuracy to 91%. Other evaluation metrics, such as precision, recall, and F1-score, also showed improvement after augmentation. This study concludes that facial accessories augmentation is more effective in improving the accuracy and robustness of face recognition models compared to geometric augmentation.

Keywords: augmentation; deep learning; face recognition; glasses.

Abstrak: Pengenalan wajah berbasis deep learning telah menjadi salah satu teknologi penting dalam berbagai aplikasi. Namun, sistem ini sering kali menghadapi tantangan dalam kondisi dunia nyata, seperti saat wajah tertutup sebagian oleh aksesori seperti masker atau kacamata. Penelitian ini bertujuan untuk mengevaluasi pengaruh augmentasi data dengan menambahkan aksesori wajah (masker, kacamata, dan kombinasi keduanya) serta augmentasi geometris terhadap akurasi sistem pengenalan wajah. Metode yang digunakan melibatkan tiga kategori dataset: dataset asli tanpa augmentasi (kategori 1), dataset dengan augmentasi aksesoris wajah (kategori 2), dan dataset dengan augmentasi geometris (kategori 3). Augmentasi data dilakukan pada dataset pelatihan untuk meningkatkan keberagaman, diikuti dengan proses deteksi wajah menggunakan SCRFD dan ekstraksi fitur dengan ArcFace. Model kemudian dilatih menggunakan Multi-Layer Perceptron (MLP). Hasil penelitian menunjukkan bahwa augmentasi aksesoris wajah (kategori 2) memberikan peningkatan signifikan pada akurasi model, mencapai 99%, sedangkan kategori 3 dengan augmentasi geometris mencapai akurasi 91%. Metrik evaluasi lainnya, seperti precision, recall, dan F1-score, juga menunjukkan peningkatan setelah augmentasi. Penelitian ini menyimpulkan bahwa augmentasi aksesoris wajah lebih efektif dalam meningkatkan akurasi dan ketahanan model pengenalan wajah dibandingkan dengan augmentasi geometris.

Kata kunci: augmentasi; deep learning; kacamata; pengenalan wajah.

DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

INTRODUCTION

Classroom management and student attendance monitoring are critical aspects in achieving effective educational goals. Studies show a positive correlation between attendance rates and academic achievement, where accurate attendance monitoring ensures active student participation in learning. However, many schools still rely on conventional systems, such as manual recording in attendance books, which are prone to human error, data manipulation, and time inefficiency. Therefore, this research aims to propose a technology-based solution to improve the accuracy and efficiency of the attendance recording system in the school environment.

Advancements in artificial intelligence (AI) have fostered the emergence of several breakthroughs, particularly in deep learning-based facial recognition [1]. The implementation of this technology in attendance recording systems has the potential to improve efficiency and accuracy, by enabling real-time and automated identification processes. Using deep neural networks architecture, deep learning has proven to be effective in various applications, such as face recognition, speech recognition, and natural language processing [2], [3].

Recent literature supports the effectiveness of deep learning in face detection and recognition, especially when addressing masked or partially occluded faces. One of the latest innovations, a method that combines one-stage and two-stage detectors with ResNet50 to detect faces without masks, achieved 98.2% accuracy [4]. Furthermore, a face mask detection model that combines deep learning and Internet of Things (IoT) technology, using ResNet50 and MobileNetV2 achieved high accuracy [5]. Furthermore, the development of a CNN-based masked face recognition system achieved 90.4% accuracy with the VGGFace2 and MaskedFace-Net datasets [6]. In addition, the utilization of deep CNN and staged image processing for automatic face mask detection also achieved 97.5% accuracy [7].

Face recognition for attendance systems has seen rapid adoption, though real-world accuracy remains challenging. While Yang & Han (2020) achieved 82% accuracy using SVM-based video processing [8], Ahmed et al. (2022) reached 99.75% accuracy with a CNN approach [9], demonstrating the technology's potential despite implementation challenges.

In addition, research related to dataset augmentation has become an important topic to improve the accuracy of face recognition systems. While Wang et al. (2020) examined general augmentation techniques [10], Ayad et al. (2023) developed the CASIA-mask dataset (418,978 images) by synthetically adding masks to unmasked faces [11]. Prusty et al. (2021) further advanced this field by achieving 99.8% mask detection precision using YOLOv3 and image filtering [12], collectively demonstrating that synthetic augmentation significantly improves model performance in real-world scenarios.

