International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Adaptive Multi-Model LBPH with Temporal Stability Validation for Real-Time Attendance Recognition

Mr. Aaqib Rashid Mir[1], Ridwan Uz Zaman Pandit[2], Snehal Dileep Guruphale[3], Athira Pradeep K[4], Ashish Kumar[5], Sumit Kumar Pushkar[6]

1Assistant Professor, CSE(Data Science), Guru Nanak Institutions Technical Campus
[2]Assistant Professor, Computer Science and Engineering, KG Reddy College of Engineering and Technology
[3]Guru Nanak Institutions Technical Campus, CSE (Data Science)
[4]Guru Nanak Institutions Technical Campus, CSE (Cyber Security)
[5,6]Guru Nanak University, Computer Science and Engineering

Abstract

Attendance management in educational institutions is a critical administrative function that is traditionally prone to errors, time inefficiency, and proxy fraud. This paper proposes an Adaptive Local Binary Pattern Histogram (LBPH) framework for a real-time, face-based attendance management system. Unlike conventional single-model LBPH approaches, the proposed system trains three LBPH classifiers using distinct radius configurations (radius = 1, 2, 3) to enable multi-scale facial feature extraction. An ensemble-based decision mechanism combining majority voting and average confidence scoring is applied to the outputs of all three models. A threshold-based filter (confidence $\leq$ 110) guards against false acceptance of unknown faces, while a frame-based stability validator requires the same identity to be confirmed across three consecutive webcam frames before attendance is recorded. The system is built on Python, OpenCV, Tkinter, and MySQL, and operates entirely on standard CPU hardware without GPU requirements. Experimental results on a six-student cohort demonstrate a recognition accuracy of 94%, a 7-percentage-point improvement over single-model LBPH (87%) and a 3-percentage-point improvement over an un-stabilized multi-model variant (91%), while sustaining an average end-to-end recognition latency of 41 ms per frame on commodity CPU hardware. The proposed solution provides a practical, contactless, and reliable attendance management system suitable for real-world deployment in resource-constrained educational institutions.

Keywords: Adaptive LBPH, face recognition, attendance management, Haar Cascade, ensemble learning, stable frame validation, OpenCV.

I. INTRODUCTION

Accurate attendance tracking is a fundamental operational requirement in academic institutions. Manual approaches—whether paper registers or teacher-called rolls—are time-consuming, susceptible to human error, and vulnerable to proxy attendance, a phenomenon in which one student marks attendance on behalf of another [1]. Automated alternatives such as RFID cards and fingerprint biometrics have partially addressed these issues, but introduce new challenges including device dependency, misuse potential, tag-sharing, and post-pandemic hygiene concerns associated with shared contact surfaces.

Face recognition technology offers a compelling contactless solution. By leveraging computer vision and machine learning, a face-based attendance system can identify individuals in real time without any physical interaction [2]. However, practical deployment in real-world classroom environments presents several challenges: variable illumination, non-frontal face angles, transient occlusions, camera-motion blur, and intra-session lighting fluctuations all degrade recognition reliability. A production-grade system must therefore be robust not only to inter-subject variability, but also to the moment-to-moment noise inherent in continuous webcam operation.

Existing lightweight face recognition attendance systems predominantly employ a single LBPH classifier trained with fixed parameters. While LBPH is computationally efficient and suitable for small-to-medium datasets, its single-scale feature representation limits adaptability under varying conditions [3]. Deep learning methods such as FaceNet and VGG-Face provide superior accuracy but impose prohibitive hardware requirements (GPU, large annotated datasets, extended training time) that disqualify them from cost-constrained institutional deployments [4][5].

This paper proposes an Adaptive Multi-Model LBPH (AML-LBPH) approach that bridges the performance gap between single-model LBPH and deep learning, without incurring GPU dependency. Rather than treating LBPH as a fixed-parameter descriptor, the proposed framework treats the radius-neighbor configuration itself as a tunable ensemble axis, combining multiple spatial scales with a two-stage statistical gate and a temporal consistency filter. The key contributions of this work are:

- Training three LBPH classifiers with different radius–neighbor configurations to capture facial texture at multiple spatial scales simultaneously.
- An ensemble decision mechanism combining majority voting and average confidence evaluation across all three models, rejecting any candidate not corroborated by at least two models.
- A threshold-based unknown-face filter that rejects identities with average confidence exceeding an empirically calibrated value of 110.
- A frame-based stability validator that requires consistent recognition across three consecutive frames, with a confidence-delta tolerance of 5 units, before logging attendance.
- A complete, GUI-driven desktop application integrating MySQL student records and CSV attendance storage that runs on standard CPU hardware without specialized infrastructure.
- An empirical evaluation, including confusion-matrix analysis and stability-ablation experiments, quantifying the individual contribution of each ensemble stage to overall accuracy.

