Green Intelligent Systems and Applications

Research Article
Volume 6(1), 2026, 161−173
https://doi.org/10.53623/gisa.v6i1.1197

Preliminary Study of ResNet-Based Facial Identification for Access Control Systems

Andre Kurniawan*, Ery Hartati

Faculty of Computer Science and Engineering, Universitas Multi Data Palembang, South Sumatra, Indonesia

*Correspondence: andrekurniawan@mhs.mdp.ac.id

SUBMITTED: 17 May 2026; REVISED: 14 June 2026; ACCEPTED: 16 June 2026

ABSTRACT: Facial recognition is an important application of artificial intelligence (AI) and computer vision in modern security systems, particularly for automated access control. This study aimed to implement and evaluate a Residual Network (ResNet)-based facial identification system for door access control applications. A publicly available facial image dataset was used and was divided into training (70%), validation (20%), and testing (10%) subsets. The proposed methodology consisted of data preprocessing, ResNet-based model training, and performance evaluation using accuracy and loss metrics. The model was trained for 10 epochs to assess its initial learning capability. The experimental results showed relatively low performance, with training accuracy ranging from 3.6% to 3.8% and validation accuracy of approximately 3.6%, while loss values remained high throughout the training process. These findings indicated that the model was unable to effectively learn discriminative facial features from the dataset and exhibited signs of underfitting. The limited performance was likely associated with insufficient dataset diversity, suboptimal preprocessing procedures, and non-optimized training parameters. The study highlighted the challenges of implementing ResNet-based facial recognition systems under constrained training conditions. Future work should focus on expanding the dataset, applying data augmentation techniques, optimizing hyperparameters, and utilizing pretrained models to improve recognition performance and system reliability.

KEYWORDS: Face recognition; ResNet; deep learning; access control; computer vision

1. Introduction

The rapid development of information technology and artificial intelligence (AI) had significantly transformed various sectors, particularly security systems. Conventional access control methods such as physical keys, access cards, and passwords were still widely used; however, these approaches had several limitations, including the risk of loss, theft, duplication, and unauthorized access. As security requirements continued to increase, there was a growing demand for authentication systems that were more secure, efficient, and convenient for users [1, 2].

Biometric technologies had emerged as promising alternatives to traditional authentication methods because they utilized unique physiological or behavioral characteristics for identity verification. Among various biometric approaches, facial recognition had gained considerable attention due to its non-contact nature, ease of use, and ability to perform automatic identification in real time. Facial recognition systems analyzed facial features captured by cameras and compared them with stored facial information to determine an individual's identity [3, 4].

Recent advances in deep learning had significantly improved the performance of facial recognition systems. Convolutional Neural Networks (CNNs) had become one of the most widely used approaches because of their capability to automatically extract complex and discriminative features from image data. Among CNN-based architectures, Residual Network (ResNet) had demonstrated excellent performance by addressing the vanishing gradient problem through residual learning and skip connections, allowing deeper neural networks to be trained more effectively [1, 5]. Consequently, ResNet had been successfully applied in various face recognition applications and had achieved high accuracy on several benchmark datasets [6, 7].

Despite these achievements, implementing facial recognition systems in real-world entrance control applications remained challenging. Unlike controlled laboratory environments, practical systems had to operate under varying lighting conditions, pose differences, facial expression changes, image quality degradation, and partial occlusions caused by accessories such as glasses or masks [8, 9]. Furthermore, many real-world implementations were constrained by limited training datasets, imbalanced data distributions, and hardware limitations, which could significantly affect recognition accuracy and system reliability [10]. These challenges often resulted in a substantial performance gap between experimental studies and practical deployment scenarios.

Previous studies had reported promising results for ResNet-based facial recognition systems, particularly in smart security and door access control applications [6, 7]. However, many of these studies primarily focused on achieving high accuracy using well-curated datasets under controlled conditions. Limited attention had been given to evaluating the practical limitations of ResNet-based systems when faced with dataset constraints, environmental variations, and real-world operational conditions. Therefore, further investigation was required to understand how such factors influenced system performance and deployment feasibility.

Based on these challenges, this study aimed to implement and evaluate a ResNet-based facial identification system for automated entrance control. The proposed framework consisted of face detection, image preprocessing, feature extraction using the ResNet architecture, and identity matching for access verification. The system was trained and evaluated using a publicly available facial image dataset that contained variations in lighting conditions, facial expressions, and viewing angles to simulate practical usage scenarios.