Although facial recognition technology continues to evolve, its accuracy still faces challenges under certain conditions, especially when subjects use facial accessories such as masks and glasses. Data augmentation is expected to be an effective solution to address this issue. Augmentation by adding elements such as masks and glasses to the facial dataset can increase the diversity of the training data,

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 413 – 424 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

allowing the model to be more accurate in recognizing faces under various conditions. This augmentation is expected to improve the system's ability to recognize faces even with variations in attributes such as masks and glasses that cover important parts of the face.

In addition to facial accessory augmentation, this research also explores the use of geometric augmentation as a comparison category. The purpose of adding this category is to test how geometric augmentation can improve the performance of facial recognition models and to compare its effectiveness with facial accessory augmentation. By adding the geometric augmentation category, it is hoped that deeper insights can be gained regarding the types of augmentation that are most effective in improving facial recognition accuracy, both under normal conditions and with obstructions caused by accessories.

Recent advances in deep learning have significantly improved face recognition systems. SCRFD (Sample and Computation Redistribution for Face Detection) has emerged as an efficient solution for cost-effective face detection, while ArcFace (Additive Angular Margin Loss) enhances recognition accuracy through discriminative feature embedding, particularly effective for scenarios with multiple or moving subjects [13].

Ozdemir & Hanbay (2022) developed a deep learning-based face tracking system that uses SCRFD for face detection, ArcFace for face recognition, and DeepSORT algorithm for stable face tracking. The results showed that the combination of SCRFD, ArcFace, and DeepSORT was effective for real-time face detection and tracking [14]. Furthermore, Arun Kumar et al. (2021) proposed a face recognition system with ArcFace and ensemble learning for individual identification despite wearing a mask showing that ArcFace is effective in model training, and the use of ensemble learning with soft voting increases face recognition accuracy to 93.65% [15].

The integration of SCRFD and ArcFace methods is expected to address the challenges in face detection hindered by the use of accessories such as masks and glasses. Therefore, this research aims to evaluate the impact of augmenting facial accessories, such as masks, glasses, and their combinations, on the accuracy of a facial recognition-based attendance system using deep learning. By adding these elements to the dataset, it is expected to improve the model's ability to recognize faces more accurately in various conditions, including when individuals are wearing accessories.

METHOD

The research flow utilized in this study is illustrated in Image 1. The study commenced with data collection at a selected school in Samarinda, East Kalimantan, which was chosen due to its continued use of a conventional, manual attendance system. Observations indicated that several students experienced changes in appearance over time, such as transitioning from not wearing glasses to wearing them, and the occasional use of face masks, particularly during illness. These variations in facial features reflect real-world conditions and serve as a valuable basis for evaluating the robustness of facial recognition systems under dynamic circumstances.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi)
Vol XI No 3, Juni 2025, hlm. 413 – 424
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi
ISSN 2407-1811 (Print)
ISSN 2550-0201 (Online)

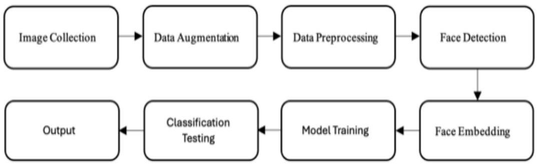

Image 1. Research Flow

During the data collection process, facial photographs were captured directly using a camera with a resolution of 500×500 pixels. A total of 20 students participated as research subjects, with each individual photographed 12 times to obtain facial images exhibiting various expressions. This process yielded 240 facial images, which were compiled into what is referred to as the Category 1 dataset, consisting of original, unaltered facial images. Examples of the collected images are presented in Image 2.

Image 2. Example of Original Face Data

Subsequently, each facial image in the Category 1 dataset underwent an augmentation process involving the addition of facial accessories, including glasses, masks, and a combination of both. The accessories used for augmentation are illustrated in Images 3.

Image 3. Glasses and Mask Accessories

The facial image augmentation stage begins with the input of original facial images, to which virtual facial accessories are added. Each image undergoes a face detection process utilizing the DLIB library with the Histogram of Oriented Gradients (HOG) method. The augmentation technique applied for embedding facial accessories is illustrated in Image 4.