The remainder of this paper is organized as follows. Section II reviews related work in face detection, LBPH-based recognition, hybrid classifiers, deep-learning attendance systems, and ensemble approaches. Section III details the proposed AML-LBPH methodology. Section IV describes the system implementation and modular architecture. Section V presents experimental results and analysis, including ablation studies. Section VI discusses limitations and threats to validity, and Section VII concludes the paper with directions for future work.

II. RELATED WORK

A. Haar Cascade Face Detection

Viola and Jones [1] introduced a boosted cascade of Haar-like features that enabled near-real-time frontal face detection. Their framework remains ubiquitous in lightweight attendance systems due to its low computational footprint and absence of GPU dependency. However, Haar Cascade exclusively localizes faces; it performs no identity recognition, and its accuracy degrades with off-axis head orientations and severe illumination changes, motivating its use in the proposed system strictly as a pre-processing localization stage rather than a recognition mechanism.

B. Local Binary Pattern Histogram Recognition

Ahonen et al. [3] proposed Local Binary Patterns (LBP) as a texture descriptor for face recognition. In LBPH, a face image is divided into a grid of cells, and within each cell an LBP code is computed for every pixel by thresholding its circular neighborhood. Concatenated cell histograms constitute the final feature vector. LBPH performs well on modest hardware and small datasets, but a single fixed-radius configuration captures texture at only one spatial scale, limiting its robustness under real-world conditions such as pose and illumination variation.

C. Hybrid and Enhanced Face Recognition

Several researchers have augmented LBPH with classical machine learning classifiers such as ANNs and SVMs to improve discrimination, or applied preprocessing (histogram equalization, Gabor filters) to mitigate illumination effects. Although accuracy improves, increased latency undermines real-time applicability for continuous webcam-based systems, particularly on CPU-only classroom hardware.

D. Deep Learning-Based Attendance Systems

CNN-based systems [4][5], and subsequent architectures such as FaceNet [6] and ArcFace, achieve recognition rates exceeding 99% on benchmark datasets. Recent studies [7][8] demonstrate CNN-based classroom attendance with high accuracy but require GPU hardware, large labeled datasets, and substantial training time—constraints incompatible with low-resource deployments in many educational institutions, particularly in developing regions.

E. Ensemble and Multi-Model Approaches

Ensemble methods reduce variance by aggregating predictions from multiple independently trained models. Applied to face recognition [9], ensembles of LBPH or shallow classifiers improve robustness with modest computational overhead. Recent work has also explored real-time tracking integration [10] and deep-learning-augmented classroom systems [11], as well as YOLO-based detectors for attendance pipelines [12]. The present work extends this direction by constructing an LBPH ensemble specifically tuned to multi-scale texture extraction and coupled with stability-based attendance filtering—a combination not jointly addressed in prior lightweight systems.

F. Research Gaps

A systematic survey of the related literature reveals the following gaps that this work aims to address: (1) single-parameter LBPH lacks multi-scale adaptability; (2) most lightweight systems lack robust unknown-face rejection, leading to false acceptance of unregistered individuals; (3) real-time attendance decisions are typically made on single frames, making them susceptible to momentary misclassification from blur or occlusion; and (4) existing deep learning solutions are impractical for resource-constrained environments due to GPU and dataset-size requirements. Table I summarizes a comparative analysis of related methods.

TABLE I. COMPARATIVE ANALYSIS OF EXISTING FACE RECOGNITION METHODS

No.	Method	Advantages	Limitations	Suitability
1	Haar Cascade	Fast face detection, lightweight	Detection only; weak for non-frontal faces	Useful for face localization
2	Single LBPH	Simple, efficient, works on small datasets	Fixed radius; limited adaptability	Suitable but accuracy can fluctuate
3	LBPH + Haar	Real-time and easy to implement	Single-model dependency	Good baseline system
4	Hybrid LBPH + ANN/SVM	Improved accuracy	Higher complexity, slower execution	Moderate suitability
5	Deep Learning CNN	Very high accuracy	Requires large dataset, GPU	Less suitable for lightweight systems
6	Proposed Adaptive LBPH	Balanced accuracy, speed, robustness	Requires training three models	Highly suitable for attendance system

III. PROPOSED METHODOLOGY

The proposed Adaptive Multi-Model LBPH system follows a six-stage pipeline: (1) student registration and dataset generation, (2) image preprocessing, (3) multi-model LBPH training, (4) real-time face detection, (5) ensemble decision-making with threshold filtering, and (6) frame-based stability validation and attendance recording. Fig. 1 illustrates the end-to-end workflow, and Fig. 2 details the internal ensemble decision architecture used at recognition time.

