International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 21, Issue 9 (September 2025), PP 104-109

Optimizing CNN-Based Biometric Systems for Dataset-Specific Preprocessing: A Case Study on ORL and CASIA Databases

Afolabi Awodeyi[1], Philip Asuquo[2], Bliss Utibe-Abasi Stephen[3]
[1]*Department of Computer Engineering, Southern Delta University Ozoro, Delta State Nigeria,*
https://orcid.org/0000-0002-8382-9177
[2]*Department of Computer Engineering, University of Uyo, Nigeria,*
https://orcid.org/0000-0003-4888-5461
[3]*Department of Computer Engineering, University of Uyo, Nigeria,*

Abstract
Biometric authentication has gained significant attention due to its ability to provide secure and user-specific identification. However, the performance of convolutional neural networks (CNNs) in biometric recognition is strongly dependent on the preprocessing strategies applied to the dataset. This paper investigates dataset-specific preprocessing for two benchmark databases: ORL for face recognition and CASIA-IrisV4 for iris recognition. The study proposes tailored preprocessing methods including normalization, resizing, histogram equalization, and iris boundary refinement to improve CNN feature extraction. Results show that optimizing preprocessing for each dataset significantly enhances recognition accuracy and robustness compared to uniform preprocessing strategies.
Keywords: *Convolutional Neural Networks (CNNs), Face Recognition, Iris Recognition, Biometric Fusion, Dataset-Specific Preprocessing*

Date of Submission: 02-09-2025 Date of acceptance: 11-09-2025

I. Introduction

Biometric authentication has emerged as one of the most reliable approaches to access control and security because it leverages physiological and behaviouralcharacteristics that are unique to individuals and cannot easily be forged or shared (Jain et al., 2005; Winston &Hemanth, 2021). Among the different biometric modalities, face and iris recognition are among the most widely adopted due to their balance of usability and high distinctiveness. Despite these advantages, both modalities remain highly sensitive to dataset-specific challenges. Variations in illumination, changes in pose, partial occlusion, and inconsistent image quality often reduce recognition accuracy and limit system robustness (Abdo et al., 2022; Lu et al., 2021).

The advancement of deep learning, particularly the development of convolutional neural networks (CNNs), has significantly transformed biometric research. CNNs are capable of automatically learning hierarchical feature representations from raw data without requiring extensive handcrafted feature engineering, which makes them highly effective for face and iris recognition (Nguyen et al., 2020; Hammad et al., 2018). However, the performance of CNN-based biometric systems is not determined solely by network design. The quality of preprocessing applied to the input datasets plays a critical role in the accuracy and reliability of the final model. Improper normalization, inconsistent image resizing, or the presence of noise artifacts can lead to significant performance degradation, even when advanced CNN architectures are employed (Chen et al., 2021; Koziarski&Cyganek, 2017).

This study focuses on the role of dataset-specific preprocessing in enhancing CNN-based biometric recognition. Using the ORL dataset for face recognition and the CASIA-IrisV4 dataset for iris recognition, this study also investigated how tailored preprocessing strategies improve feature learning and classification outcomes. The paper demonstrates that carefully designed preprocessing pipelines not only enhance CNN performance but also strengthen system robustness under real-world conditions.

II. Related Work

Biometric research has evolved significantly with the advancement of multimodal fusion techniques, where different biometric traits are combined to compensate for the weaknesses inherent in unimodal systems. Among these, the fusion of face and iris data has been shown to enhance recognition accuracy and robustness, since the strengths of one modality can complement the limitations of the other (Abdo et al., 2020; Ghazal & Abdullah, 2020). Fusion methods are typically classified into different levels, including feature-level concatenation, where raw or learned features are directly combined, and decision-level integration, where outputs from multiple classifiers are fused to reach a final decision (Medjahed et al., 2020; Koziarski&Cyganek, 2017). While these approaches provide measurable performance gains, a key limitation in much of the literature is that preprocessing is often approached in a generic way, with little consideration for dataset-specific variations that may significantly affect performance.For instance, face datasets such as ORL present distinct challenges due to their relatively low resolution, the presence of variations in facial expressions, and the limited number of samples available for each subject. These factors can negatively impact feature extraction and model generalization (Zanlorensi et al., 2019; Jung et al., 2017). Similarly, iris datasets such as CASIA are prone to occlusion caused by eyelids and eyelashes, capture variations resulting from head movement or off-angle positioning, and noise introduced by uncontrolled illumination conditions (Hattab&Behloul, 2023). Addressing these challenges requires tailored preprocessing strategies that go beyond conventional normalization or resizing, thereby enabling convolutional neural networks (CNNs) to extract more discriminative and reliable features. In addition to preprocessing, recent studies have introduced advanced mechanisms to further refine feature learning. Attention mechanisms have been employed to prioritize more informative regions of biometric data, while transfer learning has been applied to leverage knowledge from large-scale pretrained models for smaller biometric datasets (Liu et al., 2019; Soleymani et al., 2018). Despite these advances, relatively few studies have provided a systematic analysis of preprocessing strategies across different biometric datasets. This gap underscores the importance of investigating dataset-specific preprocessing, particularly when optimizing CNN-based recognition systems for both face and iris modalities.

