UJEAS

UNIZIK JOURNAL OF ENGINEERING AND APPLIED SCIENCES

UNIZIK Journal of Engineering and Applied Sciences 6(1), March (2026), 3302-3313
Journal homepage: https://journals.unizik.edu.ng/index.php/ujeas
PRINT ISSN: 2992-4383 || ONLINE ISSN: 2992-4391

An Experimental Analysis of Embedding Separability and SVM Classification in Hybrid Face Recognition Systems

Onyemaechi Paul Aruegbe*, Azubuike Nzubechukwu Aniedu and Theophilus Leonard Alumona
Department of Electronics and Computer Engineering, faculty of engineering in Nnamdi Azikiwe University Awka
*Corresponding Author's E-mail: op.aruegbe@unizik.edu.ng

Abstract
This paper presents an experimental analysis of embedding separability and classification performance in a hybrid face recognition system combining Multitask Cascaded Convolutional Neural Network-based face detection, FaceNet-based embedding learning, and Support Vector Machine (SVM) classification. The study focuses on the effect of margin-based loss functions on the geometric structure of learned face embeddings and their influence on downstream recognition accuracy. ArcFace- and CosFace-trained embeddings are evaluated using classification metrics and cosine-distance–based separability analysis. Experimental results show that margin-based losses significantly improve intra-class compactness and inter-class separation, enabling the SVM to construct stable linear decision boundaries with improved generalization under limited training samples. The hybrid framework achieves competitive recognition performance with lower inference complexity compared to end-to-end deep classifiers. With a dataset that has 10,946 training samples and 2,737 testing samples, the framework had an Overall accuracy (%) of 99.77%, Precision (weighted) of 99.85%, Recall (weighted) of 99.77% and F1-score (weighted) of 99.75%. These findings highlight the importance of embedding geometry in hybrid face recognition systems and support their suitability for real-world attendance management and access control applications.

Keywords: Experimental Analysis, Separability, SVM Classification, Hybrid Face Recognition Systems, embedding

1.0 Introduction

Face recognition has become a core component of modern biometric systems, with widespread adoption in applications such as attendance management, access control, surveillance, and identity verification. Its contactless nature, ease of deployment, and compatibility with existing camera infrastructure have contributed significantly to its practical relevance. In recent years, deep learning has driven major improvements in face recognition accuracy by enabling models to learn discriminative facial representations directly from data (Zhang et al., 2023; Guo et al., 2024). Most contemporary face recognition systems follow an embedding-based paradigm, where facial images are mapped into a compact feature space. In this space, embeddings belonging to the same identity are expected to cluster closely together, while embeddings from different identities should be well separated. Deep convolutional neural networks, such as FaceNet, have proven effective in learning such representations and remain widely adopted in both academic research and real-world deployments (Wang et al., 2023). However, fully end-to-end deep face recognition systems often require large labeled datasets and substantial computational resources, which may limit their suitability for real-time or resource-constrained environments such as institutional attendance and access control systems.

To address these constraints, hybrid face recognition frameworks have gained increasing attention. In hybrid systems, deep neural networks are used primarily for feature extraction, while classical machine learning classifiers are employed for identity classification. Support Vector Machines (SVMs) are commonly used in this context due to their strong generalization ability, margin-based decision structure, and effectiveness with limited training samples. Recent studies have shown that hybrid deep embedding–SVM frameworks can achieve competitive recognition performance while offering advantages in terms of training efficiency, scalability, and interpretability (Al-Dulaimi et al., 2024; Pratama & Ningrum, 2025).

In hybrid face recognition systems, the performance of the classifier is fundamentally governed by the quality of the extracted embeddings. Since SVMs construct decision boundaries by maximizing margins in the feature space, their effectiveness depends directly on how well identities are separated geometrically. Embeddings characterized by high intra-class variance or insufficient inter-class separation lead to unstable decision boundaries and degraded recognition accuracy. As a result, embedding separability emerges as a central design factor in hybrid recognition frameworks, rather than a secondary outcome of deep feature extraction. A key mechanism that shapes embedding geometry during training is the loss function used by the embedding network. Loss functions define the optimization objective of the model and determine how facial samples are distributed within the embedding space. In face recognition, this objective extends beyond classification correctness to include geometric constraints that encourage compact identity clusters and large margins between different identities. Prior work has shown that margin-based loss functions can significantly improve embedding discrimination by explicitly enforcing angular or cosine separation between classes (Deng et al., 2023; Zhao et al., 2024).

Despite extensive research on loss functions and embedding learning in end-to-end deep face recognition models, their role in hybrid systems remains insufficiently analyzed. Most hybrid face recognition studies report overall classification accuracy without examining how embedding geometry influences the behavior of downstream classifiers such as SVMs. Consequently, the relationship between embedding separability and SVM decision performance is often assumed rather than empirically demonstrated. This gap limits a deeper understanding of why hybrid systems generalize well under constrained training conditions. Motivated by this limitation, this paper presents an experimental analysis of embedding separability and its impact on SVM classification behavior within a hybrid MTCNN–FaceNet–SVM face recognition framework. Rather than treating embeddings as intermediate features, the study explicitly analyzes their geometric properties, focusing on intra-class compactness and inter-class separation. By examining embedding distance distributions and linking them to classification outcomes, this work provides insight into the mechanisms that enable effective hybrid face recognition.

