INTECHBiZ2025
International Conference on Industrial Technology and Business

Performance Evaluation of Face Recognition Using ResNet Architecture Under Different Lighting Conditions

Marzuarman[1], Zainal Abidin[2], Andri Nofiar Am[3], Azizul[4], Desqi Baidillah[5], Fandy Octafiandri[6], Elya Santi Pane[7]
[1,2,3,4,5,6,7]*Jurusan Teknik Elektro, Politeknik Negeri Bengkalis, Indonesia*)
marzuarman@polbeng.ac.id[1], zainal@polbeng.ac.id[2], andrinofiar90@gmail.com[3]

Abstract

Face recognition has become a prominent application area within the field of computer vision, particularly in security and automatic identification systems. This research investigates the performance of the ResNet50 architecture for face recognition tasks across different illumination settings. The face recognition system was built using the DeepFace framework in Python and executed in real time via a webcam interface. The dataset comprised facial images collected under three distinct illumination settings bright, normal, and dim. Each image underwent face detection, alignment, and normalization procedures to match the model's required input format. The system's performance was assessed using accuracy, precision, recall, F1-score, and mean recognition time. The experimental findings indicate that illumination variations have a notable impact on recognition accuracy. The highest accuracy of 98% was achieved under bright lighting, while performance dropped to 84% under dim conditions. The average recognition time ranged between 140–160 milliseconds, indicating that the system performs efficiently in real time. Overall, the ResNet50 architecture demonstrates strong capability in facial recognition under optimal lighting but requires improvement in preprocessing to handle extreme illumination variations.

Keywords : face recognition, ResNet, CNN, lighting, performance.

1. INTRODUCTION

Recent developments in Artificial Intelligence (AI) and computer vision have significantly accelerated the evolution of face recognition technology. This technology is widely used in security systems, automatic attendance, device authentication, and behavioral analysis (Abate et al., 2007; Zhao et al., 2003). Compared to other biometric methods, face recognition is non-invasive, easy to use, and can be implemented using standard camera devices (Wang & Deng, 2021). However, the system's effectiveness largely relies on the quality of the input images, which can fluctuate based on environmental factors like illumination (Jiang et al., 2021; Tan & Triggs, 2010).

Lighting variation remains one of the main challenges in facial recognition systems. Differences in light intensity, direction, and shadow can alter the pixel distribution of an image, making key facial features harder to identify (Tan & Triggs, 2010). Under poor or excessive lighting, feature extraction becomes inconsistent, reducing system accuracy. Hence, robust feature representation methods are needed to ensure stability under varying illumination (Masi et al., 2018; Wang & Deng, 2021).

The Residual Network (ResNet) is recognized as one of the most powerful deep learning architectures for image recognition tasks, introduced by (He et al., 2016). Using skip connections, ResNet allows deep networks to maintain stable gradient propagation and avoid degradation. Although ResNet has demonstrated high accuracy on standard facial datasets (Deng et al., 2019; Parkhi et al., 2015), further investigation is needed to assess its robustness under challenging lighting conditions.

This study investigates the performance of a ResNet based face recognition system under various illumination conditions. The evaluation focuses on the model's ability to maintain recognition accuracy and stability under different illumination settings. Results are expected to contribute to developing more reliable and adaptive face recognition systems for real-world applications (Alzahra & Mahmood, 2023; Wang & Deng, 2021).

2. REVIEW OF LITERATURE

Face Recognition

Face recognition refers to a biometric approach that identifies or confirms an individual by analyzing unique facial traits (Abate et al., 2007; Zhao et al., 2003). A typical face recognition system operates through three fundamental phases: detection, feature extraction, and classification (Masi et al., 2018). Detection locates the face within an image, extraction converts facial features into numerical representations, and classification compares them with stored references to determine identity (Wang & Deng, 2021).

Although convenient and sensor-free, facial recognition accuracy depends heavily on environmental conditions (Jiang et al., 2021; Tan & Triggs, 2010). Variations in facial expression, viewpoint, and especially lighting can significantly affect recognition reliability (Wang & Deng, 2021). Thus, continuous research focuses on developing methods that adapt to environmental changes (Alzahra & Mahmood, 2023; He et al., 2016).

Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a specialized form of artificial neural network optimized for processing and analyzing image data (Chollet, 2018; Krizhevsky et al., 2012). CNNs employ convolutional layers to autonomously capture spatial characteristics, building hierarchical feature representations from simple edges to complex facial structures (Masi et al., 2018; Wang & Deng, 2021).

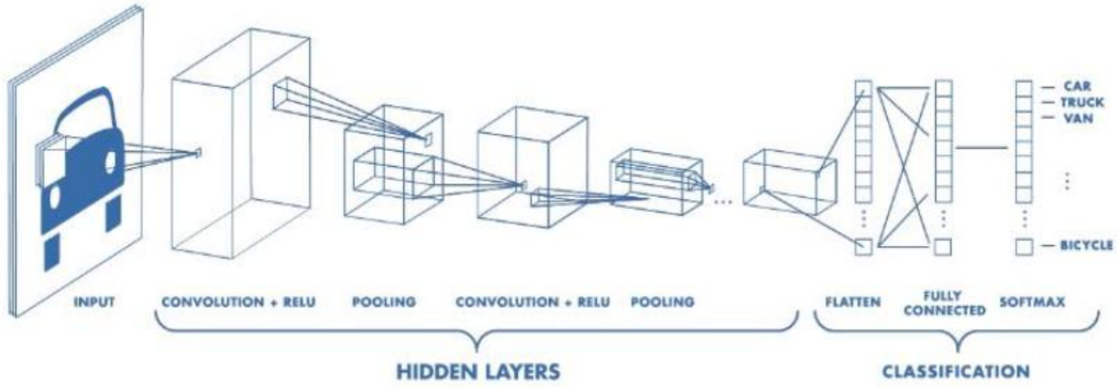

Fig. 1. Architecture of a CNN (Purwono et al., 2022).

The primary advantage of CNNs is their hierarchical learning capability, where initial layers detect simple features such as edges, and deeper layers identify complex abstractions like facial components (Masi et al., 2018). However, deeper CNNs risk experiencing vanishing gradients, which can hinder training efficiency (He et al., 2016; Wang & Deng, 2021).

Residual Network (ResNet)

ResNet, proposed by (He et al., 2016), overcomes the vanishing gradient challenge using shortcut connections, which facilitate smoother gradient propagation between network layers. Variants include ResNet18, ResNet34, ResNet50, ResNet101, and ResNet152, each differing in depth and complexity (He et al., 2016). ResNet has achieved top performance in image recognition challenges like ImageNet and is widely used for extracting robust facial embeddings for identity matching (Deng et al., 2019; Parkhi et al., 2015; Wang & Deng, 2021).

Lighting Influence on Face Recognition

Lighting variation is among the most critical factors affecting face recognition accuracy (Jiang et al., 2021; Tan & Triggs, 2010). Changes in illumination direction and intensity can cause shadows, reflections, and contrast inconsistencies, which hinder feature extraction (Wang & Deng, 2021). Techniques such as histogram equalization or illumination normalization can mitigate these effects, but deep learning models like ResNet offer greater adaptability due to their high-level feature learning capabilities (Alzahra & Mahmood, 2023; He et al., 2016).

3. METHOD

Research Design

This research was conducted to assess the effectiveness of a ResNet-based face recognition system when tested across varying illumination conditions (Kim & Kittler, 2005). The research employs an experimental design, in which the system is directly tested using facial image data captured through a camera with varying light intensities (Huang et al., 2007; Jiang et al., 2021). The primary goal of this study is to examine the impact of illumination variations on the accuracy and consistency of a deep learning based face recognition system (Dey & Samanta, 2019).

In general, the developed face recognition system consists of several main stages: image acquisition, preprocessing (face detection and alignment), feature extraction using the ResNet50 architecture, matching, and identity determination (Schroff et al., 2015). The entire process is carried out in real time using the DeepFace library, which provides multiple pre-trained models for facial feature extraction, including ResNet50 (Baltrusaitis et al., 2016).

The overall system design is illustrated in Figure 2. The camera serves as the primary input device that captures the user's facial images. The captured image data then proceed to the detection and alignment stage to ensure that the face is properly located and proportionally positioned (Simonyan & Zisserman, 2015). Subsequently, the system extracts distinctive facial features using the ResNet50 model, producing a 2048-dimensional embedding vector that represents the numerical identity of the detected face (Deng et al., 2019; Liu et al., 2017).