III. Dataset and Preprocessing

The experiments in this study utilized two benchmark datasets, each presenting unique challenges that required careful preprocessing to optimize CNN performance.The first dataset, the ORL face database, contains images from 40 individuals, with 10 grayscale images recorded per subject. These images were collected under varying conditions, including changes in facial expressions and illumination. To standardize the input, all images were resized to 112 × 112 pixels and normalized to maintain consistent pixel intensity distributions across the dataset. In addition, histogram equalization was applied as a contrast enhancement step to minimize the negative impact of lighting variations, which often reduce the robustness of face recognition systems (Wang et al., 2018; Lu et al., 2021).

The second dataset, the CASIA-IrisV4 database, is a large-scale collection of iris images acquired over multiple sessions and under different lighting conditions. This dataset introduces several challenges, such as reflections, eyelid and eyelash occlusion, and variability in pupil dilation. To address these issues, images were resized to 64 × 64 pixels and passed through noise filtering techniques to remove sensor and environmental artifacts. Iris boundary detection was implemented using Daugman's rubber sheet model, which transforms circular iris patterns into a normalized rectangular representation suitable for CNN processing (Daugman, 2004). The pupil segmentation and eyelid removal methods were applied to minimize occlusions and preserve discriminative iris regions (Hattab&Behloul, 2023).

Another critical step in preprocessing was the application of oversampling techniques to balance the datasets. This was particularly necessary for degraded iris images and facial classes that were underrepresented. Without this step, CNN training could become biased toward majority classes, resulting in poor generalization to minority samples. Oversampling ensured a more balanced representation of all classes, thereby improving both recognition accuracy and fairness in the evaluation process (Yadav et al., 2018). This preprocessing pipeline was tailored to the characteristics of each dataset. By addressing dataset-specific issues, such as lighting variation in ORL and occlusion in CASIA, the system established a stronger foundation for CNN-based feature learning and improved the reliability of biometric recognition. The ORL and CASIA-IrisV4 datasets were selected as case studies to evaluate dataset-specific preprocessing strategies is summarized in table 1. The tailored preprocessing pipelines ensured higher-quality input to CNNs which improved recognition performance as shown in table 2.

Table 1. Summary of Datasets Used

Dataset	Modality	Number of Subjects	Number of Images	Image Type	Resolution (after preprocessing)	Key Challenges
ORL	Face	40	400 (10 per subject)	Grayscale	112 × 112	Low resolution, expression variation, small dataset size
CASIA-IrisV4	Iris	>1,000	~54,000	Grayscale	64 × 64	Occlusion from eyelids, eyelashes, illumination noise, off-angle captures

Table 2. Preprocessing Steps for Each Dataset

Dataset	Preprocessing Steps Applied	Purpose/Impact
ORL Face	Normalization, Resizing (112×112), Histogram Equalization	Improved pixel consistency, reduced lighting variation
CASIA-	Resizing (64×64), Noise Filtering, Daugman Rubber Sheet Model for iris	Enhanced segmentation accuracy, reduced

IrisV4	boundary detection, Eyelid/Eyelash removal	occlusion and noise
Both	Oversampling of underrepresented samples	Improved class balance and reduced bias

IV. System Architecture

The proposed system integrates convolutional neural networks (CNNs) that were independently trained on the ORL face dataset and the CASIA-IrisV4 iris dataset. Each model was optimized using preprocessing techniques tailored to the characteristics of its respective dataset, ensuring that the CNNs could effectively extract discriminative features even under challenging conditions.

