Comparative Analysis of Facial Recognition Based Attendance System Using Machine Learning

Pradyumna D[1*] *and Dr. Balasubramanian* Prabhu Kavin B[1]

[1]Department of Data Science and Business Systems, SRM Institute of Science and Technology, Kattankulathur, India

Abstract. Automated attendance systems benefit from robust, scalable facial recognition models capable of handling real-world classroom variability illumination, pose, and occlusion. This study conducts a systematic comparison of four pipelines ArcFace, SFace, GhostFaceNet, and Dlib using the DeepFace and face recognition frameworks on a custom dataset of approximately 13,000 images representing 1,600 identities. Evaluation metrics include identification accuracy (Top-1), verification reliability (AUC, TAR@FAR), latency, and throughput (FPS). Dlib achieved the best overall performance (Top-1 = 87.93%, AUC = 0.9952, 11.53 FPS), outperforming deeper CNN-based embeddings in both speed and accuracy. The benchmark highlights trade-offs between accuracy, discriminability, and real-time efficiency, providing practical deployment insights for classroom automation.

1 Introduction

Biometric attendance systems reduce manual errors and administrative friction but demand recognition models that are both highly accurate and computationally efficient. Real-world classroom environments introduce several challenges variable illumination, motion blur, occlusions, and heterogeneous camera viewpoints—that degrade deep learning performance. While substantial prior work explores face recognition algorithms, there remains a lack of systematic benchmarking tailored to operational attendance systems deployed in large academic settings.

To address this gap, this study provides a unified evaluation of four face recognition backbones ArcFace, SFace, GhostFaceNet, and Dlib—using a standardized experimental pipeline. The novelty of the work lies in conducting the first systematic comparative benchmark of these four models specifically for largescale classroom attendance systems, including deep verification metrics, latency profiling, and deployment-focused engineering insights. Existing studies often evaluate models in isolation or on generic datasets, whereas this work offers a comprehensive real-world analysis that captures the trade-offs between accuracy, verification reliability, and inference efficiency required for scalable attendance automation.

*Corresponding author: pd3771@srmist.edu.in

1.1 Objectives and Contributions

This work makes the following contributions:

- A reproducible benchmarking framework for folder-organized datasets (known faces/person/*.jpg) applicable to real-world attendance systems.
- A systematic comparison of four widely used recognition backbones (Ar- cFace, SFace, GhostFaceNet, Dlib) across identification, verification, and runtime metrics.
- A rigorous technical analysis including ROC curves, correlation heatmaps, latency–accuracy trade-offs, and weighted efficiency scores.
- Deployment-oriented insights covering threshold calibration, alignment reliability, and throughput optimization for classroom automation.
- A novel benchmark perspective: this is the first work providing an integrated comparison of these models specifically for attendance pipelines, offering engineering-focused takeaways beyond prior standalone studies.

2 Literature Review

Early face recognition relied on handcrafted features (LBP, HOG, SIFT) and shallow classifiers, which failed under pose or illumination variation. Deep metric learning introduced discriminative embeddings optimized directly on large-scale datasets. ArcFace [1] introduced an additive angular margin loss that improved separability, while ElasticFace [2] and AdaFace further incorporated adaptive margins to increase robustness.

SFace [3] employs a sigmoid-constrained hypersphere loss, enabling smoother separation of low-quality samples. GhostFaceNet [4] optimizes efficiency using Ghost modules for lightweight CNN computation. Classical approaches such as Dlib remain relevant for low-latency applications. Although several frameworks (DeepFace, face recognition) provide standard pipelines, few studies address the specific requirements of classroom attendance systems—namely speed, through- put, and robustness under academic environmental constraints. This motivates the present benchmarking effort.