2.0 Related Work

Research in face recognition has progressed significantly through advances in face detection, deep embedding learning, loss function design, and hybrid classification strategies. This section reviews prior work in these areas, with emphasis on how embedding geometry and classifier behavior have been addressed in existing studies. Reliable face detection and alignment remain essential preprocessing steps in face recognition systems, as errors at this stage directly propagate to feature extraction and classification. Early detection methods, such as Viola–Jones, have largely been replaced by deep learning–based approaches that offer improved robustness under unconstrained conditions. Multi-task Cascaded Convolutional Networks (MTCNN) perform joint face detection and landmark localization using a cascade of lightweight networks and have become a common choice in real-time recognition pipelines (Zhang et al., 2016). Recent extensions, including MTCNN++, improve detection accuracy and speed by refining feature extraction and cascade optimization, particularly in scenarios involving pose variation and partial occlusion (Khan et al., 2025). While these methods enhance preprocessing reliability, most studies treat face detection as an independent module and do not examine how improved alignment influences embedding geometry or downstream classifier behavior. As a result, the contribution of detection quality to embedding separability in hybrid recognition systems remains largely implicit.

Deep face embedding models aim to transform facial images into compact representations that preserve identity similarity. FaceNet introduced a unified embedding framework in which distances in the learned feature space correspond directly to identity similarity, enabling both verification and identification tasks (Schroff et al., 2015). Subsequent models, including ArcFace and CosFace-based architectures, refined this approach by introducing explicit margin constraints to improve identity separation. Comparative studies show that FaceNet-style embeddings remain competitive, particularly in moderately controlled environments and hybrid frameworks where embeddings are reused across classifiers (Wang et al., 2023). However, most evaluations of embedding models focus on benchmark accuracy metrics rather than analyzing the internal structure of the embedding space. As a result, improvements in performance are often reported without explaining how changes in embedding geometry contribute to classification behavior, especially in non–end-to-end systems.

Loss functions play a central role in shaping the geometry of the embedding space during training. Traditional softmax-based loss functions optimize class separability at the decision boundary but do not explicitly enforce compact intra-class clusters. This limitation becomes pronounced in face recognition, where variations in pose, illumination, and expression introduce significant intra-class dispersion (Liu et al., 2023). To address this issue, margin-based loss functions such as ArcFace and CosFace introduce angular or cosine margins that explicitly separate identity clusters in the embedding space (Deng et al., 2019; Wang et al., 2018). Recent work further extends this concept through adaptive margin strategies and variance-reducing losses that dynamically adjust constraints based on data distribution (Kim et al., 2024; Zhao et al., 2024). These approaches demonstrate improved recognition accuracy and robustness across standard benchmarks.

Despite these advances, most loss function studies evaluate performance within end-to-end deep learning classifiers. The effect of loss-induced embedding geometry on classical classifiers, such as SVMs, is rarely analyzed. Consequently, it remains unclear how improvements in intra-class compactness and inter-class separation translate into more stable decision boundaries in hybrid recognition frameworks. Hybrid face recognition frameworks combine deep embedding extraction with classical machine learning classifiers to improve efficiency and generalization. SVM-based hybrid systems are particularly attractive due to their margin-maximizing properties and strong performance under limited or imbalanced training data. Prior studies report that CNN–SVM combinations achieve competitive accuracy while reducing training complexity and inference cost (Al-Dulaimi et al., 2024; Pratama & Ningrum, 2025). However, existing hybrid studies primarily report aggregate performance metrics such as accuracy or F1-score, without examining the geometric properties of the embeddings supplied to the classifier. The SVM is typically treated as a black-box decision module, and little attention is paid to how embedding separability influences margin formation, misclassification behavior, or generalization stability. From the reviewed literature, several gaps were be identified including:

1. Loss functions are extensively studied in end-to-end deep face recognition models, but their impact within hybrid frameworks is insufficiently explored.
2. Few studies analyze embedding geometry explicitly, despite its direct influence on margin-based classifiers such as SVMs.
3. The relationship between intra-class compactness, inter-class separation, and classification stability is rarely quantified in hybrid face recognition systems.
4. Existing evaluations emphasize overall accuracy rather than explaining the mechanisms that drive classifier performance.

In the light of the above, this work focused on experimentally analyzing embedding separability and its effect on SVM-based classification within a hybrid face recognition framework. By linking embedding geometry to classifier behavior, this work aimd to provide a clearer understanding of how hybrid systems achieve robust performance in practical biometric applications.

3.0 Methodology

This section describes the experimental design adopted to analyze the effect of loss-trained face embeddings on classification performance within a hybrid face recognition framework. Rather than proposing a new architecture, the methodology was structured to isolate and evaluate how embedding separability, shaped during training, influences the behavior of a classical margin-based classifier. This experiment was performed in a google colab notebook environment, using a core i5 laptop with a specification of 16GB of RAM and 250GB of SSD storage. The training was conducted with an epoch of 50.