This study was based on the hypothesis that the ResNet architecture could learn discriminative facial features for entrance control applications; however, its performance might be significantly influenced by dataset quality, data diversity, and training configurations. Therefore, evaluating the effectiveness and limitations of ResNet under constrained training conditions was necessary to determine its suitability for practical deployment.

The main contribution of this research was the implementation and performance evaluation of a ResNet-based facial recognition system for automated entrance control under realistic dataset conditions. In addition, this study identified key challenges related to dataset limitations, environmental variability, and model training, providing insights and recommendations for future improvements in facial recognition-based access control systems.

2. Materials and Methods

2.1. Facial recognition.

Facial recognition is a biometric identification technology in the field of computer vision that identifies or verifies individuals based on unique facial features extracted from digital images or video streams. Owing to its contactless operation, ease of use, and high level of user convenience, facial recognition has been widely adopted in applications such as access control, surveillance, smart security systems, and human–computer interaction [1, 2]. Early facial recognition systems relied on handcrafted feature extraction methods, including Eigenface, Fisherface, and Local Binary Pattern Histogram (LBPH), which utilized statistical and appearance-based representations of facial images [3,4]. While these approaches achieved acceptable performance under controlled conditions, their accuracy often declined in the presence of variations in lighting, facial expressions, head pose, occlusion, and image quality, limiting their robustness in real-world applications [5]. The emergence of deep learning has substantially advanced facial recognition technology. In particular, Convolutional Neural Networks (CNNs) have demonstrated superior performance by automatically learning hierarchical and discriminative facial features directly from image data [6]. Among CNN architectures, Residual Network (ResNet) has become one of the most widely adopted models due to its residual learning framework, which facilitates the training of deeper networks while mitigating the vanishing gradient problem. This capability enables more effective feature extraction and improves recognition accuracy [7]. As a result, ResNet-based facial recognition systems have been increasingly integrated into practical applications, including automated access control, smart surveillance, and biometric authentication, where reliable identification is required under diverse environmental and operational conditions [8,9].

2.2. Residual Network (ResNet).

Residual Network (ResNet) is a deep Convolutional Neural Network (CNN) architecture introduced by He et al. [1] to address the vanishing gradient problem that commonly occurs in very deep neural networks. As network depth increases, conventional CNNs often experience performance degradation due to difficulties in gradient propagation during training. ResNet mitigates this issue through residual learning, which enables the network to learn residual functions rather than directly mapping inputs to outputs. The core component of ResNet is the residual block, which incorporates skip (shortcut) connections that allow information and gradients to bypass one or more layers. These connections facilitate gradient flow during backpropagation, improve training stability, and enable the development of substantially deeper networks without significant loss of performance [1]. The output of a residual block can be expressed as:

$$F(x) + x$$

where (x) represents the input, (F(x)) denotes the residual function learned by the network, and (y) is the output of the residual block.

By preserving information from earlier layers while learning increasingly complex feature representations, ResNet achieves improved convergence and feature extraction capabilities compared with conventional CNN architectures. Consequently, ResNet has become a widely adopted backbone model in numerous computer vision applications, including image classification, object detection, and facial recognition. In facial recognition systems, ResNet has demonstrated strong performance in learning discriminative facial features from large-scale image datasets. Its deep architecture enables robust feature extraction under challenging conditions, including variations in illumination, facial expression, head pose, and partial occlusion [2, 3]. These characteristics make ResNet well suited for biometric authentication and access control applications that require reliable identification across diverse operating environments. Due to its effectiveness and scalability, ResNet remains one of the most widely used deep learning architectures in modern facial recognition research and serves as the foundation for many state-of-the-art face recognition frameworks [4, 5]. Figure 1 illustrates the general architecture of ResNet and the residual block mechanism.