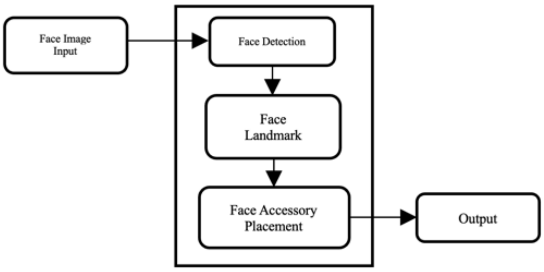

Image 4. Flow of Face Image Augmentation

After the face is detected as shown in Image 5, the next step is to predict the facial landmark points. This model can recognize important points on the face, such as the eyes, nose, mouth, and facial contours, with a total of 68 facial points as shown in Image 6.

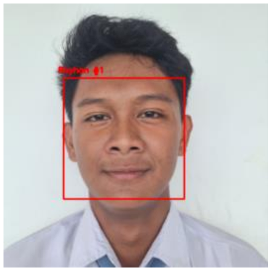

Image 5. Detected Face Image

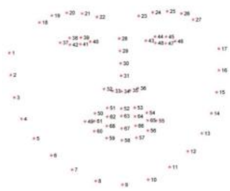

Image 6. Facial Coordinate Points

After successfully obtaining the predicted facial landmark points with the appropriate coordinates as shown in Image 7, the height and width of the facial accessory to be embedded are measured, and then the facial accessory is placed at the coordinate point that best

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 413 – 424 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

matches the original position of the facial accessory. The output is an original face image with added facial accessory augmentation, resulting in new face image data.

Image 7. Face Image Coordinate Points

For each original facial image, 15 augmented images were generated by applying variations of facial accessories. The accessory variations consisted of three types of eyeglasses—thin round glasses, thick round glasses, and rectangular glasses—and three types of face masks, including surgical masks, duckbill masks, and KF94 masks.

In addition, nine combinations were created by pairing each type of eyeglasses with each type of mask, resulting in a comprehensive set of accessory combinations. This augmentation process aimed to simulate real-world variability in facial appearance and enhance the robustness of the facial recognition system. The addition of accessories was performed on 240 facial datasets, resulting in a total of 3600 new facial images.

Thus, the total original dataset, when combined with this augmented data, amounts to 3840 facial images. This dataset was then named category 2 dataset, which includes original facial images plus augmented facial images with glasses, masks, and combinations of both. Examples of the dataset that have undergone the augmentation process can be seen in Image 8.

Image 8. Face Accessory Augmentation Dataset

As an objective comparison to the facial accessory augmentation dataset, this research develops a category 3 dataset that adopts conventional geometric augmentation. Each facial image in the category 1 dataset underwent an augmentation process through six random transformation techniques, applying one type of augmentation from the six techniques: 20% zoom, 20° rotation, 20% shift, horizontal flip, noising, and blurring. From 1 facial image, 15 new facial images are generated.

The addition of conventional geometric augmentations is applied to 240 facial datasets, resulting in a total of 3600 new facial images. Therefore, the total original dataset plus the augmented data in this category 3 dataset amounts to 3840 facial images. Examples of the dataset that have undergone the augmentation process can be seen in Image 9.

Image 9. Example of Conventional Geometric Data Augmentation Dataset

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 413 – 424 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

The dataset to be used has 3 categories: category 1, which is the original dataset consisting of 240 facial images; category 2, which is the original dataset augmented with facial accessory augmentation, totaling 3840 facial images; and category 3, which is the original dataset augmented with conventional geometric augmentation, also totaling 3840 facial images.

The dataset underwent data preprocessing by performing. The dataset was resized to 150x150 pixels. The purpose of this resizing is to ensure that all images have consistent dimensions. This allows the model to focus on important structural features and ensures more consistent data, thereby facilitating and speeding up the training and prediction processes.

After the data pre-processing, the next step is to split the data into training and validation sets with a ratio of 75:25. For the test data, another dataset consisting of 20 classes with a total of 4 photos from each class is provided, resulting in a total of 80 facial image photos for the entire test data.

