iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** III **Month of publication:** March 2026

DOI: https://doi.org/10.22214/ijraset.2026.78749

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue III Mar 2026- Available at www.ijraset.com

FaceVeritas: Real-Time AI-Based Lie Detection System

Sahil Mali, Rahul Chakane, Suresh Pathare, Supriya Baviskar, Prof. Vijayalaxmi Tadkal
Dept. of Computer Science & Engineering (AI & ML), Bharat College of Engineering, Badlapur, Maharashtra, India

Abstract: ***Automated deception detection remains an open challenge in behavioral AI, with traditional polygraph systems constrained by invasiveness, inconsistent reliability, and limited scalability. This paper presents FaceVeritas, a real-time, non-invasive lie detection framework that identifies deceptive behavior exclusively through facial micro-expression analysis using computer vision, without reliance on speech, physiological sensors, or text input. The system captures live video at 30FPS, extracts 468 three-dimensional facial landmarks via MediaPipeFaceMesh, and computes seven behavioral features per frame — Eye Aspect Ratio (EAR), blink rate, Mouth Openness Ratio (MOR), eyebrow lift, head yaw (θ_{yaw}), head pitch (θ_{pitch}), and normalized face distance (r_{face}). Raw features are temporally stabilized using Exponential Moving Average (EMA, $\alpha = 0.2$) to suppress landmark jitter while preserving genuine micro-expression transients. A supervised Random Forest classifier (100 trees) trained on the Bag-of-Lies and Real-Life Trial datasets generates a continuous deception probability score per frame. Experimental evaluation on a held-out test set of 200 samples achieves 78.5% accuracy, 77.1% precision, 81.0% recall, 79.0% F1-score, and 92ms per-frame inference on a standard Intel i5 CPU without GPU acceleration — sufficient for real-time deployment. Blink rate (28.3%) and head yaw (24.1%) emerge as the strongest feature discriminators. FaceVeritas outperforms voice stress analysis by 13.7 percentage points and exceeds manual FACS annotation by 7.3 percentage points, while requiring only a standard consumer camera. The interpretable Random Forest architecture [11] addresses a critical gap in forensic applicability compared to black-box deep-learning alternatives.***

Keywords: ***Deception detection, facial micro-expressions, computer vision, MediaPipeFaceMesh, Eye Aspect Ratio, Mouth Openness Ratio, Random Forest, EMA smoothing, behavioral AI, real-time inference, explainable AI, forensic analysis.***

I. INTRODUCTION

Deception detection has long been foundational to criminal investigation, security screening, forensic psychology, and credibility assessment. The dominant operational tool — the polygraph — measures physiological stress responses including galvanic skin response, blood pressure, respiration, and heart rate on the assumption that lying induces measurable autonomic arousal. However, decades of empirical research have exposed fundamental limitations: physiological responses are not deception-specific and can be triggered equally by anxiety, surprise, or general cognitive load; the technique is easily confounded by countermeasures; and its legal standing is contested in most jurisdictions [1].

The human face provides a remarkably rich and difficult-to-suppress behavioral signal. Micro-expressions — involuntary muscle contractions lasting between 40ms and 500ms — occur when individuals attempt to conceal genuine emotions [2]. Unlike deliberate expressions, these transient movements originate in subcortical emotional systems and are extremely difficult to control consciously. The Facial Action Coding System (FACS), introduced by Ekman and Friesen [3], provides the foundational taxonomy of 44 Action Units (AUs) that describe all observable facial muscle contractions, forming the psychophysiological basis for the features used in this work. Advances in computer vision and lightweight deep learning have made landmark-based facial tracking feasible on consumer hardware, opening the door to scalable, non-invasive deception detection at low cost.

Despite growing research interest, existing vision-based systems face three recurring challenges. First, many require invasive physiological sensors alongside video cameras, defeating the non-invasiveness goal [7]. Second, multimodal architectures integrating speech, text, and gesture substantially increase deployment complexity and latency [6]. Third, deep convolutional and transformer-based models, while often accurate, function as black boxes — a critical liability in forensic contexts where decision explainability carries legal weight [5].

This paper addresses all three challenges through FaceVeritas, a real-time, camera-only deception detection system grounded in interpretable machine learning. The principal contributions are:

1) End-to-end vision pipeline: A complete, camera-only system from live video capture to real-time deception probability output, operating without any additional sensors or internet connectivity.