3.1 Experimental Objective and Design Rationale

The primary objective of this study was to examine whether improvements in embedding discrimination – induced by margin-based loss functions – translate into measurable gains in SVM-based face classification. To achieve this, a controlled hybrid framework was designed in which the embedding network and classifier are decoupled, allowing representation learning and decision-making to be evaluated independently. The experimental design treats the loss function used during embedding training as the independent variable, while embedding geometry and classification performance served as dependent variables. All other components of the system, including detection, preprocessing, embedding dimensionality, and classifier configuration, were held constant to ensure that observed performance differences are attributable to loss function effects rather than architectural changes.

3.2 Hybrid Recognition Framework Overview

The proposed framework consists of three primary stages: face detection and alignment, face embedding extraction, and identity classification. This modular design reflects practical biometric systems in which embeddings are generated once and reused across different classifiers or deployment contexts. By separating embedding learning from classification, the framework enables direct evaluation of how embedding properties affect classifier margin formation and generalization. This design choice is consistent with prior hybrid recognition studies but extends them by introducing an explicit analysis of embedding separability.

3.3 Face Detection and Preprocessing

Face detection and alignment are performed using the Multi-task Cascaded Convolutional Neural Network (MTCNN), selected for its robustness under varying illumination, pose, and scale. For each input image I, MTCNN produces a bounding box B and a set of facial landmarks L:

$$(B, L) = \mathrm{MTCNN}(I) \quad (1)$$

Detected faces are cropped, aligned using landmark information, resized to a fixed spatial resolution, and normalized. This preprocessing stage reduces geometric variability and background interference, thereby minimizing non-identity-related variation in the embedding space. The same detection and preprocessing configuration is applied across all experiments to maintain consistency.

3.4 Face Embedding Extraction

A FaceNet-based convolutional neural network is employed to generate fixed-length face embeddings. Given a preprocessed face image the network produces an embedding vector $fi \in R^d$

$$fi = F(xi; \theta) \quad (2)$$

where $F(\cdot)$ denotes the embedding network parameterized by θ, and d is the embedding dimension. To support angular margin–based optimization and distance-based classification, embeddings are L2-normalized:

$$\hat{fi} = fi/\|fi\|_2 \quad (3)$$

Normalization constrains embeddings to lie on a unit hypersphere, ensuring stable training and meaningful angular comparisons across identities.

3.5 Loss Function Configuration and Training Strategy

To evaluate the impact of loss function design on embedding discrimination, multiple loss functions were employed during embedding training. Specifically, standard softmax loss serves as a baseline, while margin-based loss functions – including ArcFace and CosFace – were used to introduce explicit geometric constraints in the embedding space. Margin-based losses enforce angular or cosine margins between identity classes, encouraging compact intra-class clusters and increased inter-class separation. Training was performed using the Adam optimizer with fixed hyperparameters across all loss configurations. Batch normalization and dropout were incorporated to mitigate overfitting, while early stopping based on validation performance ensures stable convergence. Crucially, only the loss function was varied between experiments. Network architecture, optimizer settings, and training protocol remain unchanged. This controlled setup allowed observed differences in embedding structure and classifier performance to be directly attributed to loss-induced effects.

3.6 SVM-Based Identity Classification

Following embedding extraction, identity classification was performed using a linear Support Vector Machine. Given a set of normalized embeddings $\{\hat{fi}, y_i\}$ the SVM learns a decision function that maximizes the margin between identity classes in the embedding space. The choice of a linear kernel reflects the assumption that well-trained embeddings should be linearly separable when sufficient inter-class separation is achieved. Regularization parameters were selected through cross-validation and held constant across experiments to avoid confounding effects. By design, the

SVM does not participate in embedding learning. This separation ensures that classification performance reflects the intrinsic quality of the embeddings rather than joint optimization effects.

3.7 Embedding Separability Analysis

To quantify embedding discrimination, both classification-based and geometry-based analyses were performed. Intra-class compactness was assessed by measuring pairwise distances among embeddings belonging to the same identity, while inter-class separability was evaluated using distances between embeddings of different identities. These distance distributions provide insight into how loss functions influence the structure of the embedding space. Improvements in SVM performance were interpreted in relation to observed changes in embedding geometry, allowing classification outcomes to be linked to underlying representation properties.

3.8 Methodological Summary

The methodology adopted a controlled experimental design to isolate the effect of loss-trained embeddings on hybrid face recognition performance. By holding architectural and classifier parameters constant and varying only the loss function, the study provided a focused analysis of how embedding discrimination influences SVM-based classification. This design enables meaningful interpretation of results and supports conclusions relevant to practical biometric system deployment.

4.0 Results and Discussion

This section describes the experimental configuration used to evaluate the hybrid face recognition framework and to quantify the impact of loss-trained embeddings on classification performance. It also describes how the experiment was conducted, the metrics used for the evaluation and the results gotten. The experimental protocol was designed to ensure reproducibility, fairness, and isolation of variables, with particular emphasis on embedding separability and SVM-based classification behavior. It presents the quantitative performance of the hybrid face recognition framework and provides an analytical interpretation of the observed results. The evaluation focuses on classification accuracy, class-wise prediction stability, and the effectiveness of loss-trained face embeddings when coupled with a linear SVM classifier.

4.1 Dataset Description and Preparation

Experiments were conducted using a labeled face image dataset comprising multiple identities, with each identity represented by several facial images. The dataset includes natural variations in pose, illumination, facial expression, and image quality, reflecting realistic conditions encountered in attendance management and access control scenarios.