These include original facial images without facial accessories, facial images with original glasses, facial images with original masks, and facial images with both original glasses and original masks. Example of the test data can be seen in Image 10. The three categories of the dataset were tested for face detection and recognition with separate models, the flow of face detection and recognition can be seen in Image 10.

Image 10. Example Test Data

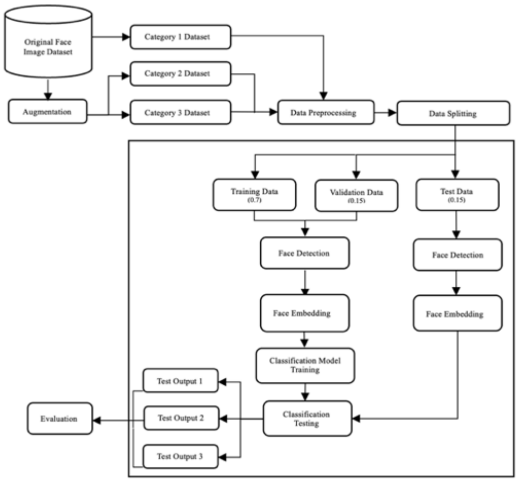

Image 11. Face Detection and Recognition Flow

In the first category trial, the original dataset underwent data preprocessing, followed by data splitting. However, for the second and third category trials, the original dataset underwent data augmentation first. Category 2 involved adding facial accessories such as glasses, masks, and combinations of both, resulting in a total addition of 3840 new facial images.

Meanwhile, Category 3 involved adding conventional geometric augmentations (20% zoom, 20° rotation, 20% shift, horizontal flipping, noising, and blurring), also resulting in a total addition of 3840 new facial images. The dataset was then subjected to data preprocessing and data splitting. The dataset is divided into training and validation data with a ratio of 75:25.

The dataset was then subjected to face detection using SCRFD to determine the positions of faces in the dataset, resulting in bounding boxes for each detected face, which were then cropped based on the bounding boxes. After all the facial image data has been successfully cropped, feature extraction and face embedding formation are performed using ArcFace, resulting in a feature vector. The next step is to train the classification model using MLP with the feature vector. The trained models consist of 3 models: model 1 obtained from training using the category 1 dataset, model 2 obtained from training using the category 2 dataset, and model 3 obtained from training using the category 3 dataset. These three models were tested for classification using test data to determine the accuracy in face recognition and mean average precision (MAP) in face detection for evaluation.

After the dataset is collected according to the requirements, the process of testing face detection and recognition is carried out using the SCRFD and ArcFace methods. The testing was conducted on the original data and then compared with the data that had been augmented with a combination of augmentations for further analysis using mean average precision (MAP).

At the analysis stage, a comparison of the results of face detection and recognition tests using the SCRFD and ArcFace methods was conducted on original data (category 1), data that has been augmented with glasses, masks, and combinations of both (category 2), and data that has been augmented with conventional geometric augmentations such as 20% zoom, 20° rotation, 20% shift, horizontal flipping, noising, and blurring (category 3).

The evaluation is divided into three parts: model evaluation with category 1 facial image data, model evaluation with category 2 facial image data, and model evaluation with category 3 facial image data. The evaluation of object detection conducted uses mean average precision (MAP). The metrics used to analyze the system are accuracy, precision, recall, and F1-score.

RESULT AND DISCUSSION

Evaluation of Category 1 Dataset (Without Augmentation)

The evaluation metric in Image 13 shows the model's performance on the category 1 dataset, which is the original dataset without augmentation, with a test accuracy of 0.89. Metrics such as precision, recall, and F1-score show variation between classes. The highest precision reached 1.00 in several classes, but classes 9 and 10 had lower precision (0.67), indicating the model's difficulty in distinguishing faces under certain conditions.