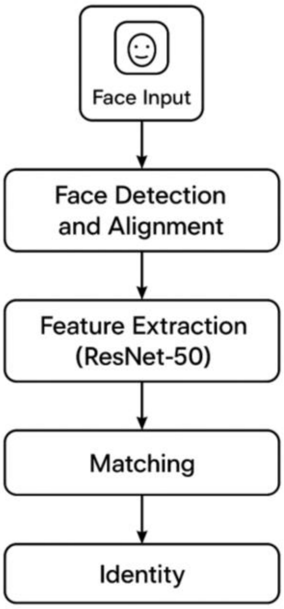

Fig. 2. Architecture diagram of the face classification system.

The extracted feature vectors are then compared with the facial data stored in the database (dataset) using the Euclidean distance calculation method. If the distance between the embeddings is below a predetermined threshold, the face is considered to match one of the registered identities in the database, and the system displays the user's name. If no match is found, the system labels the face as "Unknown."

The stages in this research design also include performance testing of the system under three different lighting conditions: bright light, normal light, and dim light. Each lighting condition was tested separately to obtain data on recognition accuracy and processing time. The results of each experiment were then analyzed to determine the effect of lighting variations on the performance of the ResNet50 model.

This design is intended to offer insights into the stability and robustness of the ResNet model across various lighting environments and to establish a basis for developing more adaptive, illumination-resilient face recognition systems.

System Workflow

The ResNet50 based face recognition workflow is structured to perform real-time individual identification using a camera as the main input device. The system utilizes the DeepFace library, which integrates multiple stages of the face recognition process, including detection, feature extraction, and identity matching.

In general, the system's operational workflow can be described through the following main stages:

1. Face Capture
   The camera captures real-time facial images of the user. Each captured image is temporarily stored in an image file format to be processed further by the system.
2. Face Detection and Alignment
   In this phase, the system identifies the position of the face in the image through the integrated face detection module available in the DeepFace library. After detecting the face, the system performs an alignment procedure to normalize facial orientation especially aligning the eyes and nose to enhance the accuracy of feature extraction (Li et al., 2015; Viola & Jones, 2001).
3. Feature Extraction Using ResNet50
   The aligned face image is subsequently processed by the ResNet50 model, a 50-layer convolutional neural network pre-trained on extensive facial datasets. The model generates a numerical representation of the face in the form of a 2048-dimensional embedding vector, which encodes the unique characteristics of each individual.
4. Matching Process
   The generated facial embedding is then matched against the stored dataset embeddings using similarity measures like Euclidean distance and cosine similarity. A lower distance value signifies a greater similarity between the input facial image and its corresponding record in the database.
5. Identity Recognition
   Based on the computed distance, the system determines whether the detected face matches one of the registered identities in the database. When the computed distance falls below a specified threshold, the system identifies and displays the individual's name; if not, the face is marked as "Unknown".
6. Display Output
   The final recognition result is displayed in the video window in real time, showing the identified individual's name or the label "Unknown."

ResNet50 Model Architecture

The ResNet50 (Residual Network 50 layers) architecture is one of the most popular Convolutional Neural Network (CNN) models developed by Kaiming He et al. (2015) from

Microsoft Research. This architecture is well known for its ability to overcome the vanishing gradient problem that commonly occurs in very deep neural networks. ResNet50 implements the principle of residual learning in its network architecture, in which shortcut connections or skip connections allow information and gradients to bypass one or more layers, ensuring stable and efficient training of deep networks.

This study employs the ResNet50 architecture as the primary model for extracting facial features, which are encoded as high dimensional numerical vectors uniquely characterizing each person. The feature extraction process is performed on facial images that have already undergone preprocessing, including face detection and alignment, to ensure optimal consistency before being input into the model.

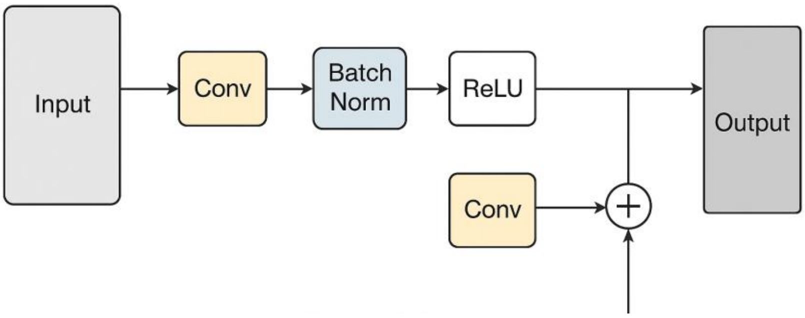

Fig. 3. ResNet50 block diagram.