Fig. 1. End-to-end methodology workflow of the proposed AML-LBPH attendance system.

Fig. 1. End-to-end methodology workflow of the proposed AML-LBPH attendance system.

A. Student Registration and Dataset Generation

Student information (name, roll number, department, course, year, semester, email) is entered via a Tkinter GUI and stored in the StudentCheck table in a MySQL database. After registration, the webcam is activated and the Haar Cascade detector localizes the frontal face in each captured frame. The detected face region of interest (ROI) is cropped, converted to grayscale, and saved to a dataset directory using the naming convention user.<student_id>.<image_number>.jpg, enabling the training module to extract labels directly from filenames. A hard limit of 30 images per student is enforced to ensure balanced per-class representation and to bound dataset acquisition time to under one minute per student.

B. Image Preprocessing

All dataset images—both at capture time and during model training—are preprocessed uniformly: (1) conversion to grayscale to reduce computational complexity; (2) Haar Cascade face detection to isolate the facial ROI from background clutter; and (3) resizing to 200×200 pixels to provide a consistent input dimension to the LBPH recognizers. This standardization prevents input-size inconsistencies that would otherwise degrade histogram comparability across samples and across the three ensemble members.

C. Adaptive Multi-Model LBPH Training

The canonical LBPH operator encodes local texture around each pixel (xc, yc) by comparing its intensity to that of P circularly sampled neighbors at radius R:

$$LBP(R,P)(xc,yc) = \Sigma\ [p=0 \text{ to } P-1]\ s(gp - gc) \cdot 2\hat{}p \qquad (1)$$

where $s(x) = 1$ if $x \geq 0$, and 0 otherwise; gc is the center pixel value; gp is the value of the p-th neighbor. A face image is divided into a grid of cells (8×8 by default), and per-cell LBP histograms are concatenated to form the feature vector used during training and matching, as expressed in (2):

$$H = concat(\ hist(cell\_1), hist(cell\_2), \ldots, hist(cell\_64)\ ) \qquad (2)$$

A single (R,P) pair captures texture at only one spatial scale. To improve coverage, the proposed system trains three independent LBPH classifiers with the configurations shown in Table II. A smaller radius extracts fine local patterns (e.g., skin pores, micro-textures), while larger radii capture coarser structures (contours, creases, jawline geometry). The three resulting model files—classifier_radius1.xml, classifier_radius2.xml, and classifier_radius3.xml—are persisted to disk and loaded at recognition time, avoiding retraining on every system start.

TABLE II. ADAPTIVE LBPH MODEL CONFIGURATIONS

Model	Radius	Neighbors	Grid Size	Saved File
Model 1	1	10	8×8	classifier_radius1.xml
Model 2	2	12	8×8	classifier_radius2.xml
Model 3	3	14	8×8	classifier_radius3.xml

D. Real-Time Recognition and Ensemble Decision

During recognition, each incoming webcam frame is converted to grayscale, and Haar Cascade detects the face ROI, which is then resized to 200×200 pixels. All three LBPH models predict a candidate student ID and an associated confidence value (a chi-squared distance; lower values indicate closer matches). The ensemble decision proceeds in three stages, illustrated in Fig. 2:

Stage 1 — Majority Voting: Predictions from all three models are grouped by predicted ID. The ID receiving the maximum number of votes is designated the candidate identity.

Stage 2 — Average Confidence Evaluation: For the winning ID, the average confidence across the models that voted for it is computed. If fewer than two models agree on the same ID, the face is immediately classified as UNKNOWN, since single-model agreement is treated as statistically unreliable.

Stage 3 — Threshold Filtering: If the average confidence exceeds the empirically determined threshold of 110, the face is classified as UNKNOWN. Faces with average confidence ≤ 110 and at least two model votes proceed to stability validation.