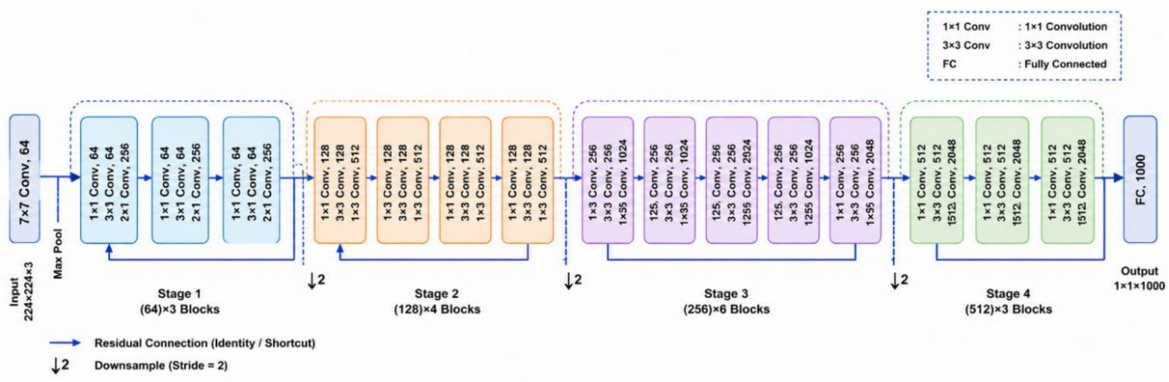

Figure 1. ResNet architecture.

2.3. Research dataset.

The face dataset used in this study was obtained from the Roboflow Universe platform under the Face Recognition category. The dataset consists of facial images labeled with individual identities (IDs), enabling each image to be associated with a specific person for recognition purposes. This dataset was used to train, validate, and evaluate the ResNet-based facial recognition model developed for the entrance security system. As shown in Figure 2, the dataset contains facial images with variations in pose, facial expression, illumination, and partial occlusion. Such variations introduce realistic identification challenges, including changes in viewing angle, lighting conditions, facial accessories (e.g., glasses and masks), and facial expressions. These characteristics are important for improving the model's ability to generalize and perform reliably in real-world environments.

Figure 2. Sample facial images from the face dataset used for training, validation, and testing of the ResNet-based facial recognition model.

All images were resized to 224 × 224 pixels to satisfy the input requirements of the ResNet architecture. The dataset was divided into training (70%), validation (20%), and testing (10%) subsets. The selected dataset provides sufficient diversity in facial appearance and environmental conditions, making it suitable for evaluating facial recognition performance under practical operating scenarios [11,n12]. Because facial data are classified as sensitive biometric information, ethical and legal considerations were taken into account during dataset selection. The dataset was used solely for research purposes and was obtained from a publicly accessible source with distribution permission provided by the dataset owner. Nevertheless, potential limitations such as identity bias, unequal data representation, and the risk of misidentification remain important considerations when utilizing public facial datasets. Therefore, dataset selection was based on data quality, legal accessibility, ethical compliance, and suitability for the objectives of this study.

2.4. System design.

The proposed facial recognition system consists of four main stages: dataset preparation, image preprocessing, feature extraction, and performance evaluation. Facial images obtained from a publicly available dataset were divided into training, validation, and testing subsets with proportions of 70%, 20%, and 10%, respectively. Prior to model training, all images underwent preprocessing, including face detection and cropping, resizing to 224 × 224 pixels, and pixel value normalization to ensure compatibility with the ResNet input requirements [11, 12]. The preprocessed images were then used to train a ResNet-based facial recognition model. During training, the validation dataset was employed to monitor model performance and reduce the risk of overfitting. Once training was completed, the testing dataset was used to evaluate the model's recognition capability. The extracted facial features were utilized for identity matching and access verification within the proposed entrance control system. Figure 3 illustrates the overall workflow of the proposed system. The workflow begins with facial image acquisition, followed by face detection, preprocessing, feature extraction using the ResNet model, identity matching, and access decision generation. This design simulates a practical facial recognition-based access control scenario and enables the evaluation of model performance under variations in facial pose, expression, illumination, and partial occlusion [13, 14].

2.5. System implementation and testing.

The proposed system was implemented using Python 3.x. OpenCV was employed for image processing and face detection, while TensorFlow/Keras was used to develop and train the ResNet model [16,17]. Experiments were conducted on a computer equipped with an Intel Core processor, 16 GB RAM, and an NVIDIA GPU when available. ResNet-50 was selected as the backbone architecture for facial feature extraction. Model training was performed using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy as the loss function. The model was trained for 500 epochs with a batch size of 32. Training and validation performance were monitored using accuracy and loss metrics. For facial recognition, facial feature vectors extracted by the ResNet model were compared with stored feature representations using a similarity-based matching approach. Identity verification was performed by selecting the highest similarity score that exceeded a predefined threshold. Similarity measurements were computed using distance metrics such as Euclidean distance or cosine similarity. The performance of the proposed system was evaluated using accuracy, precision, recall, and mean Average Precision (mAP). These metrics were used to assess the effectiveness of the model in correctly identifying facial identities and its robustness under different facial and environmental conditions [18, 19].