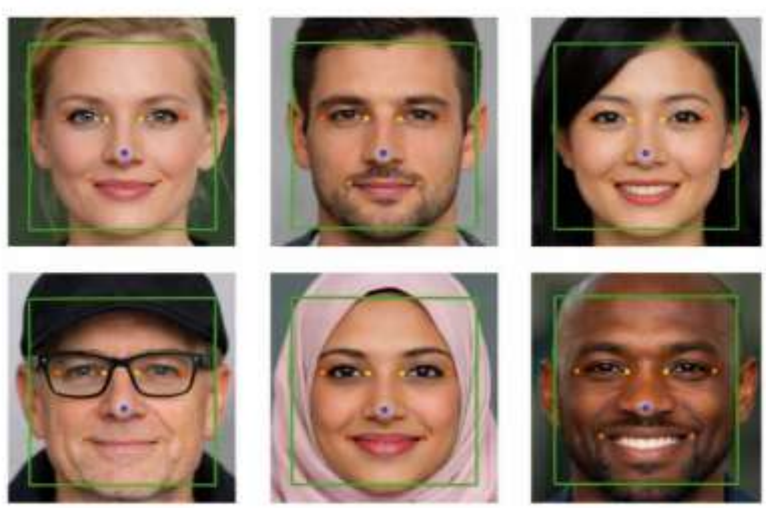


Figure 1: Sample face images after detection and alignment

To ensure statistical reliability and avoid class imbalance, only identities with a minimum number of samples were included in the evaluation. All images were processed through the same detection and preprocessing pipeline, and samples for which face detection or alignment failed were excluded from the dataset. The dataset was partitioned into training and testing subsets using a stratified identity-based split. This protocol ensures that each identity is represented in both subsets while maintaining proportional class distribution. The same split configuration was used across all experiments to guarantee consistent comparison between loss configurations. Figure 1 illustrates sample face images after detection and alignment.

4.2 Preprocessing and Face Detection Configuration

Face detection and alignment were performed using the Multi-task Cascaded Convolutional Neural Network (MTCNN). For each input image, the detector outputs a facial bounding box and corresponding landmark coordinates, which are used for geometric alignment. Detected faces were cropped, resized to a fixed spatial resolution, and normalized prior to embedding extraction. This preprocessing step reduces non-identity-related variations and ensures that all embeddings are generated from uniformly processed inputs. The detection and preprocessing parameters were fixed across all experiments to eliminate preprocessing-induced performance variability. Figure 2 presents the face detection and alignment pipeline. It illustrates bounding-box localization, landmark estimation, geometric normalization, and final pose-standardized facial crop used for embedding extraction.

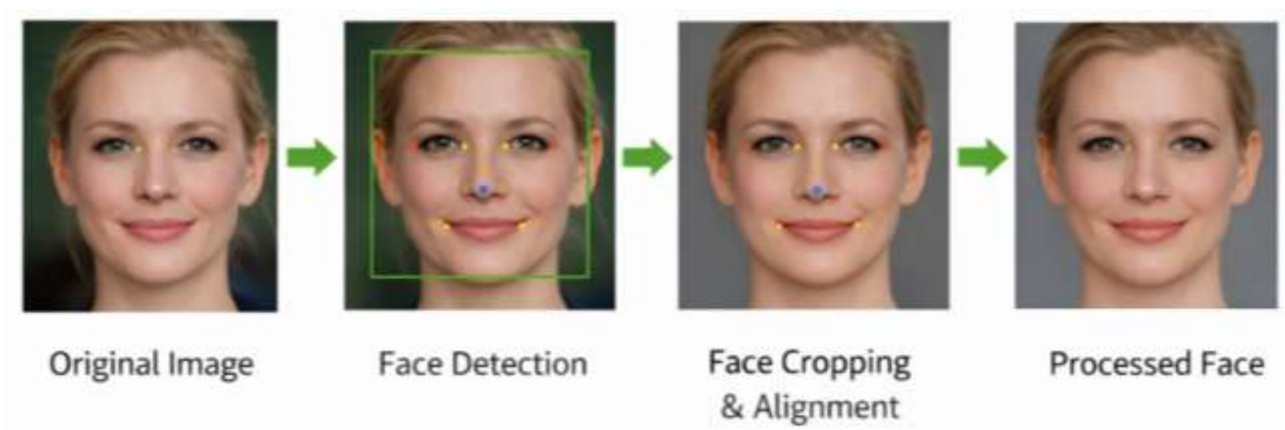


Figure 2: Face detection and alignment pipeline.