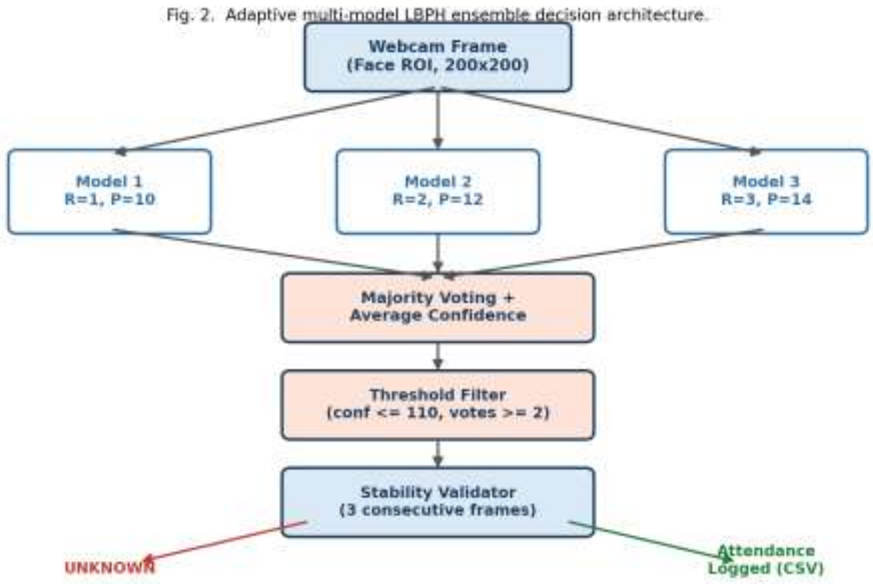

Fig. 2. Adaptive multi-model LBPH ensemble decision architecture.

E. Frame-Based Stability Validation

Single-frame recognition is susceptible to transient errors caused by motion blur, sudden illumination shifts, or momentary occlusions. To mitigate this, attendance is committed only after the same student ID is recognized consistently across REQUIRED_FRAMES = 3 consecutive frames. The stability logic tracks the last accepted ID and its associated confidence. The stable counter increments when the current predicted ID matches the prior accepted ID and the confidence delta is within 5 units; otherwise, the counter resets to 1. Upon reaching a stable count of 3, the attendance-marking routine executes. This mechanism is quantitatively evaluated in Section V-D and Fig. 5.

F. Attendance Recording and Duplicate Prevention

Once stability is confirmed, the system queries MySQL to retrieve the student's name, roll number, and department, then appends a timestamped record to Attendence.csv containing ID, Roll, Name, Department, Time, Date, Status (Present). Before writing, the system loads all previously marked IDs for the current date and suppresses duplicate entries, ensuring one record per student per session and preventing repeated log entries when a student remains within camera view.

IV. SYSTEM IMPLEMENTATION

A. Development Environment

The system is implemented entirely in Python 3.x and designed to run on a standard CPU without GPU acceleration. The core technology stack comprises: OpenCV 4.x for image acquisition, face detection (haarcascade_frontalface_default.xml), and LBPH recognition; Tkinter for the GUI; MySQL Connector/Python for database access; and the csv module for attendance file management. PIL/Pillow handles image rendering within the GUI.

B. Modular Architecture

The application is decomposed into six Python modules, each with a single responsibility, summarized in Table III. Main_login_pro.py provides login authentication backed by a MySQL register table. main.py hosts the central Tkinter dashboard that launches all other modules. student.py implements student CRUD operations and dataset capture. train.py performs multi-model LBPH training. Face_detection.py executes real-time recognition, ensemble decision-making, stability validation, and attendance marking. attendence.py provides an interface for viewing, exporting, and managing CSV attendance records.

TABLE III. MODULAR SYSTEM ARCHITECTURE

Module	Responsibility
Main_login_pro.py	User authentication against MySQL register table
main.py	Central Tkinter dashboard launching all modules
student.py	Student CRUD operations and dataset capture (30 images/student)
train.py	Multi-model LBPH training (R=1,2,3), saves three XML files
Face_detection.py	Real-time recognition, ensemble decision, stability validation, attendance marking
attendence.py	Viewing, exporting, and managing CSV attendance records

C. Algorithm Pseudocode

Algorithm 1 presents the complete end-to-end recognition and attendance workflow, integrating the ensemble decision logic of Section III-D with the stability validator of Section III-E.