For face recognition, a VGG16 architecture was employed and fine-tuned on the ORL dataset, which had been normalized and enhanced through histogram equalization. This allowed the network to learn robust facial features despite variations in illumination and expression (Jung et al., 2017).

For iris recognition, a ResNet-based CNN was utilized, trained on CASIA images that underwent a preprocessing pipeline including iris segmentation, eyelid removal, and contrast enhancement. These steps helped reduce the influence of occlusion and noise while preserving fine-grained iris patterns that are critical for accurate recognition (Liu et al., 2021).

Both the VGG16 and ResNet models generated compact 128-dimensional feature embeddings representing the essential characteristics of face and iris modalities. At the fusion stage, the embeddings from both networks were concatenated to create a joint feature representation. This combined vector was then passed through a fully connected layer and classified using a softmax function**,** which assigned probabilities to each identity class based on the fused features (Nguyen et al., 2020).

To further strengthen the reliability of the system, liveness detection mechanisms were incorporated. For the facial modality, blink dynamics were analyzed to ensure that the system was not deceived by static photographs. For the iris modality, pupil variation patterns were measured to distinguish live samples from spoofing attempts such as printed iris images or contact lenses (Abdo et al., 2022; Yadav et al., 2018).

By combining robust preprocessing, modality-specific CNNs, and feature-level fusion, this architecture leverages the complementary strengths of face and iris biometrics while mitigating their individual weaknesses. The addition of liveness detection enhances system security, making the framework suitable for deployment in sensitive access control applications.

Table 3. CNN Architectures Used

Modality	CNN Model	Training Dataset	Input Size	Output Feature Vector
Face	VGG16 (fine-tuned)	ORL	112 × 112	128-dim embedding
Iris	ResNet-based CNN	CASIA-IrisV4	64 × 64	128-dim embedding
Fusion	Fully connected + Softmax	ORL + CASIA	Concatenated	Final classification