ResNet50 is composed of 50 layers that include convolutional, batch normalization, ReLU activation, and fully connected layers. Its primary structural element, the residual block, integrates two or three convolutional layers with an identity shortcut that directly links the input to the block's output. This mechanism allows the model to maintain strong feature learning capability even as the network depth increases, avoiding performance degradation.

Within facial recognition applications, ResNet50 demonstrates strong effectiveness owing to its robustness in extracting distinctive facial features under varying illumination, poses, and expressions. These capabilities make ResNet50 highly suitable for real-time face recognition systems. In this study, the model is implemented via the DeepFace framework, utilizing a pre-trained ResNet50 trained on extensive datasets like VGGFace2 and MS-Celeb-1M to ensure robust generalization and dependable feature encoding.

Research Dataset

In this research, the dataset functions as a critical element for training and assessing the overall performance of the face recognition model. It serves as the reference database for the identity matching process (face matching). The dataset contains facial images collected from various individuals, exhibiting a range of expressions and illumination conditions.

The dataset was manually created by capturing facial images in real-time using a webcam. Each individual was photographed under three different lighting conditions bright light, normal light, and dim light to simulate real world environments. For each condition, a minimum of 10 facial images per individual were collected, resulting in a total of 30 images per person.

Each facial image was saved in JPEG (.jpg) format and labeled according to the corresponding identity. The dataset structure follows the directory format recognized by the DeepFace library, consisting of a main dataset folder and subfolders for each individual identity.

Before being used in testing, all images underwent a preprocessing stage that included *face detection*, *alignment*, and resizing to 224×224 pixels to match the input requirements of the ResNet50 model. This step ensured uniformity across all samples, enabling accurate and consistent feature extraction.

Alongside the locally collected dataset, the DeepFace framework employs a pre-trained ResNet50 model that was initially trained on extensive public datasets, including VGGFace2 and MS-Celeb-1M. These datasets encompass thousands of individual identities captured under various lighting, pose, and expression variations, enabling the model to generalize well to new, unseen faces during evaluation.

By combining the locally captured dataset and the pre-trained model, the developed face recognition system is expected to achieve high accuracy and maintain reliable performance under various lighting conditions.

Experimental Environment, Method, and Evaluation Parameters

This study was carried out in an experimental setup intended to replicate real-world scenarios for deep learning based facial recognition systems. The testing environment encompasses both hardware and software setups, along with the experimental procedures used to assess the ResNet50 model's performance across different lighting conditions.

1. Experimental Environment

   The experiment was carried out using the following hardware and software specifications:

   - Prosesor: Ryzen 5 PRO 4650U
   - RAM: 32 GB
   - GPU: Features the AMD Radeon Graphics Processor with the Ryzen 5 PRO 4650U
   - Operating System: Windows 10 Pro 64-bit
   - Camera: Logitech C920 HD Webcam (1080p, 30 fps)
   - Programming Language: Python 3.10
   - Libraries Used:
     - OpenCV (for image acquisition and real-time display)
     - DeepFace (for facial feature extraction and recognition)
     - NumPy, Pandas, dan Matplotlib (for data analysis and visualization)

The testing environment was set up in a controlled room with adjustable lighting using white LED lamps rated at 10 W. The lighting intensity was manually varied to represent three different conditions: bright (≈1000 lux), normal (≈500 lux), and dim (≈150 lux).

2. Experimental Method

   The testing method was conducted using a direct experimental approach (real-time testing) with a camera. Every subject in the dataset was evaluated under three distinct illumination conditions. For each scenario, facial images were captured in real time and matched against the stored reference data in the database. To maintain reliability, each experiment was performed five times per lighting condition.

3. Evaluation Parameters

   System performance was assessed based on four primary metrics: Accuracy, Precision, Recall, and F1-Score. Furthermore, recognition time was measured as an additional factor to evaluate the system's real-time processing efficiency.