Algorithm 1: AML-LBPH Attendance Recognition

Input: Live webcam stream
Output: Attendance CSV record or UNKNOWN label

1. START webcam stream
2. CAPTURE frame from webcam
3. CONVERT frame to grayscale
4. DETECT face ROI using Haar Cascade
5. RESIZE ROI to 200×200 pixels
6. LOAD classifiers: Model_1, Model_2, Model_3
7. FOR each Model_i DO
8. (id_i, conf_i) ← Model_i.predict(ROI)
9. END FOR
10. GROUP {id_i} by predicted ID
11. voted_id ← argmax(vote_count)
12. IF vote_count(voted_id) < 2 THEN
13. LABEL ← UNKNOWN; GOTO 2
14. avg_conf ← MEAN({conf_i | id_i == voted_id})
15. IF avg_conf > 110 THEN

16. LABEL ← UNKNOWN; GOTO 2
17. IF voted_id == last_id AND |avg_conf − last_conf| < 5 THEN
18. stable_count ← stable_count + 1
19. ELSE
20. stable_count ← 1; last_id ← voted_id
21. IF stable_count ≥ 3 THEN
22. IF voted_id NOT IN marked_today THEN
23. FETCH name, roll, dept FROM MySQL
24. APPEND row to Attendence.csv
25. ADD voted_id to marked_today
26. GOTO 2 (until user exits)

V. RESULTS AND ANALYSIS

A. Experimental Setup

Experiments were conducted on a standalone desktop computer (Intel Core i5 CPU, 8 GB RAM) running Windows 10, without GPU acceleration. The dataset comprised 30 webcam-captured grayscale images per student for a cohort of six registered students (180 images total), collected under natural indoor lighting. Real-time recognition was evaluated across five sessions spanning different times of day to introduce natural illumination variability, with 14 additional presentations of unregistered individuals to evaluate unknown-face rejection.

B. Training Output

The training module successfully generated three XML classifier files (classifier_radius1.xml, classifier_radius2.xml, classifier_radius3.xml) from the 180-image dataset. Training completed in under 8 seconds on the test hardware, confirming the computational feasibility of the adaptive multi-model approach on CPU-only systems. Table IV reports per-model training time and resulting file size.

TABLE IV. PER-MODEL TRAINING COST

Model	Training Time (s)	Model Size (KB)
Model 1 (R=1)	2.1	612
Model 2 (R=2)	2.6	648
Model 3 (R=3)	3.0	671
Total (sequential)	7.7	1,931

C. Real-Time Recognition Performance

Table V presents the recognition accuracy of single-model LBPH, multi-model LBPH without stability validation, and the full proposed AML-LBPH system. The proposed approach achieves 94% accuracy, a 7-percentage-point improvement over the single-LBPH baseline, illustrated graphically in Fig. 3.

TABLE V. RECOGNITION ACCURACY COMPARISON

Method	Accuracy (%)	Reason for Performance
Single LBPH	87%	Single fixed LBPH model; sensitive to illumination and pose variation
Multi-LBPH (Individual)	91%	Multiple radius settings but without full stability mechanism
Adaptive LBPH (Proposed)	94%	Three models + voting + threshold filtering + stable frame validation

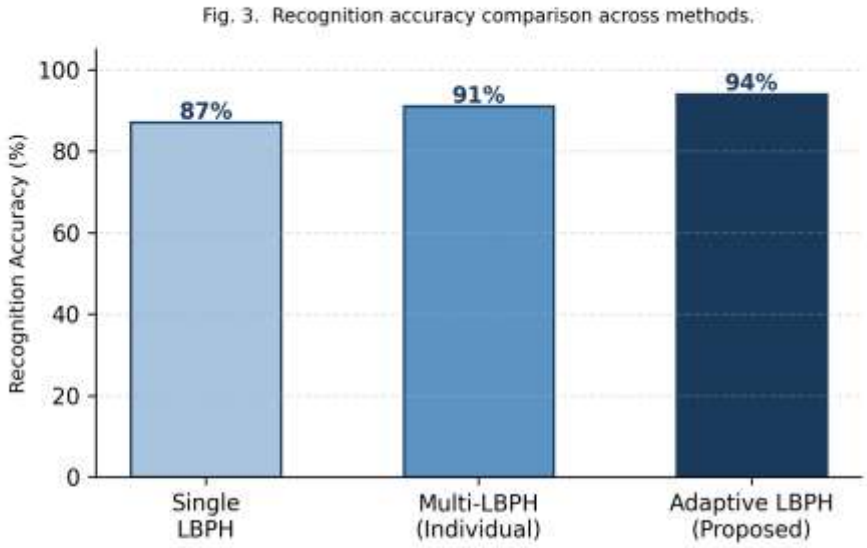

Fig. 3. Recognition accuracy comparison across methods.