High recall in most classes, including those with low precision, indicates that the model is effective in detecting relevant faces, with a recall of 1.00 in many classes. The F1-score approaching 1.00 in most classes reflects overall good performance, although there is still room for improvement in classes with lower precision and recall.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	4
1	1.00	0.50	0.67	4
2	1.00	0.50	0.67	4
3	1.00	1.00	1.00	4
4	1.00	0.75	0.86	4
5	0.80	1.00	0.89	4
6	1.00	0.75	0.86	4
7	1.00	1.00	1.00	4
8	0.80	1.00	0.89	4
9	0.67	1.00	0.80	4
10	0.67	1.00	0.80	4
11	1.00	1.00	1.00	4
12	1.00	1.00	1.00	4
13	1.00	1.00	1.00	4
14	0.75	0.75	0.75	4
15	1.00	1.00	1.00	4
16	0.80	1.00	0.89	4
17	0.75	0.75	0.75	4
18	1.00	1.00	1.00	4
19	1.00	0.75	0.86	4
accuracy			0.89	80
macro avg	0.91	0.89	0.88	80
weighted avg	0.91	0.89	0.88	80

Image 12. Evaluation Metrics of Category 1

Overall, although the model's accuracy is good, improvements can be achieved by addressing the imbalance between precision and recall through dataset adjustments or the application of augmentation techniques to enhance the

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 413 – 424 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

diversity of the training data, particularly for classes that are more difficult to recognize.

Evaluation of Category 2 Dataset (Accessory Augmentation)

Image 14 shows the evaluation metrics for category 2 for the model trained with the augmented facial accessory dataset (mask, glasses, and their combination). Overall, the model showed very good results with an accuracy of 0.99 on the test data. This shows that augmenting facial accessories plays a significant role in enhancing the model's ability to recognize faces even when wearing masks and glasses.

Other metrics, such as precision, recall, and F1-score, show almost perfect results, with most classes having precision, recall, and F1-score values reaching 1.00 each. This shows that the model successfully classified most faces very well despite variations due to the use of accessories.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	4
1	1.00	1.00	1.00	4
2	1.00	1.00	1.00	4
3	1.00	1.00	1.00	4
4	1.00	1.00	1.00	4
5	1.00	1.00	1.00	4
6	1.00	1.00	1.00	4
7	0.80	1.00	0.89	4
8	1.00	1.00	1.00	4
9	1.00	1.00	1.00	4
10	1.00	1.00	1.00	4
11	1.00	1.00	1.00	4
12	1.00	1.00	1.00	4
13	1.00	1.00	1.00	4
14	1.00	1.00	1.00	4
15	1.00	1.00	1.00	4
16	1.00	1.00	1.00	4
17	1.00	0.75	0.86	4
18	1.00	1.00	1.00	4
19	1.00	1.00	1.00	4
accuracy			0.99	80
macro avg	0.99	0.99	0.99	80
weighted avg	0.99	0.99	0.99	80

Image 13. Category 2 Evaluation Metrics

However, there is one class (class 7) that has a lower precision (0.80), although its recall and F1-score are still quite high (1.00 and 0.89, respectively). This may indicate the model's difficulty in recognizing certain faces under more complex facial accessory conditions. Overall, although there are some minor variations, the significant improvement in precision, recall, and F1-score across most classes indicates that augmenting facial accessories has helped the model recognize faces better in more diverse real-world conditions, which is important for broader facial recognition applications.

Evaluation of Category 3 Dataset (Geometric Augmentation)

Image 15 shows the evaluation. metrics for category 3, using geometric augmentations such as zoom, rotation, shifting, flipping, noise, and blurring. Overall, the model shows a fairly good accuracy of 0.91 on the test data, which is slightly lower than category 2. Other metrics, such as precision, recall, and F1-score, show quite good variation between classes, with precision and recall values mostly approaching 1.00.

Some classes, such as class 2 and class 5, showed a decrease in precision and recall (0.67 and 0.50), which may indicate the model's difficulty in recognizing individuals under complex augmentation conditions. Nevertheless, the overall high F1-score indicates that the model continues to perform well in most classes. The increase in recall indicates that geometric augmentation helps the model recognize more positive instances.

