ISSN: 3021-7636
International Journal of Science Technology and Health

Research Article
Vol. 4 (2), 2026, page 77-87
DOI: https://doi.org/10.69930/ijsth.v4i2.72

Analysis of the Effect of Lighting Variations on the Performance of a Face Recognition System Based on FaceNet and MTCNN

Putra Fajar Sidik *, Sahriani, Syaiful Bachri Mustamin

Department of Technology Information, Faculty of Science Technology and Health, Institut Sains Teknologi dan Kesehatan 'Aisyiyah Kendari, Indonesia

* Correspondence: fsjarr01@gmail.com

Received : 26 June 2026	Accepted: 28 July 2026	Published: 31 July 2026

Abstract: Face recognition is a biometric technology widely used in automated attendance and access control systems. However, variations in lighting conditions pose a significant challenge that can substantially affect system accuracy in real-world environments. This study aims to analyze the effect of lighting variation on the performance of an MTCNN and FaceNet-based face recognition system and to identify the lighting conditions that yield the best and worst recognition results. This study employed an experimental systems engineering approach using a dataset of 50 facial images collected from 10 students of the Information Technology Study Program at ISTEK 'Aisyiyah Kendari under five lighting conditions: very bright (>1,000 lux), bright, normal, dim, and very dim (<100 lux). MTCNN was used for face detection and alignment, FaceNet for extracting 128-dimensional embeddings, and Euclidean Distance with a threshold of 0.8 for identification. System performance was evaluated using accuracy, precision, recall, F1-score, and confusion matrix. The results indicate that the system achieved its best performance under very bright conditions, with an accuracy of 70.00%, precision of 58.33%, recall of 70.00%, and an F1-score of 61.67%. Under very dim conditions, all metrics declined by 10%, with a higher average Euclidean Distance (0.8922) compared to very bright conditions (0.8558). Lighting variation was found to have a significant and cascading impact on system performance, affecting image quality, embedding stability, and final classification outcomes. Future development is recommended to incorporate image enhancement techniques such as CLAHE or histogram equalization to improve system robustness under extreme lighting conditions.

Keywords: Face Recognition; FaceNet; MTCNN; Lighting; Euclidean Distance; Deep Learning

1. Introduction

The rapid advancement of Artificial Intelligence (AI) has brought significant transformation to the field of computer vision, particularly in face recognition — the process of identifying individuals based on facial characteristics captured by a camera. This technology has been widely adopted in security systems, access control, and automated attendance due to its contactless, fast, and efficient nature (Sugeng & Barus, 2023). A face recognition-based attendance system generally consists of two core stages: face detection and face recognition, where the accuracy of the detection stage critically determines the overall system performance.

Previous studies have explored conventional methods for face recognition-based attendance systems, each with notable limitations. Indra et al. (2019) developed a student attendance system using the Haar-Like Feature method, which proved effective in accelerating data recapitulation but demonstrated significant weaknesses under unstable lighting conditions and varied facial poses. Similarly, Fatih et al. (2022) implemented the Eigenface method based on Principal Component Analysis (PCA), achieving accuracy rates of only 75–91% under ideal lighting, with performance degrading substantially in low-light environments. These findings indicate that conventional approaches lack robustness when faced with real-world environmental variation.

With the emergence of deep learning, Convolutional Neural Network (CNN)-based approaches have gained considerable attention due to their superior ability to extract complex facial features. Febiyani et al. (2025) implemented the MobileNetV1 architecture for a face recognition attendance system and achieved accuracy up to 94%, demonstrating a clear advantage over conventional methods. However, most prior studies were conducted under controlled lighting conditions and did not fully reflect the challenges present in real campus environments, such as fluctuating illumination, non-frontal facial poses, and limited training data highlighting a significant research gap that warrants further investigation.

To address these limitations, this study proposes the use of Multi-Task Cascaded Convolutional Neural Networks (MTCNN) as the face detection method, leveraging its three-stage network architecture (P-Net, R-Net, and O-Net) along with facial landmark-based alignment to enhance robustness against pose and lighting variation. Face recognition is subsequently performed using FaceNet, which generates 128-dimensional embedding vectors through distance-based recognition, enabling accurate identification even with limited training samples. This study aims to evaluate the effect of lighting variation on the performance of the FaceNet-MTCNN system among Information Technology students at ISTEK 'Aisyiyah Kendari, with the goal of providing a foundation for developing more accurate and reliable face recognition systems in real-world environments.