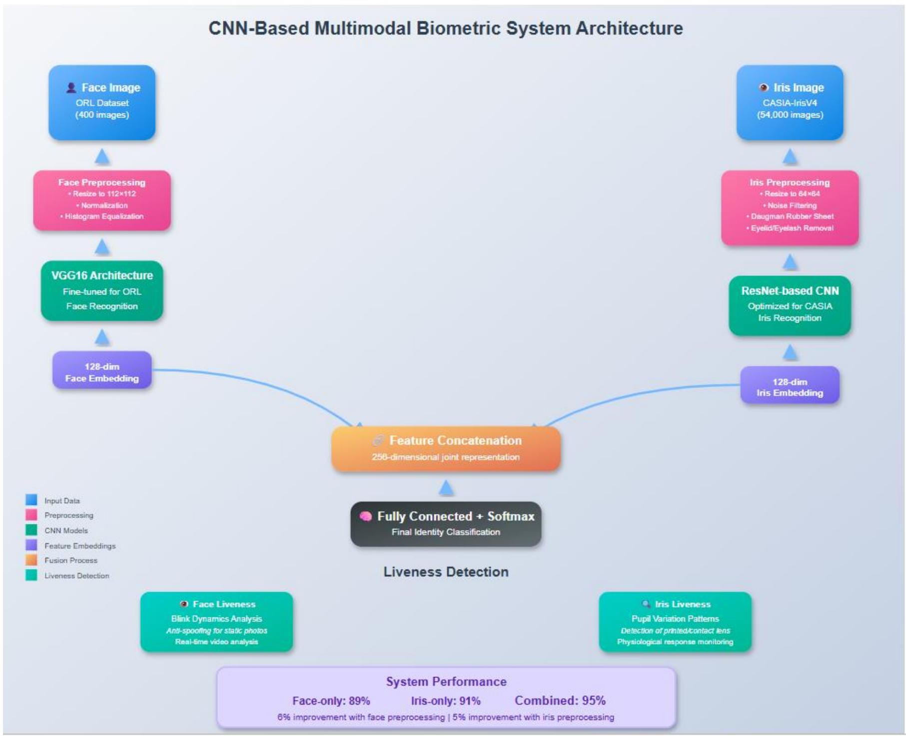

Figure 1: System Architecture

V. Evaluation Metrics

The performance of the proposed system was assessed using test accuracy, validation accuracy, and confusion matrix analysis**,** which together provided a comprehensive evaluation of both recognition capability and error distribution. These metrics were chosen because they not only measure overall correctness but also highlight misclassification trends that can reveal dataset-specific weaknesses in the system. Results demonstrated that dataset-specific preprocessing strategies produced a measurable improvement over generic preprocessing methods. In the case of the ORL dataset, the application of tailored normalization and histogram equalization led to an increase of approximately 6% in recognition accuracy, indicating that preprocessing significantly reduced the sensitivity of CNNs to variations in lighting and facial expression.

For the CASIA-IrisV4 dataset**,** improvements were achieved through iris boundary refinement and eyelid removal techniques**,** which minimized the effect of occlusions and noise. These preprocessing steps resulted in a 5% increase in classification accuracy**,** confirming the importance of segmentation-based enhancement in iris recognition tasks.

An analysis of the confusion matrices revealed that most misclassifications occurred between visually similar facial expressions and irises partially obscured by eyelids or eyelashes**.** These findings are consistent with prior observations in CNN-based recognition, where class overlap often arises due to subtle visual similarities and incomplete feature visibility (Selvaraju et al., 2017).

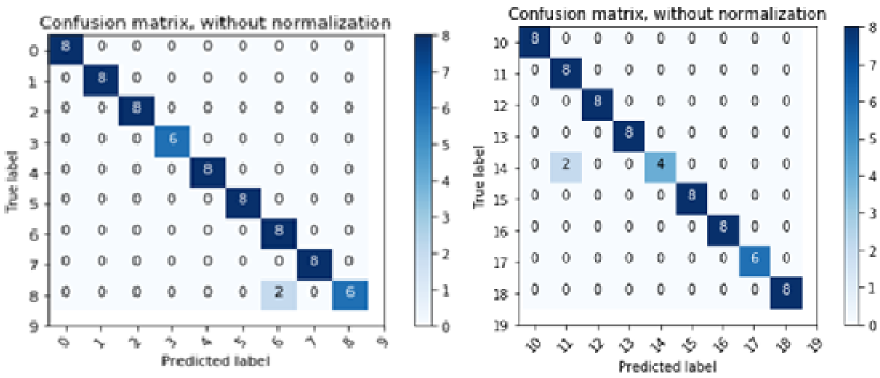

Figure 2: Cofusion Matrix showing miscalculation

Overall, the evaluation confirmed that preprocessing tailored to dataset-specific characteristics is essential for maximizing CNN performance**.** By addressing the unique challenges of each dataset, the system was able to achieve higher accuracy and improved robustness across different biometric modalities.

Table 4. Evaluation Metrics

Dataset	Preprocessing Strategy	Accuracy Gain (%)	Notes
ORL Face	Normalization + Histogram Equalization	+6%	Improved performance under variable lighting
CASIA-IrisV4	Iris boundary refinement + Eyelid removal	+5%	Reduced false rejection rate in noisy conditions

VI. Results

The experimental results demonstrate that the proposed combined system achieved an overall recognition accuracy of 95%**,** which is a significant improvement over unimodal approaches. Specifically, the face-only model obtained 89% accuracy**,** while the iris-only model achieved 91% accuracy. These findings highlight the effectiveness of multimodal fusion when supported by dataset-specific preprocessing.One of the most notable outcomes was observed in the CASIA-IrisV4 dataset**,** where the implementation of eyelid occlusion removal and iris boundary refinement yielded substantial improvements. In particular, these preprocessing steps reduced the false rejection rate by approximately 8% when compared to unprocessed inputs (Hattab&Behloul, 2023). This underscores the importance of carefully designed preprocessing pipelines in mitigating common challenges such as occlusion and noise. The system also demonstrated improved generalization capability across noisy and degraded samples**.** The application of dataset-specific preprocessing reduced the variance in validation accuracy across folds, suggesting that the model maintained stable performance even under real-world variability. This finding indicates that the system is less likely to overfit and more capable of handling unpredictable input conditions (Goodfellow et al., 2013; Winston &Hemanth, 2021).