3. Results and Discussion

3.1. Comparison Experiments: Training Performance of the ResNet Model

This study evaluated the performance of the ResNet-based facial recognition model using the facial image dataset described in Section 2. The training process was conducted for 500 epochs using the preprocessing procedures outlined previously, including face cropping, image resizing to 224 × 224 pixels, and pixel normalization.

3.1.1. Results for initial training phase.

The initial training results indicate that the ResNet model achieved relatively low classification performance throughout the training process. As shown in Figure 4, both training and validation accuracy remained nearly constant during most epochs, fluctuating between approximately 3.8% and 4.7% for training accuracy and between 3.6% and 4.3% for validation accuracy. Similarly, SM Table 1 shows that although the training loss gradually decreased from 8.4549 at Epoch 1 to approximately 6.3448 at Epoch 500, the corresponding improvement in accuracy was minimal. The learning curves presented in Figure 4 reveal that the model failed to achieve substantial performance gains despite prolonged training. Training accuracy increased only slightly from 3.80% to 4.62%, while validation accuracy remained around 3.80% at the end of training. At the same time, validation loss exhibited considerable fluctuations and generally remained high throughout the experiment. This pattern suggests that the model was unable to learn sufficiently discriminative facial representations from the available training data.

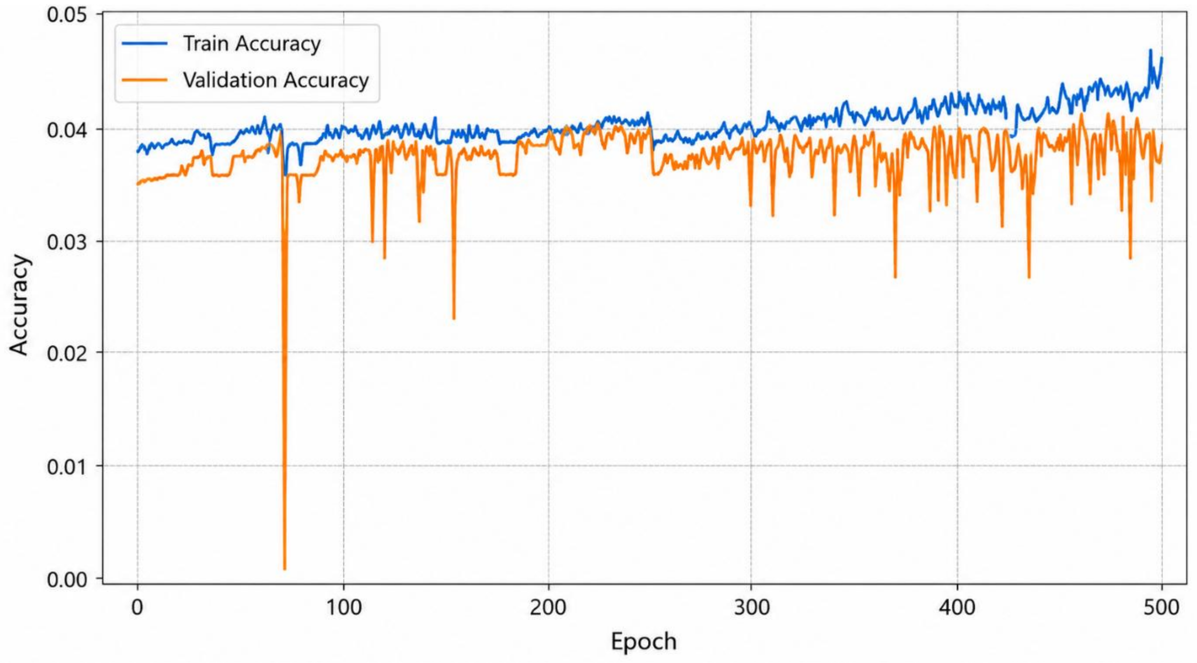

Figure 4. Training and validation accuracy and loss curves of the ResNet facial recognition model over 500 training epochs.

The detailed results in SM Table 1 further confirm this observation. Although a gradual reduction in training loss can be observed across epochs, the corresponding accuracy values remained close to random classification performance. This indicates that the optimization process was able to reduce the loss function to some extent but did not successfully improve the model's ability to distinguish between different facial identities. Such behavior is commonly associated with underfitting, where the model fails to capture the underlying patterns present in the dataset. Several factors may have contributed to this outcome. First, the dataset contains significant variations in facial pose, illumination, facial expression, and partial occlusion, which increase the complexity of the recognition task. Examples shown in Figure 2 illustrate the diversity of facial appearances and environmental conditions that the model must learn to handle. These variations may have hindered the extraction of consistent facial features, particularly when the number of samples available for each identity was limited. Second, although standard preprocessing techniques were applied, more advanced methods such as facial alignment, illumination normalization, contrast enhancement, and extensive data augmentation were not implemented. Previous studies have demonstrated that such preprocessing techniques can substantially improve facial recognition performance by reducing intra-class variability and increasing the robustness of learned features [15–17].