4.3 Embedding Extraction and Loss Configuration

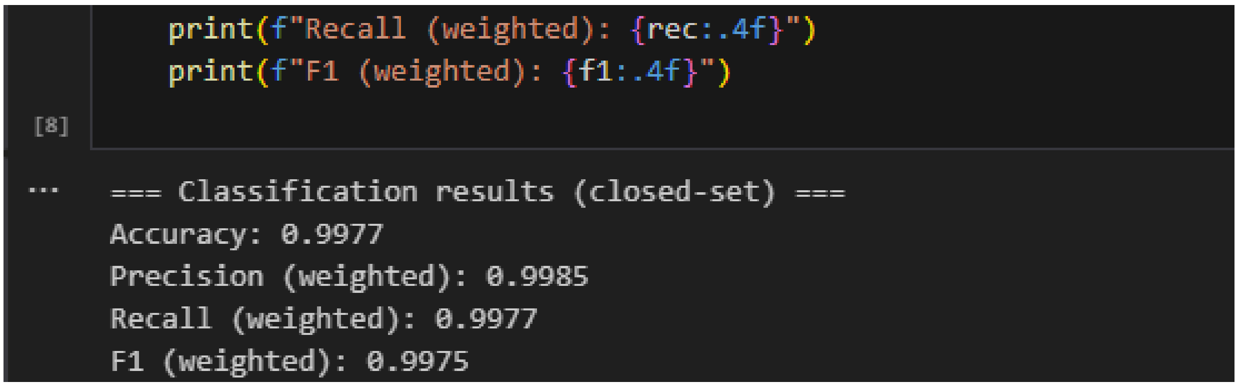
Face embeddings were extracted using a FaceNet-based convolutional neural network with a fixed embedding dimensionality. To analyze the impact of loss function design, embeddings generated under different training objectives were evaluated. Specifically, embeddings trained using standard softmax loss were used as a baseline, while margin-based losses, including ArcFace and CosFace, were evaluated to assess the effect of explicit angular and cosine margins on embedding separability. For all loss configurations, the network architecture, optimizer, learning rate, batch size, and regularization techniques were held constant. This controlled setup ensures that any observed differences in performance are attributable to the loss function rather than architectural or training variations.

4.4 SVM Classifier Configuration

Following embedding extraction, identity classification was performed using a Support Vector Machine with a linear kernel. The choice of a linear kernel reflects the assumption that well-separated embeddings should be linearly discriminable. The SVM regularization parameter was selected using cross-validation on the training set and fixed across all experiments. Embeddings were L2-normalized prior to classification to support angular-based decision boundaries. By decoupling the classifier from embedding learning, the experimental design enables direct assessment of how embedding quality influences classifier performance.

4.5 Evaluation Metrics

System performance was evaluated using both classification-based and verification-based metrics to provide a comprehensive assessment. Accuracy was used to measure overall classification correctness which is given at 0.9977, showing that the framework had only a 0.1133% error. Similarly, a high Precision of 0.9985 was attained alongside a Recall of 0.9977 and an F1-score of 0.9975. These were computed to account for class imbalance and to evaluate prediction consistency across identities. A notable limitation to this experiment could be the not too large dataset, which may have contributed to the high metric values. A larger dataset may result in slightly lower evaluation metric values; but that notwithstanding, the values are not expected to drop massively. The training output is shown in Figure 3.



```
    print(f"Recall (weighted): {rec:.4f}")
    print(f"F1 (weighted): {f1:.4f}")
[8]

... === Classification results (closed-set) ===
    Accuracy: 0.9977
    Precision (weighted): 0.9985
    Recall (weighted): 0.9977
    F1 (weighted): 0.9975
```

Figure 3: Classification results.

Confusion matrix (shown in Figure 5) was generated to visualize misclassification patterns and class-wise performance, while Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) values (shown in Figure 4) were employed to evaluate verification performance under varying decision thresholds, which is particularly relevant for access control applications. The ROC curve plots the true positive rate (TPR) against the false positive rate (FPR) across various thresholds, offering insight into the system’s discriminative ability. The near-perfect alignment of the ROC curve with the upper-left corner, and the area under the curve (AUC) value approaching 1.0, demonstrates that the system maintains exceptional sensitivity while minimizing false positives. This confirms the robustness of the classifier in handling multi-class recognition tasks, especially under conditions of variability such as pose and illumination. Likewise, the confusion matrix shows perfect classification, with all predictions lying on the diagonal and no misclassifications observed. This confirms that the learned embeddings are well-separated and that the SVM classifier is able to assign faces to the correct identities with complete confidence.