	precision	recall	f1-score	support
0	1.00	1.00	1.00	4
1	0.80	1.00	0.89	4
2	0.67	1.00	0.80	4
3	1.00	0.75	0.86	4
4	1.00	1.00	1.00	4
5	1.00	0.50	0.67	4
6	1.00	0.75	0.86	4
7	0.80	1.00	0.89	4
8	1.00	1.00	1.00	4
9	1.00	1.00	1.00	4
10	1.00	0.75	0.86	4
11	1.00	1.00	1.00	4
12	0.80	1.00	0.89	4
13	0.80	1.00	0.89	4
14	1.00	0.75	0.86	4
15	1.00	1.00	1.00	4
16	0.80	1.00	0.89	4
17	1.00	0.75	0.86	4
18	1.00	1.00	1.00	4
19	1.00	1.00	1.00	4
accuracy			0.91	80
macro avg	0.93	0.91	0.91	80
weighted avg	0.93	0.91	0.91	80

Image 14. Category 3 Evaluation Metrics

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol XI No 3, Juni 2025, hlm. 413 – 424 ISSN 2550-0201 (Online)
DOI: https://doi.org/10.33330/jurteksi.v11i3.3846
Available online at https://jurnal.stmikroyal.ac.id/index.php/jurteksi

Overall, although there were some decreases in precision in several classes, geometric augmentation proved to enhance the model's ability to generalize and handle greater variations in facial conditions, which is important for facial recognition applications in real-world situations.

Comparison of Model Performance Before and After Augmentation

Table 1 shows a comparison of the model's performance before and after augmentation, both involving facial accessory augmentation (mask, glasses, and their combination) and using geometric augmentations such as zoom, rotation, shift, flip, noise, and blurring.

It can be seen that there is a 10% increase in accuracy, indicating that the model is more effective in classifying faces after being augmented with masks and glasses or a combination of both. Additionally, there is a slight accuracy improvement of 2% in geometric augmentations such as zoom, rotation, shifting, flipping, noise, and blurring.

This is in line with the research conducted by Pham et al. (2024) which combines deep learning-based methods with data augmentation to improve performance in scenarios where people wear masks and glasses [16]. This approach achieves high accuracy even with limited training data. Similarly, the research conducted by Alamsyah et al. in 2022 on facial data augmentation by creating variations in facial poses through angular head movements, such as yaw, pitch, and roll [17].

The research results show the best outcomes for all models. The model's accuracy increased from 0.8108 to 0.8719 after applying facial data augmentation techniques. This increase in accuracy is due to data augmentation enhancing the model's ability to apply the knowledge learned to new data that it has never encountered before. As a result, this leads to better performance on the test dataset [18], [19].

Table 1. Comparison of Model Performance Before and After Augmentation

Metric	Without Augmentation	With Augmentation	
	Kategori 1	Kategori 2	Kategori 3
Accuracy	0.89	0.99	0.91
Precision	0.88	0.99	0.92
Recall	0.85	0.98	0.91
F1-Scoce	0.86	0.98	0.91

Additionally, the augmented dataset makes the model more resilient to real-world variations and disturbances, which is crucial in facial recognition, especially concerning lighting, expressions, and obstructions that can affect performance [20], [21].

Data augmentation techniques, such as the addition of masks and glasses, enrich the diversity in the training dataset. This diversity allows the model to learn how to recognize faces under various conditions and obstructions, which in turn strengthens the model's resilience and generalization capability [16], [22]. By mimicking real-world conditions, where faces are often partially obstructed, the model can better handle these variations during the inference process [16], [22]. Data augmentation with masks and glasses also forces the model to focus on visible and crucial facial features that are not covered. This can result in better feature extraction and improved facial recognition performance [23], [24].

In addition, data augmentation techniques—particularly geometric augmentation methods—can significantly improve accuracy compared to training without augmentation. Geometric augmentation introduces variations within the dataset, enabling the model to generalize more effectively. By incorporating different perspectives and distortions, the model becomes better equipped to recognize faces under varying conditions, thereby enhancing both robustness and accuracy [17], [25].

Moreover, geometric augmentation improves feature extraction. Techniques such as rotation, zooming, and shifting enable the model to learn more diverse features from the images. This increased variety in features enhances the model's ability to distinguish between different faces, which contributes to improved accuracy [17], [26].

Overall, it can be concluded that facial augmentation using accessories plays a more significant role in improving the accuracy and robustness of facial recognition models compared to geometric augmentation. This is because accessory augmentation directly addresses issues of occlusion and enhances the model's ability to generalize across various real-world conditions. By training the model with images containing facial accessories, the model becomes more resilient to scenarios in which faces are partially obscured [22], [23]. In contrast, geometric transformations primarily address pose variations but do not effectively mitigate the issue of occlusion, which is frequently encountered in real-world environments [22].