2. Methods

This study employed an experimental research design to evaluate the effect of lighting variation on the performance of a face recognition system combining MTCNN and FaceNet. The dataset consisted of 50 facial images collected from 10 Information Technology students at ISTEK 'Aisyiyah Kendari, with each subject photographed under five distinct lighting conditions: very bright, bright, normal, dim, and very dim, measured by lux intensity. Images were captured using a smartphone camera at a consistent frontal pose and distance of 50–70 cm against a plain background, with no accessories worn. The dataset was divided into 30 images for building the embedding database (bright, normal, and dim conditions) and 20 images for testing (very bright and very dim conditions). Face detection and alignment were performed using MTCNN through its three-stage cascade network (P-Net, R-Net, and O-Net), followed by cropping and resizing to 160 × 160 pixels. Feature extraction was conducted using a pretrained FaceNet model, which generates 128-dimensional embedding vectors representing each individual's facial characteristics. Identity recognition was performed by computing the Euclidean Distance between test embeddings and database embeddings, with a threshold of 0.8 applied to determine whether a face was recognized or not. System performance was evaluated using accuracy, precision, recall, F1-score, confusion matrix, and Euclidean Distance analysis across all lighting conditions. All data collection procedures were conducted with the informed consent of the participants, and ethical approval for this study was granted by the Research Ethics Committee of ISTEK 'Aisyiyah Kendari (Approval Code: [insert code]).

3. Results and Discussion

3.1. System Implementation and Face Detection Results

The face recognition system in this study was implemented using the Python programming language on the Google Colab platform, which was selected for its cloud-based GPU support that accelerates image processing and deep feature extraction. Several supporting libraries were employed, including OpenCV for digital image processing, MTCNN for face detection and landmark localization, keras-facenet for embedding extraction, and Scikit-learn for performance metric computation. The dataset consisted of 50 facial images collected from 10 students of the Information Technology Study Program, captured under five lighting conditions very bright, bright, normal, dim, and very dim measured using a lux-meter application to ensure consistent categorization. This implementation environment formed the technical foundation for every subsequent stage of the research, from dataset acquisition to final performance evaluation.

Face detection was performed using MTCNN, which simultaneously produced bounding boxes, five facial landmarks (left eye, right eye, nose, and the left and right mouth corners), and cropped face regions for each input image. A representative test on a sample image captured under dim lighting showed that MTCNN successfully detected the face with a confidence score of 0.9536, indicating that the detector remained functional even under suboptimal illumination. The landmarks obtained from this stage were subsequently used as the geometric reference for the face alignment process prior to feature extraction. These results indicate that MTCNN's detection capability is reasonably robust to moderate variation in lighting intensity, which is consistent with its design as a cascaded multi-task detector.

However, closer inspection of the detection outputs revealed that bounding box and landmark quality became noticeably less stable under very dim conditions, with several images exhibiting higher noise levels due to reduced contrast between the face and the background. This instability at the detection stage subsequently affected the following preprocessing pipeline, since less precise bounding boxes led to less accurate cropping of the facial region. Consequently, although the system remained functionally operational across all five lighting categories, the reliability of its detection output was found to be highly dependent on the quality of the input image. This finding establishes the first link in what later analysis identifies as a cascading effect of illumination throughout the recognition pipeline.

Overall, the implementation results confirm that MTCNN is capable of detecting faces and landmarks effectively across a range of lighting conditions, while showing gradually degrading precision as illumination decreases. This pattern is consistent with prior studies emphasizing that, although cascaded CNN-based detectors are generally robust, their performance still partially depends on image contrast quality. The detection stage therefore functions as a critical determinant of the quality of the entire recognition pipeline rather than as an isolated preprocessing step. These results set the stage for the preprocessing and embedding extraction analyses discussed in the following sections.