Another possible explanation relates to the model configuration. While ResNet is capable of extracting highly discriminative features, its effectiveness depends on appropriate hyperparameter settings, sufficient training data, and optimized learning strategies. The learning rate, batch size, and training configuration used in this study may not have been optimal for the characteristics of the dataset. Furthermore, the large number of facial identities relative to the available samples per class may have increased classification difficulty and limited the model's generalization capability. The architecture of ResNet, illustrated in Figure 1, relies on residual learning and deep feature extraction to identify complex facial patterns. However, the experimental results suggest that the available dataset and training configuration were insufficient to fully exploit the advantages of the architecture. Consequently, the residual blocks were unable to learn sufficiently discriminative representations to achieve reliable facial classification performance. Overall, the results indicate that the initial training configuration did not produce a satisfactory facial recognition model. Although a gradual reduction in training loss was observed, the consistently low training and validation accuracies suggest that the model suffered from underfitting and failed to learn meaningful facial representations. Future improvements should focus on increasing dataset diversity, balancing the number of samples per identity, implementing advanced preprocessing and data augmentation techniques, and optimizing training hyperparameters to enhance recognition performance..

3.1.2. Results of the Extended Training Phase

To determine whether a longer training duration could improve facial recognition performance, the ResNet model was trained for an extended number of epochs beyond the initial training configuration. The results of this experiment are summarized in Table 1 and compared with the learning behavior observed in the initial training phase. As shown in Table 1, extending the training duration produced only a marginal improvement in model performance. Training accuracy increased slightly, reaching values above 4% in some epochs, while validation accuracy remained approximately 3–4% throughout the training process. Although the training loss gradually decreased, both training and validation loss values remained relatively high, indicating that the model had not achieved satisfactory convergence.

Table 1. Summary of model performance during the extended training phase.

Metric	Initial Training
Epochs	500
Training Accuracy	3.6–4.1%
Validation Accuracy	~3.6%
Training Loss	High
Validation Loss	High
Convergence Status	Not Converged
Learning Behavior	Underfitting

The limited improvement suggests that the primary performance bottleneck was not the number of training epochs. If insufficient training time were the main issue, a more substantial increase in both training and validation accuracy would be expected as learning progressed. Instead, the consistently low accuracy values indicate that the model was unable to learn sufficiently discriminative facial representations from the available dataset. This observation is consistent with the learning curves presented in Figure 4, where both training and validation accuracy remained nearly flat despite prolonged training. Several factors may explain this behavior. First, the facial dataset contains considerable variations in pose, illumination, facial expression, and partial occlusion, which increase classification complexity. While such variability is desirable for evaluating model robustness, it also requires a sufficiently large and balanced dataset to enable effective feature learning. When the number of samples per identity is limited, additional training epochs may provide little benefit because the model repeatedly encounters the same information during optimization [20, 21]. Second, class imbalance may have affected the learning process. Unequal representation among facial identities can bias the model toward majority classes while reducing its ability to recognize minority classes. Consequently, the learned feature space may not adequately separate individual identities, resulting in poor classification performance and limited generalization capability [22]. Third, the training configuration may not have been optimal for the dataset characteristics. Hyperparameters such as learning rate, batch size, optimizer settings, and regularization strength directly influence convergence behavior. Suboptimal hyperparameter selection can prevent the model from effectively minimizing classification errors even when the training duration is increased [23].

Furthermore, the persistent gap between the expected performance of the ResNet architecture and the observed results suggests that the network was operating in an underfitting regime. The residual learning mechanism illustrated in Figure 1 is designed to facilitate deep feature extraction and efficient gradient propagation. However, the low training accuracy observed throughout the experiment indicates that the model failed to capture meaningful relationships within the dataset. This conclusion is further supported by the high loss values reported in Table 1, which demonstrate that the optimization process did not successfully learn representative facial features. Overall, the results indicate that extending the training duration alone was insufficient to overcome the limitations imposed by the dataset and training configuration. Although a slight reduction in loss was observed, the model continued to exhibit underfitting characteristics, as evidenced by low training accuracy, low validation accuracy, and poor convergence behavior. Therefore, future improvements should focus on increasing dataset diversity, balancing class distributions, applying advanced preprocessing and data augmentation techniques, optimizing hyperparameters, and incorporating transfer learning strategies rather than solely increasing the number of training epochs [20, 23].