4. RESULT & DISCUSSION

Experimental Results

The experimental testing aimed to assess the ResNet50 model's performance in facial recognition across three lighting environments bright, normal, and dim. For each participant, the tests were repeated five times under every lighting condition. The results include recognition accuracy, average processing time, and the number of recognition errors (*misrecognition*). Comprehensive performance results of the system under the three lighting conditions are summarized in Table 1.

Table 1. Face Recognition System Performance Using ResNet50

Lighting Condition	Total Tests	Correct (TP)	Incorrect (FP/FN)	Accuracy (%)	Average Time (ms)
Bright	50	49	1	98	142
Normal	50	47	3	94	148
Dim	50	42	8	84	161

As shown in the table, the system achieved its highest accuracy of 98% under bright lighting, whereas performance dropped to 84% in dim light conditions. The average recognition time also tended to increase under lower lighting levels due to the model's difficulty in optimally extracting facial features. These results demonstrate that the ResNet50 model performs optimally in well-lit environments but experiences a noticeable decline in accuracy when illumination is reduced. Nevertheless, the system maintains recognition speeds within 140–160 milliseconds per frame, which confirms its capability to operate in real time. Figure 4 below illustrates a comparison of recognition accuracy across different lighting conditions.

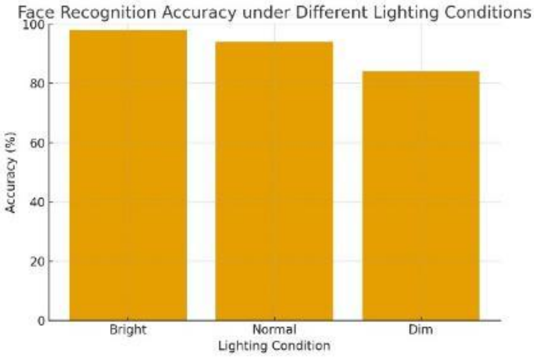

Fig. 4. Comparison of Accuracy under Different Lighting Conditions.

System Accuracy under Three Lighting Conditions, showing the performance degradation of the ResNet50 model as the light intensity decreases.

Analysis of Experimental Results

The experimental findings indicate that the performance of the ResNet50 model is strongly affected by variations in lighting conditions. In bright and normal illumination, facial attributes including eye shape, nose structure, and facial contours are effectively extracted, leading to a high level of recognition accuracy. However, under dim lighting, the image quality decreases and visual noise increases, leading to a higher rate of detection errors.

The ResNet50 model employs a residual learning approach that effectively overcomes performance degradation in deep neural networks. Nevertheless, since the system does not incorporate additional preprocessing techniques such as histogram equalization or illumination normalization, a reduction in accuracy still occurs under extreme lighting conditions.

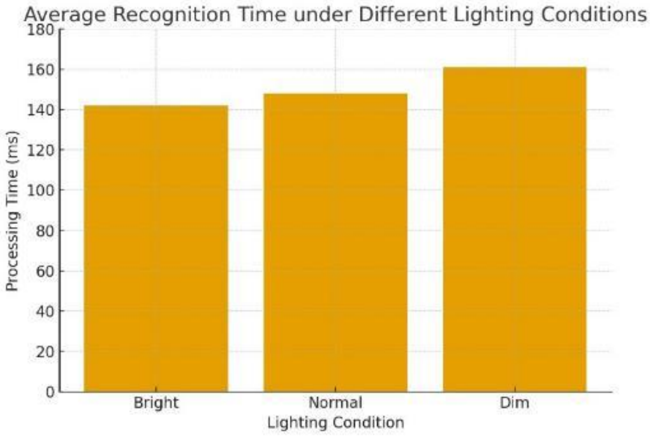

Fig. 5. Graph of average face recognition time under different lighting conditions.

Furthermore, the recognition time results indicate that although ResNet50 has a deep and complex architecture, the system is capable of operating with an average processing time of 140–160 milliseconds per frame, thus meeting the criteria for real-time recognition. Overall, the findings show that the ResNet50 architecture provides excellent face recognition performance under optimal lighting conditions, achieving an average accuracy of 92% across all tested lighting scenarios.

The graph in Figure 5 shows that as the lighting level decreases, the face recognition processing time increases due to the model's difficulty in optimally extracting facial features.