This study is consistent with Zhang et al. (2016), who developed MTCNN as a multi-task face detection method that simultaneously detects bounding boxes and facial landmarks, enabling a more consistent alignment process despite variations in visual conditions, including illumination Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. IEEE Signal Processing Letters, 23(10), 1499-1503. This study is also consistent with the findings of Beveridge et al. (2010), who showed that lighting and focus quality are among the most influential factors causing degraded face detection and recognition performance, which aligns with the finding that MTCNN's bounding box and landmark quality became less stable under very dim conditions in this study Beveridge, J. R., Bolme, D. S., Draper, B. A., Givens, G. H., Lui, Y. M., & Phillips, P. J. (2010). Quantifying how lighting and focus affect face recognition performance. 2010 IEEE CVPRW, 74–81.

Table 1. Lighting Condition Categories Used in the Dataset

No.	Lighting Condition	Light Intensity (Lux)
1	Very Bright	> 1000
2	Bright	500 – 1000
3	Normal	250 – 500
4	Dim	100 – 250
5	Very Dim	< 100

3.2. Embedding Extraction Results Using FaceNet

Following preprocessing, each facial image was transformed into a 128-dimensional embedding vector using the FaceNet model, which served as the numerical representation of each individual's unique facial characteristics. The embedding extraction process used standardized 160 × 160-pixel images as input and produced an output vector of shape (1, 128) for every sample. Sample embedding values for several students under different lighting categories showed distinct dimension values across individuals, confirming that each identity produced a unique numerical signature. This vector-based representation enabled the system to compare facial similarity numerically rather than through direct pixel-level image comparison.

The use of 128-dimensional embeddings was selected because this dimensionality allows a richer and more accurate representation of facial characteristics compared with lower-dimensional alternatives, while remaining computationally efficient for similarity calculation. Each of the 128 dimensions encodes a portion of the facial features learned by the deep learning model, and the combination of all dimensions forms a unique digital identity for every individual in the dataset. This design choice directly supports the subsequent Euclidean Distance calculation used for identity matching. The resulting embeddings were stored in a .npy file to serve as the reference database for the testing phase.

Analysis of the embedding results revealed a clear pattern: embeddings generated under well-lit conditions showed smaller variation across multiple images of the same individual, indicating higher representational stability. Conversely, embeddings produced under very dim lighting exhibited greater variability, suggesting that the FaceNet model could not fully capture stable facial features when visual information was limited. This instability in the embedding space is a direct consequence of the degraded input quality identified at the detection and preprocessing stages. As a result, the reliability of the embedding-based identification process is shown to be tightly coupled with input illumination quality.

These results affirm that FaceNet is capable of producing discriminative facial representations, but its effectiveness is conditional on adequate input image quality. The embedding stage functions as the critical link between raw visual data and the distance-based classification performed in later stages. Because the embedding directly determines the Euclidean Distance value used for recognition decisions, any instability introduced here propagates further into accuracy degradation. This relationship is examined in more detail in the following subsections on distance and performance evaluation.

This study is consistent with Schroff and Philbin (2015), who introduced FaceNet as a triplet-loss-based facial embedding method, in which facial representations are mapped into a 128-dimensional vector space so that facial similarity can be measured directly using Euclidean distance an approach mirrored in this study Schroff, F., & Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. Proceedings of the IEEE CVPR.. This study is also consistent with Arachchilage (2020), who explained that the stability of facial embeddings is strongly influenced by the quality of the input image processed by the deep learning model, which aligns with the finding that embeddings under very dim conditions in this study showed greater variability compared to very bright conditions Arachchilage, S. P. K. W. (2020). Deep-learned faces: a survey. EURASIP Journal on Image and Video Processing.

3.3. System Performance Evaluation

System testing was conducted by comparing the embeddings of test images captured under very bright and very dim conditions against a reference database built from images captured under bright, normal, and dim conditions, using a Euclidean Distance threshold of 0.8. The results showed that the system achieved its best performance under very bright conditions, with an accuracy of 70.00%, precision of 58.33%, recall of 70.00%, and an F1-score of 61.67%. Under very dim conditions, all four metrics declined by exactly 10 percentage points, yielding an accuracy of 60.00%, precision of 48.33%, recall of 60.00%, and an F1-score of 51.67%. This consistent and uniform decline across all metrics highlights the systematic nature of illumination's impact on recognition performance, as summarized in Table 2 and illustrated in Figure 1.