3.1.3. Discussion of model performance.

The experimental results demonstrated that the ResNet-based facial recognition model achieved limited recognition performance under the current training configuration. As summarized in Table 2 and illustrated in Figure 4, both training and validation accuracies remained below 5% throughout the training process, while loss values remained relatively high. Although a gradual decrease in training loss was observed, the corresponding improvement in accuracy was minimal, indicating that the model failed to learn sufficiently discriminative facial representations. The observed learning behavior suggests that the model suffered from underfitting. Despite the residual learning mechanism of ResNet, shown in Figure 1, the network was unable to effectively capture the complex facial patterns present in the dataset. This conclusion is supported by the nearly flat accuracy curves and the absence of stable convergence during training.

Several factors may have contributed to the low performance. First, the dataset size and class distribution may have been insufficient for training a deep neural network. Facial recognition models generally require large and balanced datasets containing substantial variations in pose, illumination, facial expression, and occlusion to achieve robust feature learning [25,26]. Limited samples per identity can restrict the model's ability to learn representative facial characteristics and generalize to unseen data. Second, the preprocessing pipeline consisted primarily of face cropping, resizing, and normalization. While these procedures standardize image inputs, more advanced techniques such as facial alignment, illumination correction, and data augmentation were not implemented. Previous studies have shown that these techniques can significantly improve feature consistency and recognition performance by reducing intra-class variability [27]. Third, the selected hyperparameter configuration may not have been optimal for the dataset characteristics. Learning rate, batch size, optimizer settings, and training duration strongly influence convergence behavior and model performance. Inadequate hyperparameter tuning can limit the effectiveness of feature learning even when a powerful architecture such as ResNet is employed [28].

Compared with previous studies, the performance obtained in this work was substantially lower. For example, FaceNet and ArcFace have reported recognition accuracies exceeding 99% on benchmark datasets using large-scale training data and optimized deep-learning frameworks [29,30]. Similarly, ResNet-based facial recognition systems have achieved high recognition rates when trained using balanced datasets and comprehensive preprocessing strategies [31]. The discrepancy between these results and those obtained in the present study highlights the importance of dataset quality, training configuration, and computational resources in determining facial recognition performance. Nevertheless, the findings remain valuable because they demonstrate the practical challenges associated with deploying facial recognition systems in resource-constrained environments. The results show that the effectiveness of deep learning models is highly dependent on the availability of representative training data and appropriate optimization strategies, even when advanced architectures such as ResNet are utilized.

3.1.4. Overall evaluation.

Based on the experimental results, the proposed ResNet-based facial recognition system has not yet achieved the level of performance required for practical deployment in automated entrance control applications. The consistently low accuracy values and high loss values indicate that the model was unable to reliably distinguish between different facial identities under varying environmental and facial conditions. From an operational perspective, such performance may result in frequent misidentification, false acceptance, and false rejection events, which could compromise both security and user convenience. Therefore, the current implementation should be regarded as a proof-of-concept rather than a deployable security solution. Several limitations were identified during the study. These include the limited size and diversity of the facial dataset, the absence of advanced preprocessing and augmentation techniques, and the lack of systematic hyperparameter optimization. Collectively, these factors restricted the model's ability to learn robust facial representations and achieve stable convergence. To improve performance, future research should focus on expanding dataset diversity, balancing class distributions, and implementing advanced preprocessing techniques such as facial alignment and illumination normalization. In addition, data augmentation methods, including rotation, scaling, brightness adjustment, and horizontal flipping, may improve model robustness by increasing the variability of training samples [34]. Another promising strategy is the adoption of transfer learning using pretrained ResNet models. Transfer learning enables the model to leverage feature representations learned from large-scale image datasets, thereby improving learning efficiency and recognition performance when only limited facial data are available [35, 36]. Furthermore, systematic optimization of training parameters, including learning rate, batch size, optimizer selection, and regularization settings, may enhance convergence and classification accuracy. Overall, the study demonstrates both the potential and the limitations of applying ResNet-based facial recognition to entrance control systems. While the current results are not yet suitable for real-world deployment, the findings provide a useful foundation for future improvements aimed at developing a more accurate, reliable, and robust biometric access control system.