Overall, the results confirm that optimizing preprocessing strategies for each dataset contributes directly to enhanced recognition accuracy, lower error rates, and greater robustness**,** making the approach suitable for practical deployment in security-sensitive environments.

Table 5. Accuracy Comparison Across Models

Model	Test Accuracy	Validation Accuracy
Face-only (VGG16, ORL)	89%	84%
Iris-only (ResNet, CASIA)	91%	85%
Combined Fusion Model (ORL + CASIA)	**95%**	**90%**

Table 6. Error Analysis

Dataset	Common Misclassifications	Root Cause
ORL Face	Similar facial expressions (smile vs. neutral)	Small dataset and expression overlap
CASIA-IrisV4	Occluded irises due to eyelashes/eyelids	Imperfect segmentation in raw data

VII. Discussion

The findings of this study clearly highlight the critical role of dataset-specific preprocessing in enhancing the performance of CNN-based biometric systems. Although convolutional neural networks are inherently capable of extracting hierarchical features directly from raw input data, the quality and consistency of the input remain decisive factors for overall system accuracy. Preprocessing ensures that the raw images fed into the network are standardized, noise-free, and better structured, thereby reducing the computational burden on the network and improving learning efficiency (Chen et al., 2021; Liu et al., 2019). The results confirm that optimized preprocessing pipelines, such as normalization, histogram equalization, iris boundary refinement, and eyelid occlusion removal, directly contribute to robust recognition accuracy and reduced error rates**.** By tailoring these preprocessing steps to the specific characteristics of the ORL and CASIA datasets, the proposed framework achieves better generalization and resilience under challenging conditions. This reinforces the argument that preprocessing should not be viewed as a generic step but as an essential design element that must be adapted to the dataset at hand. From an application perspective, the proposed system demonstrates strong potential for deployment in high-security domains such as airports, hospitals, research laboratories, and financial institutions, where reliability, precision, and robustness are essential (Winston &Hemanth, 2021). By integrating both physiological and structural biometric modalities, the framework enhances system trustworthiness, making it a suitable solution for real-world security challenges. Despite these strengths, some challenges remain. A key limitation lies in the computational overhead introduced by multi-branch CNN architectures**,** which can increase processing time and resource consumption. This may affect scalability, particularly when extending the framework to large-scale, real-time applications. To address these issues, future research should be explored in the use of lightweight CNN models, efficient transfer learning strategies and hardware acceleration techniques**,** which can maintain high accuracy while reducing computational complexity (Siau& Wang, 2020; Soleymani et al., 2018).

VIII. Conclusion

This study has emphasized the importance of dataset-specific preprocessing as a critical factor in improving the performance of CNN-based biometric recognition systems. By using the ORL face dataset and the CASIA-IrisV4 dataset as representative case studies, it was demonstrated that tailoring preprocessing strategies to the inherent characteristics of each dataset leads to measurable gains in recognition accuracy. Techniques such as normalization, histogram equalization, iris boundary refinement, and eyelid occlusion removal were shown to significantly reduce errors and strengthen system robustness.

The findings underscore the fact that preprocessing is not simply a preparatory step but a decisive stage in the design of biometric recognition pipelines. When appropriately optimized, preprocessing enhances the generalizability and resilience of CNN models, making them more capable of handling the variability and unpredictability of real-world deployment environments, particularly in security-critical applications such as access control systems. This research provides a foundation for extending dataset-specific preprocessing strategies into multimodal biometric fusion frameworks. Future work will be exploredin the integration of additional biometric modalities, such as voice and gait, while also investigating the development of adaptive preprocessing methods that can automatically adjust to new datasets. Such efforts are expected to further strengthen system reliability and facilitate scalable, real-world implementations of CNN-based biometric security solutions.

References

[1]. Abdo, M., Mekdad, Y., & Ahmad, A. (2022). Multimodal biometric fusion based on convolutional neural networks. *Sensors, 22*(5), 2451.

[2]. Abdo, M., Mekdad, Y., Ahmad, A., & Saeed, M. (2020). Feature-level fusion of face and iris using CNN. *International Journal of Computer Vision, 128*(9), 2405–2420.

[3]. All About Vision. (2024, March 20). How the iris works. Retrieved from https://www.allaboutvision.com/

[4]. Asha, M. J., Deepa, P. S., &Madhusoodanan, N. (2023). Face and iris verification using Siamese network. *Computers & Electrical Engineering, 105*, 108534.