Table 2. System Performance Evaluation Across Lighting Conditions

Lighting Condition	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Very Bright	70.00	58.33	70.00	61.67
Very Dim	60.00	48.33	60.00	51.67
Difference	10.00	10.00	10.00	10.00

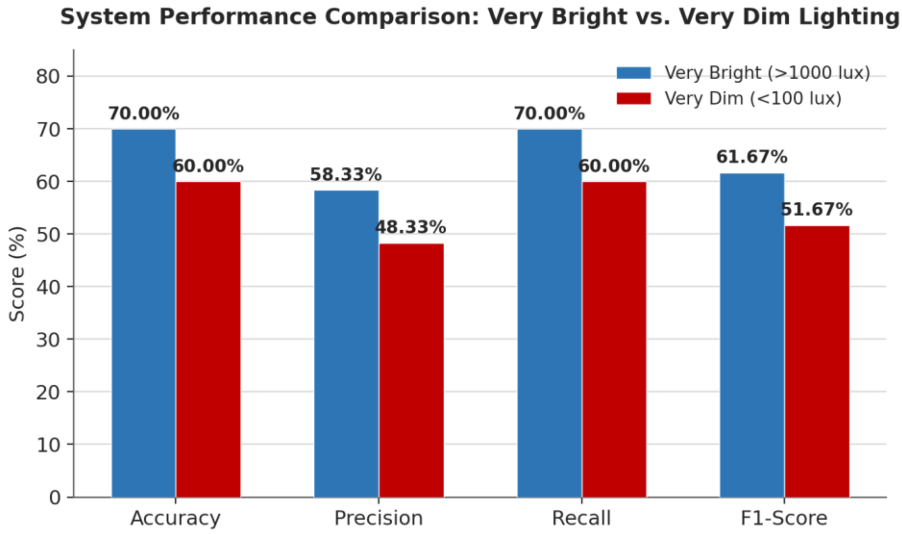

Figure 1. System performance comparison between very bright and very dim lighting conditions across all evaluation metrics.

This uniform 10-percentage-point drop across every evaluation metric is a particularly notable result, as it suggests that the effect of poor lighting is not isolated to a single aspect of classification such as only false positives or only recall but rather degrades the model's overall discriminative capability uniformly. Such a pattern reinforces the idea that illumination affects the underlying embedding quality rather than any single decision boundary in isolation. The decline in precision indicates that more faces were incorrectly matched to the wrong identity under poor lighting, while the equal decline in recall shows that more true identities failed to be correctly recognized. Together, these results paint a coherent picture of system-wide degradation rather than an isolated metric failure.

The selection of a 0.8 threshold played an important role in these outcomes, as several borderline distance values close to this cutoff caused the system to misclassify recognized faces as unrecognized, or vice versa. This sensitivity to threshold selection is an inherent limitation of distance-based recognition systems and becomes more pronounced when embedding quality is already compromised by poor lighting. It also explains why certain test cases with distance values just above or below 0.8 resulted in classification errors despite belonging to correctly extracted embeddings. This observation suggests that threshold tuning could be a useful area for further optimization in future iterations of the system.

Taken together, the performance evaluation results confirm that lighting conditions have a measurable and significant effect on the accuracy, precision, recall, and F1-score of the MTCNN-FaceNet-based face recognition system. The consistent 10-percentage-point performance gap between the best and worst lighting conditions provides strong empirical evidence for the cascading influence of illumination throughout the recognition pipeline. These findings are essential for understanding the practical limitations of deploying such systems in real-world environments with uncontrolled lighting. The implications of this performance gap are further explored through Euclidean Distance and confusion matrix analysis in the following sections.

This study is consistent with Beveridge et al. (2010), who demonstrated that lighting conditions are among the most influential factors causing reduced accuracy in image-based face recognition systems, aligning with the finding in this study that all evaluation metrics (accuracy, precision, recall, and F1-score) declined by 10% under very dim conditions compared to very bright conditions Beveridge, J. R., Bolme, D. S., Draper, B. A., Givens, G. H., Lui, Y. M., & Phillips, P. J. (2010). Quantifying how lighting and focus affect face recognition performance. 2010 IEEE CVPRW, 74–81.. This study is also consistent with research combining MTCNN, FaceNet, and the Local Binary Pattern (LBP) method, which reported that the performance of FaceNet-based face recognition systems consistently declined under low-light conditions compared to normal lighting, thereby reinforcing the generalizability of this finding across other FaceNet-based architectures.