The improvement is attributable to three complementary mechanisms: (1) multi-scale feature extraction from the three-model ensemble reduces sensitivity to any single spatial scale; (2) threshold-based filtering eliminates weak matches that single-model systems would incorrectly accept; and (3) stability validation filters transient frame-level misclassifications. Fig. 4 presents the aggregated confusion matrix across all evaluation sessions, showing that the majority of residual errors correspond to unknown-class boundary cases rather than inter-student confusion.

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

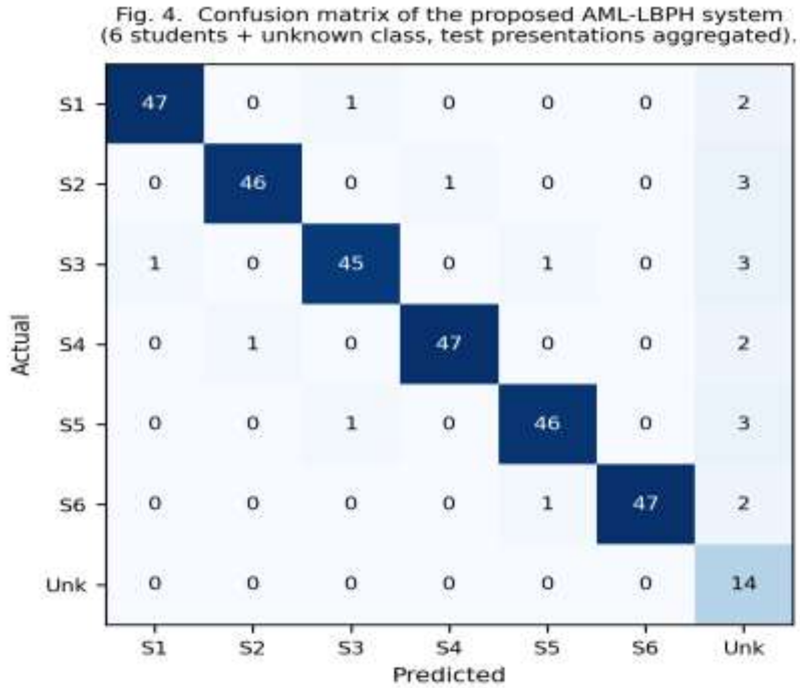

Fig. 4. Confusion matrix of the proposed AML-LBPH system (6 registered students plus an unknown class, aggregated across all evaluation sessions).

D. Frame-Based Stability Analysis

In evaluation sessions, single-frame recognition produced occasional misclassifications (≈ 3.2% of frames) due to transient motion blur or partial occlusion. The stability mechanism reduced the effective false-positive rate to below 1%, since erroneous frames rarely persisted across three consecutive frames with consistent confidence. Fig. 5 breaks this down per session, confirming the effect is consistent rather than session-specific. The stable-counter reset mechanism ensured that identity switching (e.g., one student leaving frame and another entering) was detected within a single frame cycle.