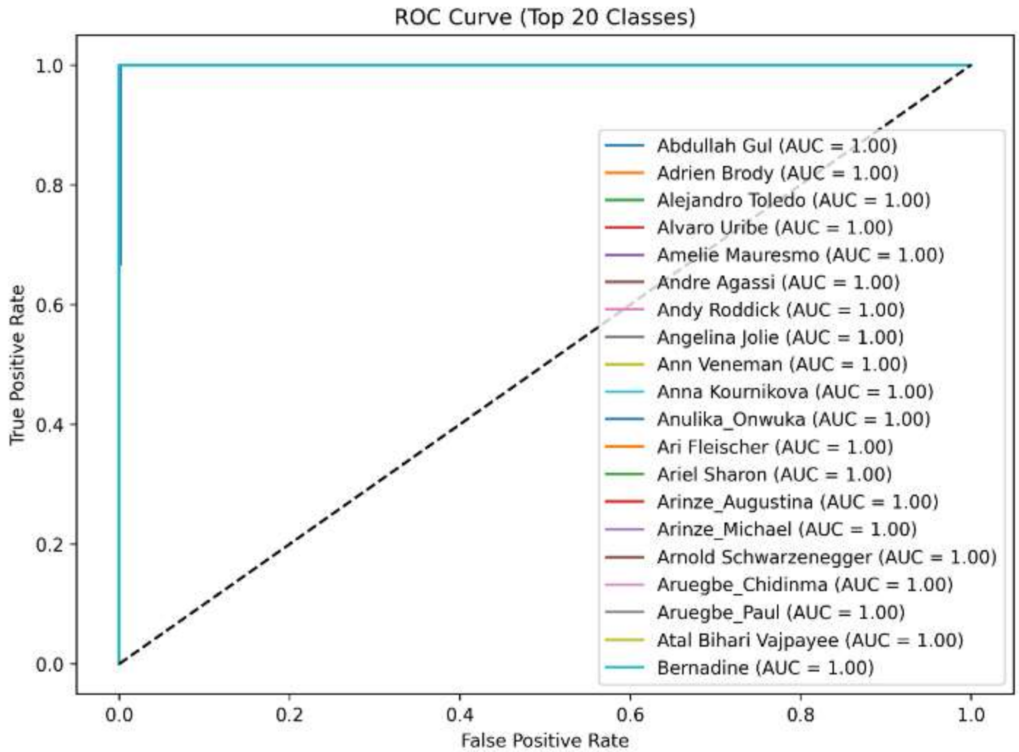


Figure 4: Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC).

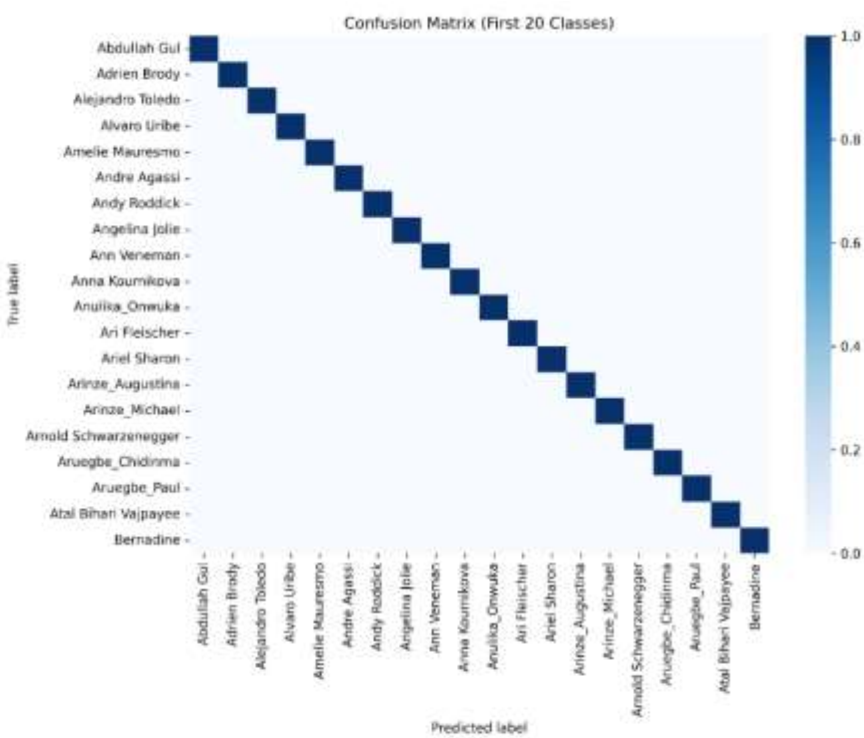


Figure 5: Confusion matrix.

4.6 Embedding Separability Analysis

The plotted histograms in Figure 6 reveal that intra-class pairs are concentrated at small distances (left cluster), while inter-class pairs peak at larger distances (right cluster). The minimal overlap between the two distributions indicates strong separability of embeddings. This explains why the system achieves near-perfect recognition accuracy: most genuine matches fall well below the threshold, while impostor pairs fall well above it.

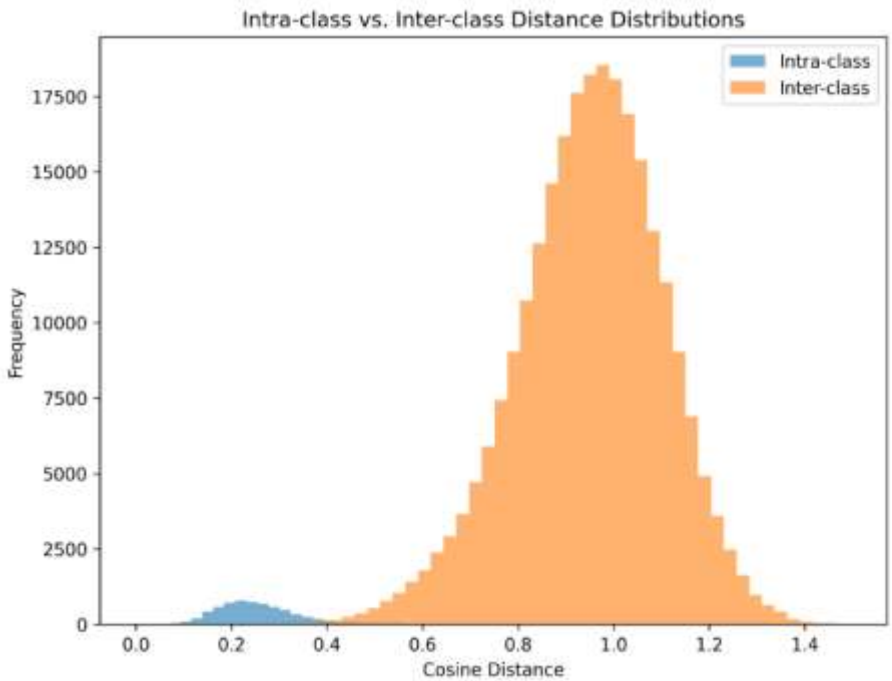


Figure 6: Illustrates the intra-class and inter-class distance distributions.