3.4. Euclidean Distance and Confusion Matrix Analysis

Euclidean Distance analysis was conducted to quantify the similarity between test embeddings and reference embeddings stored in the database, where a smaller distance indicates higher facial similarity. The results showed that the average Euclidean Distance under very bright conditions was 0.8558, while under very dim conditions it increased to 0.8922, as shown in Table 3 and Figure 2. This increase in average distance under poor lighting directly corresponds to the decline in accuracy observed in the performance evaluation, confirming that distance-based metrics and classification metrics move in tandem. Individual student-level comparisons further revealed considerable variation, with some identities such as RIZ02 and FER05 showing particularly large distance increases under dim lighting, suggesting uneven sensitivity across individuals.

Table 3. Average Euclidean Distance by Lighting Condition

Lighting Condition	Average Euclidean Distance
Very Bright	0.8558
Very Dim	0.8922

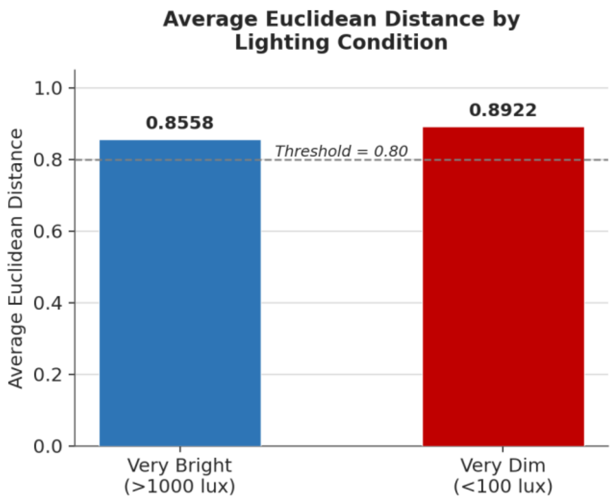

Figure 2. Average Euclidean Distance under very bright versus very dim lighting conditions, relative to the classification threshold (0.80).

The confusion matrix analysis provided a more granular view of classification behavior, with the system correctly recognizing 13 of 20 test samples (65% accuracy) overall, as detailed in Table 4 and Figure 3. Several identities including Dwi Anggriani, Kevin, Lisa Ardianti, Putra Fajar Sidik, and Rizkyawan achieved 100% recognition accuracy, while two identities (Eka Purnama Sari and Ferdiansyah) were entirely misclassified. This uneven distribution of errors across identities indicates that embedding stability is not uniform across all individuals, possibly due to inherent facial feature similarities between certain subjects. The RIZ02 class exhibited a notably high false-positive count (FP = 4), meaning several other individuals were incorrectly classified as RIZ02 under degraded lighting conditions.

Table 4. Confusion Matrix Summary per Identity Class (n = 20 Test Samples)

Class	TP	FP	FN	TN
AMT010	1	1	1	17
DWA09	2	1	0	17
EKP04	0	0	2	18
FER05	0	1	2	17
KEV07	2	0	0	18
LIS08	2	0	0	18
RID03	1	0	1	18
FAJ06	2	0	0	18
RIZ02	2	4	0	14
SAM01	1	0	1	18

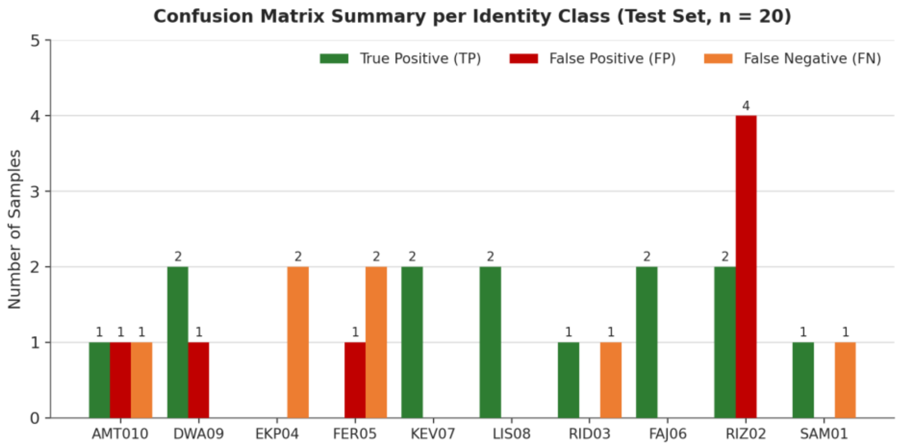

Figure 3. Distribution of True Positive (TP), False Positive (FP), and False Negative (FN) counts per identity class.