5. CONCLUSION

From the experimental outcomes, it can be inferred that ResNet50 achieves high recognition accuracy, performing best in environments with sufficient illumination. The model effectively extracts facial features through its residual learning mechanism, enabling it to maintain high accuracy without experiencing performance degradation despite its deep network structure. Experimental findings reveal that the system attained its peak accuracy of 98% under bright illumination, whereas performance dropped to 84% in dim lighting as a result of lower facial image quality.

Regarding efficiency, the system achieved an average recognition time of 140–160 milliseconds, indicating its capability to perform real-time operations on mid tier hardware. For future development, it is recommended that the system be enhanced with illumination preprocessing techniques, such as histogram equalization or adaptive gamma correction, to improve robustness against lighting intensity variations. Furthermore, increasing the diversity of the dataset including the number of individuals, facial expressions, and image capture angles would strengthen the model's generalization capability.

Future research could also explore alternative architectures, such as VGGFace2, ArcFace, or MobileFaceNet, and implement the system on edge computing devices such as Raspberry Pi or Jetson Nano to evaluate performance under resource-constrained real-world conditions. With these advancements, the ResNet50-based face recognition system is expected to operate more adaptively, efficiently, and accurately across various lighting conditions in real-world environments.

6. ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Politeknik Negeri Bengkalis, particularly to the Center for Research and Community Service (P3M) Politeknik Negeri Bengkalis, for the support and facilities provided during the completion of this research.

7. REFERENCES

Abate, A. F., Nappi, M., Riccio, D., & Sabatino, G. (2007). 2D and 3D face recognition: A survey. *Pattern Recognition Letters*, *28*(14), 1885–1906. https://doi.org/10.1016/j.patrec.2006.12.018

Alzahra, A. A., & Mahmood, M. R. (2023). Face recognition using deep residual networks and transfer learning. *International Journal of Advanced Computer Science and Applications*, *14*(2), 75–82.

Baltrusaitis, T., Robinson, P., & Morency, L. P. (2016). OpenFace: An open source facial behavior analysis toolkit. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 1–10. https://doi.org/10.1109/WACV.2016.7477553

Chollet, F. (2018). *Deep Learning with Python*. Manning Publications.

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. (2019). ArcFace: Additive angular margin loss for deep face recognition. *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4685–4694. https://doi.org/10.1109/CVPR.2019.00482

Dey, R., & Samanta, D. (2019). Face recognition using deep learning: A survey. *International Journal of Engineering and Advanced Technology*, *8*(6), 2278–3075.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Huang, G. B., Ramesh, M., Berg, T., & Learned-Miller, E. (2007). *Labeled Faces in the Wild: A database for studying face recognition in unconstrained environments* (Issues 07–49).

Jiang, Y., Li, W., & Zhao, H. (2021). Improving illumination robustness in face recognition using adaptive histogram equalization. *Sensors*, *21*(8), 2612. https://doi.org/10.3390/s21082612

Kim, T. K., & Kittler, J. (2005). Locally linear discriminant analysis for multimodally distributed classes for face recognition with a single model image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(3), 318–327.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems (NeurIPS)*, *25*, 1097–1105.

Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 5325–5334.

Liu, W., Wen, Y., Yu, Z., Li, M., Raj, B., & Song, L. (2017). SphereFace: Deep hypersphere embedding for face recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 6738–6746.

Masi, I., Wu, Y., Hassner, T., & Natarajan, P. (2018). Deep face recognition: A survey. *Proceedings of the IEEE*, *106*(8), 1283–1313. https://doi.org/10.1109/JPROC.2018.2816039

Parkhi, O. M., Vedaldi, A., & Zisserman, A. (2015). Deep face recognition. *British Machine Vision Conference (BMVC)*, 1–12.

Purwono, Ma'arif, A., Rahmaniar, W., Fathurrahman, H. I. K., Frisky, A. Z. K., & Haq, Q. M. U. (2022). Understanding of Convolutional Neural Network (CNN): A Review. *International Journal of Robotics and Control Systems*, *2*(4), 739–748. https://doi.org/10.31763/ijrcs.v2i4.888

Schroff, F., Kalenichenko, D., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823. https://doi.org/10.1109/CVPR.2015.7298682

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *International Conference on Learning Representations (ICLR)*.

Tan, X., & Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, *19*(6), 1635–1650.

Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 511–518.

Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, *429*, 215–244. https://doi.org/10.1016/j.neucom.2020.10.081

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, *35*(4), 399–458. https://doi.org/10.1145/954339.954342