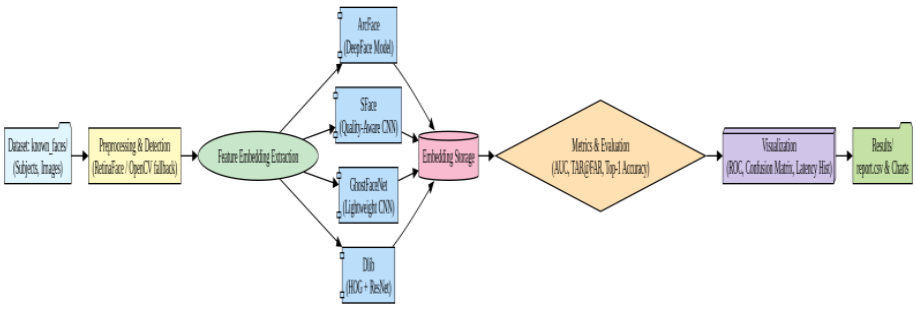

Fig. 1: System Architecture for Classroom Attendance Pipeline

3 System Architecture

The attendance system pipeline (Fig. 1) integrates several core modules: 1) **Detection and Alignment:** Using RetinaFace (with OpenCV fallback) ensures consistent facial localization. 2) **Embedding Extraction:** The aligned face is embedded via ArcFace, SFace, GhostFaceNet, or Dlib models. 3) **Matching and Decision:** Cosine similarity computes identity match scores, applying adaptive thresholds derived from ROC

calibration. 4) **Attendance Logging:** Verified faces are logged into a structured SQLite database.
The benchmark harness automates dataset evaluation, produces per-model metrics, confusion matrices, ROC plots, and consolidated CSV summaries.

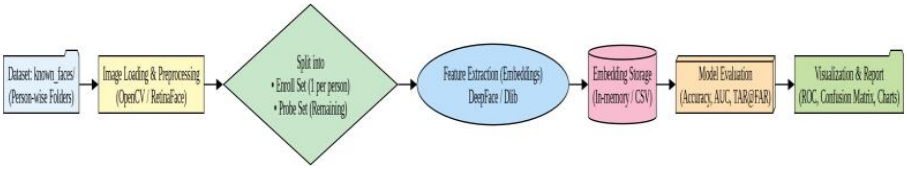

Fig. 2: Benchmark Data and Results Flow.

4 Methodology

Four pretrained backbones were tested via standardized pipelines:

- **ArcFace (DeepFace)** – Additive angular margin; strong discriminability.
- **SFace (DeepFace)** – Sigmoid-constrained hypersphere; robust to label noise.
- **GhostFaceNet (DeepFace)** – Lightweight Ghost-module CNN; optimized for speed.
- **Dlib (face recognition)** – Classical 128-D embedding; stable and efficient baseline.

Evaluation Metrics

- Identification: Top-1 accuracy (nearest-neighbor classification).
- Verification: AUC and TAR@FAR = 1e–3, 1e–4.
- Latency and Throughput: Mean per-image runtime and FPS.
- Extended Metrics: Efficiency (accuracy per latency) and weighted composite score.

Experimental Setup

- **Dataset:** ~13K images, ~1.6K identities, split into enrollment/probe sub- sets.
- **Hardware:** NVIDIA RTX 3050 Laptop GPU, Arch Linux, Python 3.10.
- **Frameworks:** DeepFace, face recognition, scikit-learn, OpenCV.

5 Results

Table 1 presents the consolidated metrics for each model. Dlib achieved the best overall accuracy and AUC, while maintaining high throughput (11.53 FPS).