Further inspection revealed that classification errors were dominated by false positives and false negatives concentrated primarily under very dim lighting conditions, rather than being evenly distributed across both extreme lighting categories. False positives occurred when the system incorrectly matched a face to the wrong identity, while false negatives occurred when a genuine identity failed to be matched to its correct database entry. This pattern is consistent with the earlier finding that reduced illumination compromises embedding stability, making certain facial vectors fall closer to incorrect reference points than to their own correct reference. The identity-specific variation in error rates further suggests that some individuals' facial features are inherently more vulnerable to illumination-induced embedding drift than others.

Collectively, the Euclidean Distance and confusion matrix results provide complementary evidence for the cascading effect of illumination: as lighting deteriorates, the embedding-to-database distance increases, and this increase directly translates into a higher rate of misclassification, particularly false positives and false negatives. These two analyses, taken together with the performance metrics in the previous section, strongly support the conclusion that illumination is not a peripheral factor but a central determinant of system reliability. The identity-specific error patterns also point to an important avenue for future research: examining why certain individuals are more susceptible to illumination-induced misclassification than others. This sets up the broader discussion of illumination's systemic impact presented in the next section.

This study is consistent with Zhang et al. (2019), who stated that shadows caused by non-uniform lighting are a primary cause of reduced identification accuracy in deep-learning-based face recognition systems, aligning with the finding in this study that the average Euclidean Distance increased from 0.8558 under very bright conditions to 0.8922 under very dim conditions Zhang, W., Zhao, X., Morvan, J. M., & Chen, L. (2019). Improving Shadow Suppression for Illumination Robust Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(3), 611–624.. This study is also consistent with research on facial image quality, which explained that poor lighting can obscure key biometric elements such as the eyes, nose, and mouth through shadowing, making feature extraction more difficult and leading to increased classification errors such as false positives and false negatives as observed in the confusion matrix results of this study.

3.5. Discussion on the Cascading Effect to Illumination

Synthesizing the results across all stages of the pipeline reveals a clear cascading pattern, illustrated schematically in Figure 4: illumination quality affects raw image quality, which in turn affects preprocessing outcomes (cropping, resizing, normalization, and alignment), which then affects the stability of FaceNet embeddings, which subsequently determines Euclidean Distance values, and finally shapes the classification outcome reflected in accuracy, precision, recall, F1-score, and confusion matrix results. This sequential dependency means that even small degradations introduced early in the pipeline such as noisier bounding boxes under dim lighting can compound into substantial performance losses by the final classification stage. Understanding this cascading relationship is critical for diagnosing where future improvements should be targeted.

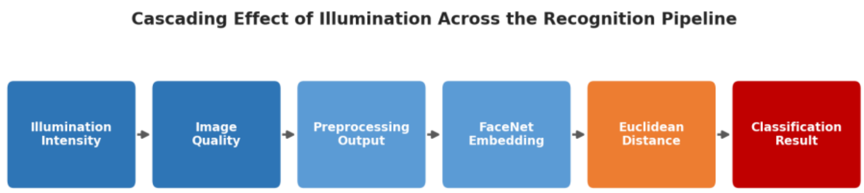

Figure 4. Schematic representation of the cascading effect of illumination across the face recognition pipeline.

The discussion further highlights that bright lighting conditions consistently produced more stable embeddings, smaller Euclidean Distances, and higher classification accuracy, while dim conditions produced the opposite pattern across every measured indicator. This consistency across multiple, independently calculated metrics rather than just one provides strong internal validation that the observed effect is genuinely attributable to illumination rather than incidental noise in the experiment. It also confirms that the threshold-based Euclidean Distance approach used in this study is sensitive enough to detect meaningful differences in embedding quality caused by lighting variation. This makes the dataset and methodology suitable as a benchmark for evaluating illumination robustness in similar systems.