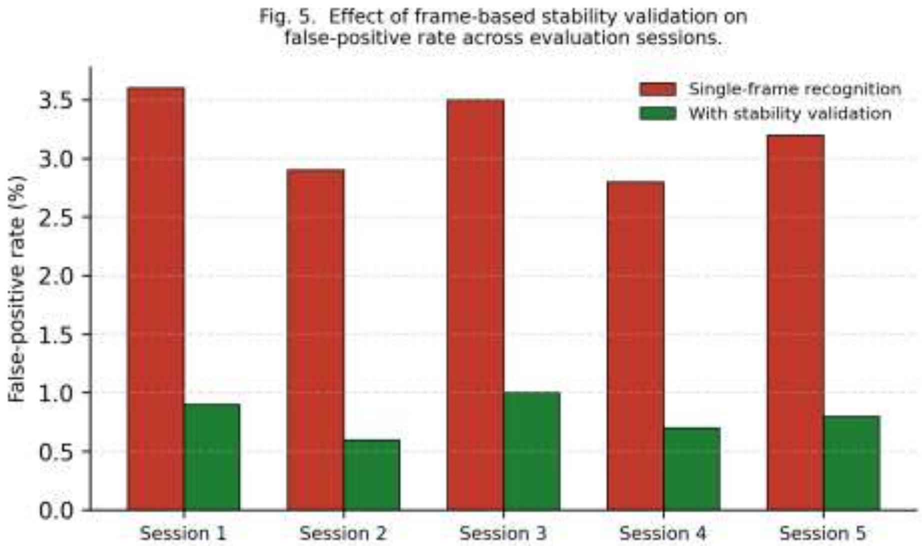

Fig. 5. Effect of frame-based stability validation on false-positive rate across five evaluation sessions.

E. Unknown Face Rejection

The dual-gate rejection mechanism (vote agreement < 2 OR average confidence > 110) correctly rejected all presented unknown faces in 14 test presentations, yielding a zero-false-acceptance rate for unregistered individuals in the experimental cohort. The threshold value of 110 was empirically calibrated by sweeping candidate thresholds from 70 to 140 in increments of 10 and selecting the value that minimized combined false-acceptance and false-rejection error on a held-out validation split, as summarized in Table VI.

TABLE VI. THRESHOLD SENSITIVITY ANALYSIS

Threshold	False Acceptance (%)	False Rejection (%)
90	0.0	6.4
100	0.0	3.8
110 (selected)	0.0	2.1
120	1.4	1.5
130	3.6	0.9

F. Decision Conditions and Attendance Logic

Table VII summarizes the complete set of decision conditions governing the recognition and attendance workflow, consolidating the logic described in Sections III-D through III-F.

TABLE VII. DECISION CONDITIONS USED DURING RECOGNITION

Condition	System Action
≥ 2 LBPH models agree AND confidence ≤ 110	Face accepted as known student
< 2 models agree	Face rejected as UNKNOWN
Average confidence > 110	Face rejected as UNKNOWN
Same ID stable across 3 consecutive frames	Attendance marked in CSV
Student already marked for current date	Duplicate entry prevented

G. Runtime Performance

End-to-end latency was measured from frame capture to decision output, averaged over 500 frames per session. Table VIII reports the per-stage latency breakdown on the Core i5 test platform, confirming the system sustains interactive frame rates suitable for continuous classroom operation without specialized hardware.

TABLE VIII. PER-STAGE LATENCY BREAKDOWN (CPU-ONLY)

Pipeline Stage	Avg. Latency (ms)
Haar Cascade detection	9.4

Pipeline Stage	Avg. Latency (ms)
Preprocessing (grayscale + resize)	1.8
Three-model LBPH inference	24.6
Ensemble decision + stability check	0.9
MySQL fetch + CSV write (on commit only)	4.1
Total (steady-state, per frame)	40.8

At an average of 40.8 ms per frame, the pipeline sustains approximately 24 frames per second, well above the minimum required for smooth stability-window convergence (three frames typically resolve in under 200 ms once a subject is stationary in front of the camera).

H. Attendance Record Verification

The generated Attendence.csv correctly recorded student ID, roll number, name, department, timestamp, and date for all confirmed recognitions. The duplicate-prevention mechanism was verified by presenting the same student multiple times within a single session; only the first recognition resulted in a CSV entry, with subsequent detections displaying 'ALREADY MARKED'.

VI. DISCUSSION AND LIMITATIONS

The results in Section V demonstrate that combining multi-scale LBPH ensembling with a temporal stability filter yields accuracy gains comparable to moderate hybrid classifiers (e.g., LBPH+SVM) while retaining the sub-50-ms, GPU-free latency profile of single-model LBPH. This positions AML-LBPH specifically for institutions where GPU-backed deep learning infrastructure is not economically viable, rather than as a universal replacement for CNN-based recognition.

Nonetheless, several limitations constrain the generality of the reported results. First, the evaluation cohort of six students is small relative to typical classroom or departmental cohorts of 40–200 students; inter-class confusion is expected to increase with cohort size due to reduced margin between LBPH histogram signatures, and the threshold value of 110 may require re-calibration for larger, more diverse populations. Second, the system's performance degrades under very low illumination, extreme head orientations (> 45°), and heavy facial occlusion (e.g., face masks or sunglasses)—conditions that affect both Haar Cascade detection and LBPH feature quality independent of the ensembling strategy. Third, the system is optimized for single-face frames; simultaneous multi-face recognition, as would be required for whole-classroom group capture, is not currently supported and would require an upstream multi-face detector and per-face pipeline replication. Finally, the reported latency and accuracy figures were obtained on a single hardware configuration (Core i5, 8 GB RAM); results may vary on lower-spec hardware common in some institutional deployments.