CONCLUSION

This research evaluates the impact of data augmentation on a deep learning-based facial recognition system through two approaches: the addition of facial accessories and geometric transformations. The results show a significant improvement - accuracy increased from 89% (without augmentation) to 99% (accessories) and 91% (geometric). Evaluation metrics (precision, recall, F1-score) also showed the highest improvement with accessory augmentation (0.99) compared to geometric (0.91-0.92). The findings prove that data augmentation, particularly facial accessories, significantly improves the accuracy and robustness of the facial recognition model compared to geometric transformations.

Although this research makes a significant contribution in improving the accuracy of face recognition systems through data augmentation with accessories and geometric changes, there are some areas that can still be improved. Future research can explore other augmentation techniques, such as color-based augmentation to create more variety. In addition, testing on more diverse datasets, including more complex real-world conditions, can improve the model's ability to generalize. Thus, these suggestions can serve as a foundation for further research that is more comprehensive and applicable.

BIBLIOGRAPHY

[1] X. Chen, "The Advance of Deep Learning and Attention Mechanism," in *Proceedings* - 2022 International Conference on Electronics and Devices, Computational Science*, ICEDCS 2022*, 2022. doi: 10.1109/ICEDCS57360.2022.00078.

[2] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Comput Sci*, vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.

[3] J. Chai, H. Zeng, A. Li, and E. W. T. Ngai, "Deep learning in computer vision: A critical review of emerging techniques and application scenarios," *Machine Learning with Applications*, vol. 6, p. 100134, Dec. 2021, doi: 10.1016/j.mlwa.2021.100134.

[4] S. Sethi, M. Kathuria, and T. Kaushik, "Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread," *J Biomed Inform*, vol. 120, p. 103848, Aug. 2021, doi: 10.1016/j.jbi.2021.103848.

[5] R. A. S. Naseri, A. Kurnaz, and H. M. Farhan, "Optimized face detector-based intelligent face mask detection model in IoT using deep learning approach," *Appl Soft Comput*, vol. 134, p. 109933, Feb. 2023, doi: 10.1016/j.asoc.2022.109933.

[6] B. Kocacinar, B. Tas, F. P. Akbulut, C. Catal, and D. Mishra, "A Real-Time CNN-Based Lightweight Mobile Masked Face Recognition System," *IEEE Access*, vol. 10, pp. 63496–63507, 2022, doi: 10.1109/ACCESS.2022.3182055.

[7] M. Umer *et al.*, "Face mask detection using deep convolutional neural network and multi-stage image processing," *Image Vis Comput*, vol. 133, p. 104657, May 2023, doi: 10.1016/j.imavis.2023.104657.

[8] H. Yang and X. Han, "Retracted: Face Recognition Attendance System Based on Real-Time Video Processing," *IEEE Access*, vol. 8, pp. 159143–159150, 2020, doi: 10.1109/ACCESS.2020.3007205.

[9] M. A. Ahmed, M. D. Salman, R. A. Alsharida, Z. T. Al-Qaysi, and M. M. Hammood, "An Intelligent Attendance System Based On Convolutional Neural Networks For Real-Time Student Face Identifications," *Journal of Engineering Science and Technology*, vol. 17, no. 5, 2022.

[10] X. Wang, K. Wang, and S. Lian, "A survey on face data augmentation for the training of deep neural networks," *Neural Comput Appl*, vol. 32, no. 19, pp. 15503–15531, Oct. 2020, doi: 10.1007/s00521-020-04748-3.

[11] Waleed Ayad, Siraj Qays, and Ali Al-Naji, "Generating and Improving a Dataset of Masked Faces Using Data Augmentation," *Journal of Techniques*, vol. 5, no. 2, pp. 46–51, Jun. 2023, doi: 10.51173/jt.v5i2.1140.

[12] M. R. Prusty, V. Tripathi, and A. Dubey, "A novel data augmentation approach for mask detection using deep transfer learning," *Intell Based Med*, vol. 5, p. 100037, 2021, doi: 10.1016/j.ibmed.2021.100037.