[5]. Chen, H., Zhang, L., & Liu, X. (2021). Adaptive CNN feature extraction for multimodal biometrics. *Pattern Recognition Letters, 146*, 31–38.

[6]. Drozdowski, P., Rathgeb, C., Dantcheva, A., Busch, C., & Gomez-Barrero, M. (2018). Demographic bias in biometrics: A survey. *ACM Computing Surveys, 54*(4), 1–35.

[7]. Ghazal, M., & Abdullah, S. (2020). A deep multimodal fusion network for iris and face biometrics. *Multimedia Tools and Applications, 79*(3), 1797–1816.

[8]. Goodfellow, I., Bengio, Y., Courville, A., &Bengio, Y. (2013). Challenges in representation learning. *Neural Networks, 64*, 59–63.

[9]. Hammad, M., Alkasassbeh, M., & Abu-Qasmieh, I. (2018). Face and iris fusion using CNN. *Applied Sciences, 8*(12), 2459.

[10]. Hattab, M., &Behloul, F. (2023). A CNN model for iris segmentation and recognition. *Journal of King Saud University–Computer and Information Sciences, 35*(2), 135–144.

[11]. Jain, A. K., Ross, A., &Prabhakar, S. (2005). An introduction to biometric recognition. *IEEE Transactions, 14*(1), 4–20.

[12]. Jayanthi, S., Kumar, M. P., &Nandhini, S. (2021). Efficient CNN models for multimodal biometrics. *International Journal of Biometrics, 13*(1), 55–72.

[13]. Jung, K., Park, K. R., & Kim, J. (2017). Attention-based biometric recognition. *Image and Vision Computing, 64*, 34–43.

[14]. Koziarski, M., &Cyganek, B. (2017). Fusion strategies for multimodal biometrics. *Pattern Analysis and Applications, 20*(2), 489–502.

[15]. Liu, Z., Luo, J., & Lin, W. (2021). Face-iris fusion with CNNs. *Sensors, 21*(1), 244.

[16]. Liu, Z., Luo, J., & Lin, W. (2019). Attention mechanisms in CNNs. *Pattern Recognition Letters, 123*, 78–84.

[17]. Lu, Y., Zhang, S., & Chen, H. (2021). Iris recognition under noise using deep CNNs. *IEEE Access, 9*, 88711–88721.

[18]. Medjahed, S. A., Lakhdar, B., Khelifi, A., &Benlamri, R. (2020). Fusion strategies in biometric systems. *International Journal of Computer Applications, 175*(15), 5–13.

[19]. Nguyen, K., Yadav, D., &Sood, M. (2020). Triplet-based fusion for face and iris. *IEEE Access, 8*, 23411–23423.

[20]. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., &Batra, D. (2017). Grad-CAM: Visual explanations from deep networks. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 618–626).

[21]. Siau, J. H., & Wang, C. (2020). Siamese CNN for biometric authentication. *Expert Systems with Applications, 140*, 112873.

[22]. Soleymani, S. A., Ghasemi, E., Ghidary, S. S., &Nasrabadi, A. M. (2018). Biometric multimodal fusion using transfer learning. *Journal of Computer Science, 14*(4), 514–524.

[23]. Wang, H., Yu, F., & Wu, Y. (2018). Attention-guided CNNs for biometric fusion. *Computer Vision and Image Understanding, 169*, 36–45.

[24]. Winston, D. A., &Hemanth, D. J. (2021). Deep learning in multimodal biometrics. *Artificial Intelligence Review, 54*(3), 2035–2064.

[25]. Yadav, D., Singh, R., Vatsa, M., &Noore, A. (2018). Presentation attack detection in biometrics. *Pattern Recognition, 78*, 73–87.

[26]. Zanlorensi, L. A., Proença, H., de Carvalho, E., & Junior, J. (2019). Deep CNNs for biometric verification. *Pattern Recognition Letters, 123*, 58–65.

[27]. Zheng, W., Lu, Y., & Chen, S. (2017). Center loss for biometric learning. *Pattern Recognition, 67*, 132–142.

[28]. Zheng, W., Lu, Y., & Chen, S. (2018). Enhancing biometric models with contrastive loss. *Pattern Recognition, 79*, 69–80.