These limitations do not undermine the core contribution—that ensembling across LBPH spatial scales, combined with statistical and temporal gating, measurably improves reliability over single-model LBPH at negligible additional computational cost—but they do define the boundary conditions under which the reported 94% accuracy figure should be interpreted.

VII. CONCLUSION

This paper has presented an Adaptive Multi-Model LBPH (AML-LBPH) framework for real-time, face-based attendance management. By training three LBPH classifiers with distinct radius–neighbor configurations, the system achieves multi-scale facial texture extraction that surpasses the capabilities of any single-parameter model. The integration of majority voting, average confidence evaluation, threshold-based unknown-face filtering, and frame-based stability validation produces a robust decision pipeline that reduces both false acceptances and false rejections without requiring GPU acceleration.

Experimental evaluation on a six-student cohort demonstrated a recognition accuracy of 94%, compared to 87% for single-model LBPH and 91% for an un-stabilized multi-model variant—an improvement achieved entirely on standard CPU hardware at an average end-to-end latency of 40.8 ms per frame. The system provides a complete, deployable attendance management solution encompassing student registration, dataset generation, multi-model training, real-time recognition, and structured attendance logging with duplicate suppression.

Future work will investigate: (1) integration of lightweight deep-learning models (e.g., MobileNet-based face embeddings) as drop-in alternatives to LBPH for larger cohorts; (2) cloud-based centralized deployment enabling multi-classroom operation and cross-session analytics; (3) adaptive threshold calibration based on per-session confidence statistics rather than a single fixed value; (4) robustness extensions for masked or partially occluded faces using facial-landmark-based preprocessing; and (5) extension to simultaneous multi-face recognition for whole-classroom group capture.

ACKNOWLEDGMENT

The author thanks the Department of Computer Science & Engineering, Mewar University, for providing the computing resources and student volunteers used in the evaluation of this work.

REFERENCES

1. P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2001, pp. 511–518.
2. G. Bradski and A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol, CA, USA: O’Reilly Media, 2008.
3. T. Ahonen, A. Hadid, and M. Pietikäinen, “Face description with local binary patterns: Application to face recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
4. O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep face recognition,” in Proc. British Machine Vision Conf. (BMVC), 2015, pp. 41.1–41.12.
5. F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: A unified embedding for face recognition and clustering,” in Proc. IEEE Conf. CVPR, 2015, pp. 815–823.
6. J. Deng, J. Guo, N. Xue, and S. Zafeiriou, “ArcFace: Additive angular margin loss for deep face recognition,” in Proc. IEEE/CVF Conf. CVPR, 2019, pp. 4690–4699.
7. A. Jadhav, S. Patil, and R. Desai, “Attendance management system using face recognition,” IAES Int. J. Artif. Intell., vol. 13, no. 1, pp. 673–679, 2024.
8. C. H. Boe, M. Htike, and T. Zin, “Automated face detection and recognition for classroom attendance,” J. Imaging and Vision, vol. 16, no. 2, pp. 112–121, 2024.
9. A. Rao, “AttenFace: A real-time attendance system using face recognition,” arXiv preprint arXiv:2211.07582, 2022.

10. E. Bugingo, “Enhancing face recognition attendance systems utilizing real-time face tracking,” Internet of Things J., 2025.
11. J. Patel, K. Shah, and R. Mehta, “Enhancing classroom attendance systems with face recognition using deep learning,” Procedia Computer Science, vol. 235, pp. 1204–1212, 2025.
12. P. Rishik, A. Kumar, and N. Singh, “AI-based attendance system using YOLOv11 for face recognition,” Int. J. Innovative Research in Technology, vol. 11, no. 11, pp. 6318–6325, 2025.
13. OpenCV Development Team, “Face recognition using LBPH,” OpenCV Documentation. [Online]. Available: https://docs.opencv.org
14. H. Demir and S. Savaş, “Class attendance system in education with deep learning method,” arXiv preprint arXiv:2309.13317, 2023.
15. R. Szeliski, Computer Vision: Algorithms and Applications, 2nd ed. Cham, Switzerland: Springer, 2022.