[13] M. F. Ozdemir and D. Hanbay, "Deep Learning Based Real-Time Face Tracking System in Multi-Camera," *SSRN Electronic Journal*, 2022, doi: 10.2139/ssrn.4098733.

[14] M. F. Ozdemir and D. Hanbay, "Deep Learning Based Real-Time Face Tracking System in Multi-Camera," *SSRN Electronic Journal*, 2022, doi: 10.2139/ssrn.4098733.

[15] R. Arun Kumar, V. A. Solayappan, S. S. T, and R. P. K, "Masked Deep Face Recognition using ArcFace and Ensemble Learning," in *2021* IEEE 2nd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET), IEEE, Dec.

2021, pp. 1–6. doi: 10.1109/TEMSMET53515.2021.9768777.

[16] P.-S. Pham, M.-K. Nguyen, M.-H. Le, M.-T. Tran, and T.-N. Le, “Masked Face Recognition on Limited Training Data,” in *2024* International Conference on Multimedia Analysis and Pattern Recognition (MAPR), IEEE, Aug. 2024, pp. 1–6. doi: 10.1109/MAPR63514.2024.10661046.

[17] T. M. S. N. Alamsyah, T. F. Abidin, R. Ferdhiana, M. Dirhamsyah, and M. Chaidir, “Analysis of Face Data Augmentation in Various Poses for Face Recognition Model,” in *2022* 7th International Conference on Informatics and Computing, ICIC 2022, 2022. doi: 10.1109/ICIC56845.2022.10006997.

[18] P. Rahmatullah, T. F. Abidin, A. Misbullah, and Nazaruddin, “Effectiveness of Data Augmentation in Multi-class Face Recognition,” in *2021* 5th International Conference on Informatics and Computational *Sciences (ICICoS)*, IEEE, Nov. 2021, pp. 64–68. doi: 10.1109/ICICoS53627.2021.9651780.

[19] A. Ma, N. Dvornik, R. Zhang, L. Pishdad, K. G. Derpanis, and A. Fazly, “SAGE: Saliency-Guided Mixup with Optimal Rearrangements,” in *BMVC 2022 - 33rd* British Machine Vision Conference Proceedings, 2022.

[20] S. Zeng, B. Zhang, Y. Zhang, and J. Gou, “Dual sparse learning via data augmentation for robust facial image classification,” International Journal of Machine Learning and Cybernetics, vol. 11, no. 8, 2020, doi: 10.1007/s13042-020-01067-w.

[21] E. Ghorbel, G. Maddouri, and F. Ghorbel, “Face Blending Data Augmentation for Enhancing Deep Classification,” in International Conference on Pattern Recognition Applications and Methods, 2024. doi: 10.5220/0012357900003654.

[22] R. Shah, M. Purohit, S. Gadhe, K. Mehrotra, and S. Mehta, “FaceReZero Transformer for Unbiased Occlusion Invariant Deep Face Recognition with Train-Time Augmentations,” in *2023* Seventh International Conference on Image Information Processing (ICIIP), IEEE, Nov. 2023, pp. 73–78. doi: 10.1109/ICIIP61524.2023.10537650.

[23] H. A. Sung, S. Kim, and E. C. Lee, “Masked Face Recognition Model with Explainable AI,” 2023, pp. 154–159. doi: 10.1007/978-3-031-27199-1_16.

[24] Z. Ye, H. Zhang, X. Li, and Q. Zhang, “DeMaskGAN: a de-masking generative adversarial network guided by semantic segmentation,” *Vis Comput*, vol. 40, no. 8, pp. 5605–5618, Aug. 2024, doi: 10.1007/s00371-023-03125-0.

[25] M. Anusha and P. Kiruthika, “A Comparative Study on Augmented Analytics Using Deep Learning Techniques,” in Smart Innovation, Systems and Technologies, 2022. doi: 10.1007/978-981-16-3675-2_9.

[26] N. J. Shoumy, L. M. Ang, D. M. M. Rahaman, T. Zia, K. P. Seng, and S. Khatun, “Improving Human Emotion Recognition from Emotive Videos Using Geometric Data Augmentation,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2021. doi: 10.1007/978-3-030-79463-7_13.