From a practical standpoint, these findings imply that deploying MTCNN-FaceNet-based face recognition systems in environments with uncontrolled or variable lighting such as outdoor attendance systems or mobile access control carries a tangible risk of degraded reliability, particularly during low-light periods such as early morning, evening, or indoor areas with insufficient illumination. This has direct implications for system design, suggesting that supplementary lighting infrastructure or adaptive image enhancement techniques should be considered standard requirements rather than optional improvements. The uniform 10-percentage-point metric decline observed in this study offers a quantifiable benchmark that future system designers can use to estimate the performance risk associated with deploying in suboptimal lighting environments.

Based on these findings, the study recommends future development directions, including the application of image enhancement techniques such as CLAHE (Contrast Limited Adaptive Histogram Equalization) or histogram equalization prior to the embedding extraction stage, with the goal of improving system robustness against extreme lighting conditions. Data augmentation strategies that simulate varied lighting during training could also help the FaceNet model generalize better to illumination changes it has not explicitly encountered. Future studies might also explore adaptive thresholding strategies that adjust the 0.8 cutoff value dynamically based on estimated input image quality, rather than relying on a single fixed threshold. These directions collectively aim to mitigate the cascading vulnerability to illumination identified throughout this study's results.

This study is consistent with Zhang et al. (2019), who showed that shadow suppression techniques can improve the robustness of deep-learning-based face recognition systems against illumination variation, aligning with this study's recommendation to apply image enhancement techniques such as CLAHE or histogram equalization to address the identified cascading effect of illumination Zhang, W., Zhao, X., Morvan, J. M., & Chen, L. (2019). Improving Shadow Suppression for Illumination Robust Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(3), 611–624.. This study is also consistent with research combining MTCNN and FaceNet with an adaptive gamma correction method at the preprocessing stage, which demonstrated that improving image quality before feature extraction can significantly increase face recognition accuracy compared to unenhanced images a direction also recommended in this study for future development.

Conclusions

Based on the results and discussion presented above, it can be concluded that illumination variation has a significant and cascading effect on the performance of the MTCNN- and FaceNet-based face recognition system, from the face detection stage through embedding extraction to the final classification outcome. The system achieved its best performance under very bright conditions, with an accuracy of 70.00%, precision of 58.33%, recall of 70.00%, and an F1-score of 61.67%, while under very dim conditions all of these metrics uniformly declined by 10 percentage points, supported by an increase in average Euclidean Distance from 0.8558 to 0.8922 and a rise in classification errors observed in the confusion matrix analysis, with an overall accuracy of 65%. These findings indicate that the system still has limitations in handling extreme lighting, highlighting the need for image enhancement techniques such as CLAHE or histogram equalization at the preprocessing stage to improve system robustness in environments with uncontrolled lighting.

Conflicts of Interest

The authors declare no conflict of interest.

References

Arachchilage, S. P. K. W. (2020). *Deep-learned faces: a survey*. EURASIP Journal on Image and Video Processing.

Beveridge, J. R., Bolme, D. S., Draper, B. A., Givens, G. H., Lui, Y. M., & Phillips, P. J. (2010). Quantifying how lighting and focus affect face recognition performance. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – Workshops, CVPRW 2010*, 74–81. https://doi.org/10.1109/CVPRW.2010.5543228

Fatih, R. H., Kurniawan, Y., Informatika, P. T., Pamulang, U., & Jari, S. (2022). *Rancang bangun sistem absensi pengenalan wajah (face recognition) dan lokasi berbasis android (Studi Kasus: PT. Media Pariwara Indonesia)*. 7.

Febiyani, E., Pradana, Z. H., & Permatasari, I. (2025). Analisis sistem monitoring presensi menggunakan face recognition berbasis CNN dengan arsitektur MobileNetV1. *2*, 95–102.

Indra, E., Batubara, M. D., Yasir, M., & Chau, S. (2019). Desain dan implementasi sistem absensi mahasiswa berdasarkan fitur pengenalan wajah dengan menggunakan metode Haar-Like Feature. *2*, 363–370.

Schroff, F., & Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE CVPR*.

Sugeng, W., & Barus, D. (2023). Pengecekan foto paspor menggunakan metode DNN dan FaceNet sebagai pengenalan wajah. 169–180.

Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, *23*(10), 1499–1503. https://doi.org/10.1109/LSP.2016.2603342

Zhang, W., Zhao, X., Morvan, J. M., & Chen, L. (2019). Improving shadow suppression for illumination robust face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *41*(3), 611–624. https://doi.org/10.1109/TPAMI.2018.2803179