4. Conclusions

This study evaluated the feasibility of implementing a ResNet-based facial recognition model for an automated door access control system. The findings revealed that the model achieved very low recognition performance, characterized by low training and validation accuracy and persistently high loss values, indicating an underfitting condition and an inability to learn sufficiently discriminative facial features from the available dataset. The results suggest that increasing the number of training epochs alone was insufficient to improve model performance when fundamental limitations existed in the dataset and training pipeline. These limitations included inadequate dataset size and diversity, limited preprocessing and data augmentation procedures, and suboptimal training configurations. Consequently, the current model cannot be considered suitable for practical deployment in real-world access control applications, where reliable and accurate identification is essential. Nevertheless, this study provides valuable insights into the challenges associated with developing facial recognition systems under constrained data conditions and highlights critical factors affecting model effectiveness. Future research should focus on improving dataset quality and diversity, incorporating advanced preprocessing and augmentation techniques, leveraging transfer learning from pretrained deep learning models, and systematically optimizing training parameters to enhance recognition accuracy, robustness, and deployment readiness in automated security systems.

Acknowledgments

Not applicable

Author Contribution

All authors contributed equally to the conception and design of the study, material preparation, data collection, and data analysis.

Data Availability

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

Competing Interest

The authors declare no conflict of interest.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.53623/gisa.v6i1.1191.

References

[1] He, K.; Zhang, X.; Ren, S.; Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/CVPR.2016.90.

[2] Diba, M.; Khosravi, H. (2024). SNResNet: A New Architecture Based on SqNxt Blocks and Rish Activation for Efficient Face Recognition. *Traitement du Signal*, *41*, 235–244. https://doi.org/10.18280/ts.410235.

[3] Deng, N.; Lin, J. (2025). Residential Access Control Facial Recognition Based on Improved ResNet Computing Model. *Traitement du Signal*, *42*, 295–304. https://doi.org/10.18280/ts.420130.

[4] Satapathy, A.; Livingston, J.L.M. (2023). A Lightweight Convolutional Neural Network Built on Inception-Residual and Reduction Modules for Deep Facial Recognition in Realistic Conditions. *Connection Science*, *35*, 1–21. https://doi.org/10.1080/13682199.2023.2176735.

[5] Xiao, L.; Li, X.; Luo, Y.; et al. (2024). Research on the Application of CNN Face Recognition Technology in the Airport. *Advances in Transportation Studies and Development Engineering*, *1*, 1–8. https://doi.org/10.3233/ATDE240470.

[6] Obaid, A.M.; et al. (2023). Exploring the Potential of A-ResNet in Person-Independent Face Recognition and Classification. *International Journal of Advances in Soft Computing and Its Applications*, *15*, 1–15. https://doi.org/10.2478/ijanmc-2023-0052.

[7] Zhang, M.; Zhang, Y.; Zhang, Q. (2023). Attention-Mechanism-Based Models for Unconstrained Face Recognition with Mask Occlusion. *Electronics*, *12*, 3916. https://doi.org/10.3390/electronics12183916.

[8] Al-Mashhadani; et al. (2024). Thermal Image Identification Against Pose and Expression Variations Using Deep Learning. *Journal of Engineering Research*, *12*, 1–13. https://doi.org/10.1016/j.jer.2023.10.043.

[9] Sharma, R.; et al. (2024). Low-Resolution Face Recognition: Review, Challenges and Research Directions. *Computers and Electrical Engineering*, *118*, 109846. https://doi.org/10.1016/j.compeleceng.2024.109846.

[10] Li, Y.; et al. (2024). Occluded Face Recognition Network Based on DCGAN and ResNet. *Procedia Computer Science*, *246*, 1203–1212. https://doi.org/10.1016/j.procs.2024.09.087.

[11] Viola, P.; Jones, M. (2004). Robust Real-Time Face Detection. In *International Journal of Computer Vision*, *57*, 137–154. https://doi.org/10.1007/978-1-4757-4032-3_24.

[12] Shorten, C.; Khoshgoftaar, T.M. (2019). A Survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*, 60. https://doi.org/10.1186/s40537-019-0197-0.

[13] Sokolova, M.; Lapalme, G. (2009). A Systematic Analysis of Performance Measures for Classification Tasks. *Information Processing and Management*, *45*, 427–437. https://doi.org/10.1016/j.ipm.2009.03.002.