Table 1: Model-wise Performance Comparison

Model	**Top-1**	**AUC**	**TAR@1e–3**	**TAR@1e–4**	**Latency (s)**	**FPS**
ArcFace	0.7727	0.9763	0.8143	0.7230	0.1582	6.32
SFace	0.5963	0.9500	0.5437	0.5090	0.1183	8.45
GhostFaceNet	0.5395	0.9373	0.5630	0.3708	0.1857	5.39
Dlib	**0.8793**	**0.9952**	**0.9305**	**0.8327**	**0.0867**	**11.53**

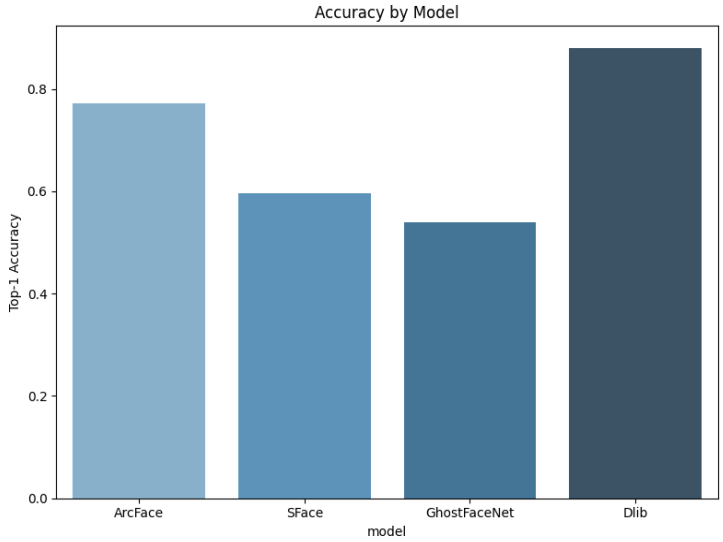

Fig. 3: Accuracy Comparison Across Models.

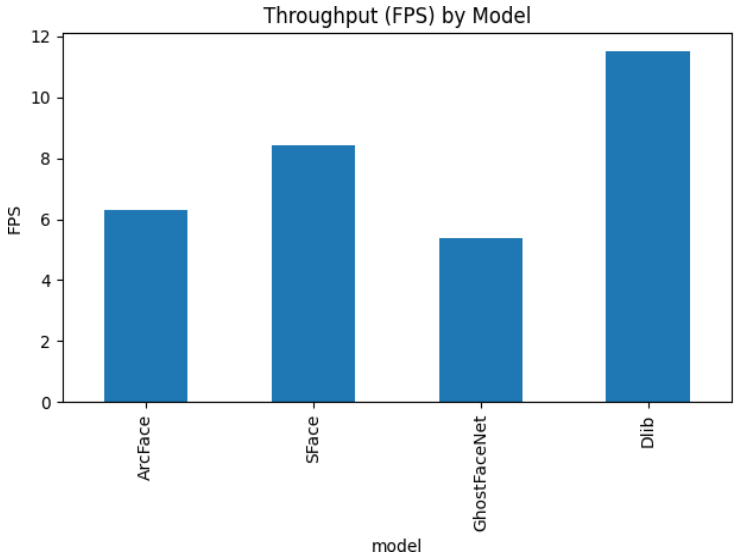

Fig. 4: Throughput (FPS) Comparison.