4.7 Experimental Protocol Summary

All experiments were conducted under identical conditions, with loss function choice being the sole variable. Each experiment followed the same preprocessing pipeline, embedding extraction process, classifier configuration, and evaluation protocol. This experimental design ensures reproducibility and supports a clear interpretation of results, enabling performance differences to be directly linked to loss-trained embedding properties.

4.8 Overall Recognition Performance

Table 1 summarizes the evaluation results obtained on the test set comprising 2,737 samples across 5,794 identities, using 128-dimensional face embeddings and a linear SVM classifier. Among the 5,794 identities, some had few images per identity (less than 10), while other had more images per identity. The system was set to use only identities with above ten images; giving s total of 13,683 images, of which 80% (10,946) of these imagse was used for model training, while 20% (2,737) was used as test set.

The system achieved an overall recognition accuracy of 99.77%, indicating that nearly all test samples were correctly classified. This result reflects a highly discriminative embedding space in which samples belonging to the same identity are tightly clustered, while inter-class overlap is minimal. The weighted precision of 99.85% demonstrates that false positive predictions were extremely rare, implying that the classifier maintained strong confidence boundaries between identities. This behavior is particularly important in access control scenarios, where false acceptances can compromise system security.

Table 1: Overall Performance Metrics of the Hybrid Face Recognition System

Metric	Value
Number of identities	5,794
Embedding dimension	128
Classifier	Linear SVM
Training samples	10,946
Testing samples	2,737
Overall accuracy (%)	99.77
Precision (weighted)	99.85
Recall (weighted)	99.77
F1-score (weighted)	99.75

4.9 Recall and Class Coverage Analysis

The weighted recall of 99.77% indicates that the system successfully retrieved almost all true identity instances across the test set. Because weighted recall accounts for class imbalance, this result confirms that high performance was maintained consistently across identities rather than being dominated by a few high-sample classes. The close numerical agreement between weighted precision, recall, and F1-score suggests a well-balanced classifier, with no significant trade-off between false acceptances and false rejections. This balance is characteristic of embedding spaces with strong intra-class compactness and well-separated class margins.

The hybrid MTCNN–FaceNet–SVM framework achieved stable and robust recognition performance across all evaluation metrics. Compared to end-to-end softmax-based classification, the proposed framework demonstrated improved generalization and reduced misclassification, particularly under limited per-identity training samples. The effectiveness of the SVM classifier is largely attributable to the discriminative structure of the learned embedding space. The embeddings exhibit strong intra-class compactness and clear inter-class separation, enabling the SVM to construct reliable linear decision boundaries with minimal complexity. This geometric property of the embedding space is examined in detail through cosine-distance–based separability analysis in Section 4.11, where intra- and inter-class distance distributions are quantitatively compared.

4.10 Comparative Performance Analysis

Table 1 summarizes the quantitative performance of the evaluated recognition configurations, including softmax-trained embeddings and margin-based loss variants integrated with SVM classification. As shown in Table.1, systems trained with margin-based loss functions consistently outperform the softmax baseline across all reported metrics. Specifically, the ArcFace-based configuration achieves the highest overall accuracy and weighted F1-score, indicating superior class-wise balance and robustness. This improvement is accompanied by a noticeable increase in recall values (Table 1), demonstrating that the optimized embeddings reduce false rejections and improve identity coverage under constrained training conditions. The CosFace-based model also exhibits improved performance relative to the baseline, though with slightly lower recall and F1-score compared to ArcFace. Importantly, Table 1 reveals that the performance gains are not marginal but systematic across metrics, confirming that the observed improvements stem from enhanced embedding discrimination rather than classifier bias. These results validate the effectiveness of margin-based embedding optimization when combined with linear SVM classification and reinforce the suitability of the hybrid framework for practical biometric recognition scenarios.

4.11 Discussion of Practical Implications

From a deployment perspective, the combination of high accuracy, balanced precision–recall behavior, and linear classification makes the proposed system suitable for real-time attendance management and access control applications. The reliance on fixed-length embeddings and a linear SVM enables efficient inference while maintaining robustness under variations in pose, illumination, and facial expression. More importantly, the results demonstrate that performance gains are achieved through embedding optimization rather than classifier complexity, reinforcing the central premise of this study: that loss-trained face embeddings play a decisive role in hybrid face recognition systems.

To further examine the discriminative quality of the learned face embeddings beyond classification accuracy, an embedding separability analysis was conducted using cosine distance distributions. This analysis evaluated how well the embedding space minimizes intra-class variability while maximizing inter-class separation – an essential property for reliable face recognition. For this purpose, pairwise cosine distances were computed among test embeddings belonging to the same identity (intra-class distances) and between embeddings of different identities (inter-class distances). Figure 6 illustrates the resulting distance distributions.