[14] Schroff, F.; Kalenichenko, D.; Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823. https://doi.org/10.1109/CVPR.2015.7298682.

[15] Goodfellow, I.; Bengio, Y.; Courville, A. (2016). Deep Learning; MIT Press: Cambridge, MA, USA.

[16] Johnson, J.M.; Khoshgoftaar, T.M. (2019). Survey on Deep Learning with Class Imbalance. *Journal of Big Data*, *6*, 27. https://doi.org/10.1186/s40537-019-0192-5.

[17] Bengio, Y. (2012). Practical Recommendations for Gradient-Based Training of Deep Architectures. In *Neural Networks: Tricks of the Trade*, 2nd ed.; Springer: Berlin, Germany; pp. 437–478. https://doi.org/10.1007/978-3-642-35289-8_26.

[18] Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.

[19] Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. (2019). ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4690–4699. https://doi.org/10.1109/CVPR.2019.00482.

[20] Pan, S.J.; Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*, 1345–1359. https://doi.org/10.1109/TKDE.2009.191.

[21] Buda, M.; Maki, A.; Mazurowski, M.A. (2018). A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks. *Neural Networks, 106,* 249–259. https://doi.org/10.1016/j.neunet.2018.07.011.

[22] Krawczyk, B. (2016). Learning from Imbalanced Data: Open Challenges and Future Directions. *Progress in Artificial Intelligence*, *5,* 221–232. https://doi.org/10.1007/s13748-016-0094-0.

[23] Smith, L.N. (2018). A Disciplined Approach to Neural Network Hyper-Parameters: Part 1—Learning Rate, Batch Size, Momentum, and Weight Decay. *arXiv* Preprint*, arXiv:1803.09820.*

[24] Parkhi, O.M.; Vedaldi, A.; Zisserman, A. (2015). Deep Face Recognition. *Proceedings of the British Machine Vision Conference (BMVC)*, *41.*1–41.12. https://doi.org/10.5244/C.29.41.

[25] Taigman, Y.; Yang, M.; Ranzato, M.; Wolf, L. (2014). DeepFace: Closing the Gap to Human-Level Performance in Face Verification. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1701–1708. https://doi.org/10.1109/CVPR.2014.220.

[26] Sun, Y.; Wang, X.; Tang, X. (2014). Deep Learning Face Representation from Predicting 10,000 Classes. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1891–1898. https://doi.org/10.1109/CVPR.2014.244.

[27] Shorten, C.; Khoshgoftaar, T.M. (2019). A Survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6,* 60. https://doi.org/10.1186/s40537-019-0197-0.

[28] Perez, L.; Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification Using Deep Learning. *arXiv Preprint, arXiv:1712.04621*

[29] Schroff, F.; Kalenichenko, D.; Philbin, J. (2015). FaceNet: A Unified Embedding for Face Recognition and Clustering. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815–823. https://doi.org/10.1109/CVPR.2015.7298682.

[30] Guo, Y.; Zhang, L.; Hu, Y.; He, X.; Gao, J. (2016). MS-Celeb-1M: A Dataset and Benchmark for Large-Scale Face Recognition. *Computer Vision – ECCV 2016 Workshops*, 87–102. https://doi.org/10.1007/978-3-319-46604-0_6.

[31] Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. (2019). ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 4690–4699. https://doi.org/10.1109/CVPR.2019.00482.

[32] Buda, M.; Maki, A.; Mazurowski, M.A. (2018). A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks. *Neural Networks*, *106,* 249–259. https://doi.org/10.1016/j.neunet.2018.07.011.

[33] Deng, J.; Guo, J.; Niannan, X.; Zafeiriou, S. (2019). ArcFace: Additive Angular Margin Loss for Deep Face Recognition. *CVPR 2019*, 4690–4699. https://doi.org/10.1109/CVPR.2019.00482.

[34] Cubuk, E.D.; Zoph, B.; Mane, D.; Vasudevan, V.; Le, Q.V. (2019). AutoAugment: Learning Augmentation Policies from Data. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 113–123. https://doi.org/10.1109/CVPR.2019.00020.

[35] Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. (2018). A Survey on Deep Transfer Learning. *Lecture Notes in Computer Science*, *11141,* 270–279. https://doi.org/10.1007/978-3-030-01424-7_27.

[36] Weiss, K.; Khoshgoftaar, T.M.; Wang, D. (2016). A Survey of Transfer Learning. *Journal of Big Data*, *3*, 9. https://doi.org/10.1186/s40537-016-0043-6.