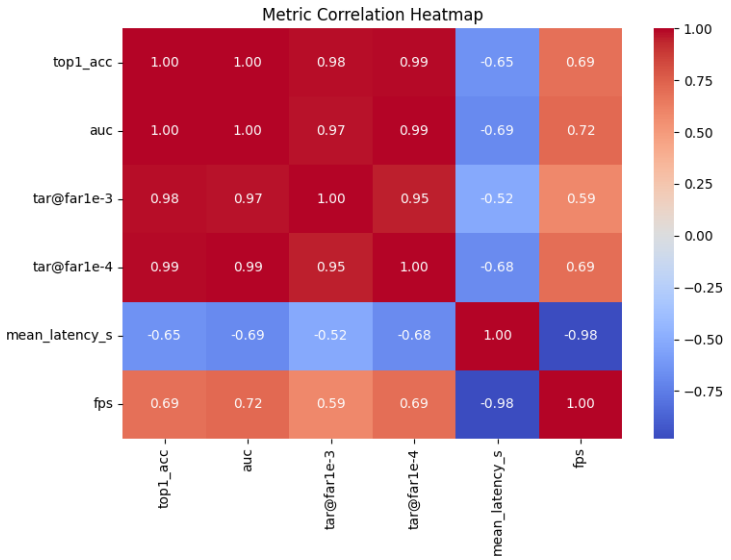

Fig. 5: Correlation Heatmap Between Metrics

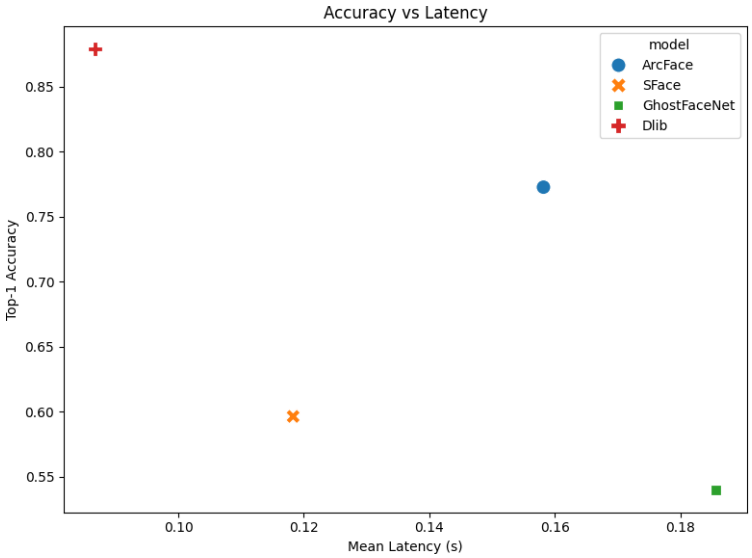

Fig. 6: Accuracy vs Latency Trade-off.

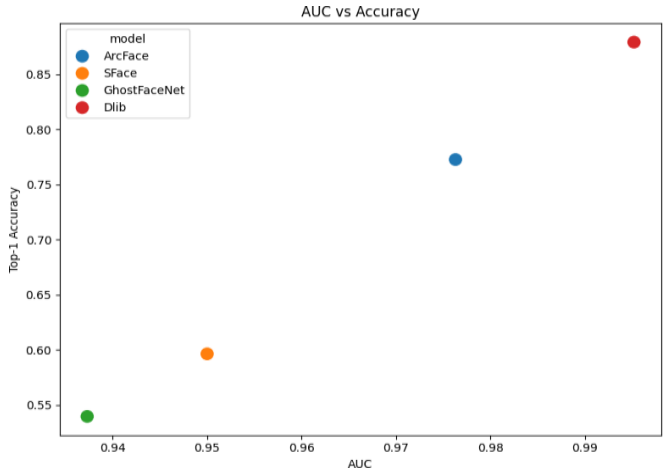

Fig. 7: AUC vs Accuracy Correlation.