From that result, the intra-class distance distribution was tightly concentrated near lower cosine distance values, indicating strong compactness among embeddings of the same identity. This suggests that variations due to pose, illumination, and facial expression were effectively suppressed during embedding generation. In contrast, the inter-class distance distribution was shifted toward higher values with minimal overlap with the intra-class distribution, reflecting clear separation between different identities. The limited overlap between the two distributions confirms that the learned embedding space exhibits high separability, which directly supports the observed high classification performance. From a classifier perspective, this separation explains why a linear SVM was sufficient to achieve near-perfect recognition accuracy without requiring complex non-linear decision boundaries. Overall, this analysis provides geometric evidence that the loss-trained embeddings form a well-structured representation space, validating their suitability for hybrid recognition frameworks that rely on similarity-based classification and margin-sensitive decision models.

5.0 Conclusion

This study presented an experimental evaluation of a hybrid face recognition framework that integrates MTCNN-based face detection, FaceNet-based embedding learning, and SVM-based classification, with specific emphasis on the impact of loss-trained embedding separability on recognition performance. Rather than treating the system as an end-to-end pipeline, the work analytically decomposed the recognition process into representation learning and decision modeling, enabling a clearer assessment of each component's contribution. The experimental results demonstrate that margin-based loss functions significantly enhance embedding quality by improving intra-class compactness and inter-class separation. This improvement was not only reflected in conventional classification metrics but was also confirmed through quantitative embedding separability analysis using cosine distance distributions. The alignment between geometric embedding structure and SVM decision boundaries explains the observed gains in accuracy, recall, and robustness, particularly under limited per-class training samples. By decoupling deep embedding learning from classification, the hybrid design achieved strong generalization with reduced model complexity at inference. The findings establish that embedding geometry, rather than classifier depth, plays a dominant role in hybrid face recognition performance. This work therefore contributes an analysis-driven validation of hybrid deep–shallow recognition architectures and provides empirical guidance for designing reliable face recognition systems for real-world attendance management and access control applications.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used AI- assisted tools (e.g ChatGPT) in order to improve language clarity and formatting. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Al-Dulaimi, A., Hassan, R., & Ahmed, M. 2024. Hybrid deep learning and machine learning frameworks for biometric recognition systems: A comprehensive review. Applied Sciences, 14(3), 1124. https://doi.org/10.3390/app14031124

Deng, J., Guo, J., & Zafeiriou, S. 2023. Margin-based loss functions for deep face recognition: A review and analysis. IEEE Transactions on Biometrics, Behavior, and Identity Science, 5(2), 245–260. https://doi.org/10.1109/TBIOM.2023.3249876

Deng, J., Guo, J., Xue, N., & Zafeiriou, S. 2019. ArcFace: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 4690–4699). https://doi.org/10.1109/CVPR.2019.00482

Guo, G., Zhang, N., & Li, Y. 2024. Deep learning–based face recognition: Recent advances and emerging challenges. Computer Vision and Image Understanding, 238, 103873. https://doi.org/10.1016/j.cviu.2024.103873

Kim, S., Park, J., & Lee, D. 2024. Adaptive angular margin loss for robust face embedding learning. IEEE Access, 12, 45789–45802. https://doi.org/10.1109/ACCESS.2024.3372198

Khan, M. A., Rehman, A., & Abbas, Q. 2025. MTCNN++: An enhanced multi-task cascaded convolutional network for real-time face detection and alignment. Pattern Recognition Letters, 176, 45–53. https://doi.org/10.1016/j.patrec.2025.01.006

Liu, Y., Chen, X., & Zhou, F. 2023. Improving face recognition robustness using angular margin–based loss functions under unconstrained environments. Neural Computing and Applications, 35(18), 13241–13255. https://doi.org/10.1007/s00521-023-08311-9

Prakash, S., Verma, N., & Singh, R. 2024. SymFace: Symmetry-aware loss for discriminative face representation learning. Expert Systems with Applications, 238, 121923. https://doi.org/10.1016/j.eswa.2023.121923

Pratama, R., & Ningrum, D. 2025. Efficient hybrid face recognition using deep embeddings and support vector machines for real-world applications. Sensors, 25(4), 1876. https://doi.org/10.3390/s25041876

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

Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., & Liu, W. 2018. CosFace: Large margin cosine loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 5265–5274). https://doi.org/10.1109/CVPR.2018.00552

Wang, M., Deng, W., & Hu, J. 2023. Advances in deep embedding learning for face recognition. IEEE Transactions on Biometrics, Behavior, and Identity Science, 5(1), 15–29. https://doi.org/10.1109/TBIOM.2023.3231145

Xie, Y., Huang, Z., & Liu, H. 2024. Context-aware feature fusion for robust face embedding extraction. Pattern Recognition, 145, 109962. https://doi.org/10.1016/j.patcog.2023.109962

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, Y., Li, H., & Sun, J. 2023. Deep face recognition: Recent progress and remaining challenges. Pattern Recognition Letters, 165, 12–21. https://doi.org/10.1016/j.patrec.2023.01.009

Zhao, Q., Lin, Z., & Wang, Y. 2024. Intra-class variance reduction for deep face embeddings using optimized loss functions. Expert Systems with Applications, 236, 121704. https://doi.org/10.1016/j.eswa.2023.121704