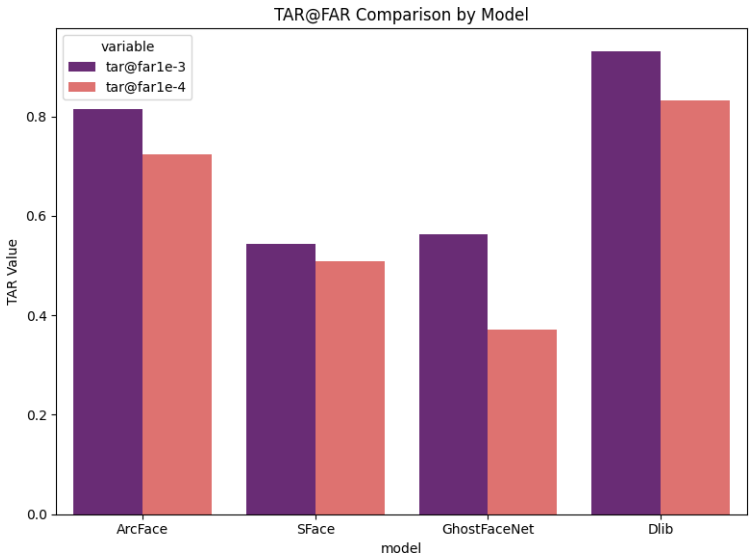

Fig. 8: TAR@FAR Comparison Across Models

Table 2: Ranked Summary of Efficiency and Weighted Scores

Model	**Top-1 AUC**	**Latency (s)**	**Efficiency**	Weighted Score
Dlib	0.8793 0.9952	0.0867	10.14	1.000
ArcFace	0.7727 0.9763	0.1582	4.88	0.632
SFace	0.5963 0.9500	0.1183	5.04	0.182
GhostFace Net	0.5395 0.9373	0.1857	2.90	0.010

6 Discussion

The analysis reveals a clear trade-off among the four backbones:

- **Dlib** achieved superior verification metrics and speed, validating its robust- ness for moderate datasets.
- **ArcFace** displayed strong discriminative characteristics but suffered from higher latency due to its deeper angular-margin network.
- **SFace** offered balanced runtime performance with moderately lower accu- racy.
- **GhostFaceNet** prioritized lightweight efficiency but showed reduced dis- criminative power.

Failure cases included visually similar subjects, low-light samples, and indi- viduals with few reference images. Dataset imbalance and environmental noise contributed to reduced performance in verification tails. These factors highlight the need for quality-aware scoring, augmentation, and robust threshold calibration.

7 Conclusion and Future Work

This study presented a comprehensive benchmark of four face recognition back- bones—ArcFace, SFace, GhostFaceNet, and Dlib—for automated attendance systems. Among the evaluated models, Dlib achieved the highest identification accuracy (87.93%), highest AUC (0.9952), and the fastest throughput (11.53 FPS). ArcFace demonstrated strong discriminability, whereas SFace and Ghost- FaceNet provided improvements in efficiency but with lower overall accuracy.

Limitations :

- The dataset, though large, contains class imbalance and limited augmenta- tion of challenging scenarios (occlusion, extreme illumination).
- Only four backbones were tested; emerging adaptive-margin and transformer- based architectures were not included.
- No liveness or spoof detection was integrated, which limits real-world secu- rity.

Future Work : The benchmark can be extended by :

- Incorporating adaptive-margin models such as AdaFace, ElasticFace, and MagFace.
- Integrating liveness detection and anti-spoofing mechanisms.
- Exploring multimodal biometric fusion (face + voice + RFID).
- Implementing privacy-preserving face recognition using federated learning.
- Deploying models on edge-optimized pipelines using ONNX/TensorRT and quantization for real-time classrooms.
- Adding temporal smoothing and tracking for video-based attendance.

Acknowledgment

The author thanks Dr. Prabhu Kavin B for his invaluable mentorship and SRM Institute of Science and Technology for supporting this work.

References

1. M. Kim, A. K. Jain, and X. Liu, “AdaFace: Quality Adaptive Margin for Face Recognition,” in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recog- nition (CVPR)*, 2022.
2. F. Boutros, M. Klemt, A. Kuijper, and N. Damer, “ElasticFace: Elastic Margin Loss for Deep Face Recognition,” in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022.
3. Y. Zhong, J. Deng, and S. Zafeiriou, “SFace: Sigmoid-Constrained Hypersphere Loss for Robust Face Recognition,” *arXiv preprint arXiv:2205.12010*, 2022.
4. GhostFaceNet Team, “GhostFaceNet: Lightweight Face Recognition via Ghost Modules,” *IEEE Access*, 2023.
5. S. I. Serengil and A. Özpinar, “DeepFace: A Lightweight Face Recognition Framework for Python,” *arXiv preprint arXiv:2303.14872*, 2023.