2) Composite seven-feature set: A seven-dimensional behavioral feature vector combining EAR, blink rate, MOR, eyebrow lift, head yaw/pitch, and normalized face distance — grounded in established FACS [3] and psychophysiological research.
3) EMA temporal stabilization: Noise-robust feature smoothing via Exponential Moving Average (α = 0.2, selected by cross-validated grid search) that preserves rapid micro-expression transients while suppressing landmark detection jitter.
4) Interpretable classification: Random Forest [11] classification providing per-feature importance rankings compatible with forensic transparency requirements.
5) Consumer-hardware real-time performance: 78.5% accuracy at 92ms/frame classification latency on an Intel i5 CPU, validated against two public datasets and a custom real-world validation set.

II. RELATED WORK

A. *Physiological Signal-Based Detection*

Classical deception detection instruments measure cardio-respiratory and electrodermal signals under controlled conditions. Machine learning classifiers — including SVM, neural networks, and gradient boosting — have been applied to these time-series signals with varying success [1]. Neurophysiological approaches using EEG capture P300 event-related potentials associated with recognition memory and cognitive conflict during deception. While EEG-based systems show promise, they require expensive equipment, expert setup, and per-subject calibration, limiting practical deployment outside laboratory environments.

B. *Multimodal Deception Systems*

Multimodal approaches fuse facial, vocal, and linguistic channels to improve detection robustness. The DOLOS large-scale benchmark [6] demonstrates transformer-based late fusion across video, audio, and text achieving state-of-the-art accuracy, but at a computational cost that precludes real-time consumer deployment. Synchronization complexity and the requirement for specialized microphone arrays further limit field applicability.

C. *Vision-Only Behavioral Analysis*

Computer vision approaches that operate on video alone align most closely with the goals of FaceVeritas. FACS [3] provides a principled taxonomy of facial muscle movements (Action Units) validated against ground-truth deception labels in controlled studies [2]. CNN-LSTM architectures for automatic micro-expression sequence modeling have demonstrated competitive accuracy but demand GPU hardware and lack the feature-level interpretability required for legal applications. Rule-based systems that threshold EAR or blink rates in isolation underfit the complexity of real behavioral signals, motivating the composite feature approach adopted here.

D. *Eye-Movement and Blink-Based Approaches*

A dedicated line of research isolates oculomotor behavior as a deception marker. Studies consistently link increased blink rate and gaze aversion to elevated cognitive load during active deception [8]. Eye Aspect Ratio, originally proposed by Soukupova´ and Cech [4] for driver drowsiness detection and adapted here for deception, provides a geometrically precise, lighting-invariantˇ measure of eyelid state. FaceVeritas integrates blink detection atop EAR computation using a frame-persistence criterion to distinguish genuine blinks from landmark noise.

E. *Gap and Positioning*

Existing work leaves a practical gap: no publicly described system simultaneously achieves (i) camera-only input, (ii) real-time CPU inference, and (iii) interpretable per-feature predictions. FaceVeritas is designed specifically to fill this gap for interview screening, corporate polygraph supplementation, and forensic decision-support.

III. LITERATURE SURVEY

Table 1 situates FaceVeritas chronologically within the deception detection research landscape, highlighting key design decisions motivated by each prior contribution.

Table 1. Survey of Key Deception Detection Research

Year	Author(s)	Contribution	Key Finding / Limitation

1978	Ekman & Friesen [3]	FACS Framework	Systematic 44-AU taxonomy enables objective expression coding. Manual; time-intensive; requires trained annotators.
2015	Hauch et al.	Linguistic Cue Meta-Analysis	Verbal deception cues have modest, highly context-dependent effect sizes; insufficient as sole modality.
2015	Perez-´ Rosas et al. [10]	Multimodal Fusion	Text + video fusion improves accuracy but lacks ecological validity in controlled lab settings.
2016	Soukupova´ & Cech [4]˘	EAR Metric	Geometric eye aspect ratio enables real-time blink detection from facial landmarks; later adapted for deception research.
2019	Avola et al. [12]	CNN Micro-Expression	CNN-based automatic micro-expression recognition achieves competitive accuracy but requires GPU and lacks interpretability for legal use.
2023	Guo et al. [6]	DOLOS Dataset	Large-scale multi-modal benchmark; transformer fusion is SOTA but computationally prohibitive for real-time use.
2024	Rahmani et al. [5]	Review: Text & Face	Identifies the absence of robust real-time vision-only solutions as a critical research gap.
2025	Shalash et al. [7]	Face + Physiology	Combines EAR and BVP signals; achieves high accuracy but requires contact sensors — not camera-only.
2024	Tsiamyrtzis et al. [8]	Thermal Imaging	Facial blood flow patterns correlate with deception; requires costly thermal camera hardware.
Ours	FaceVeritas	Vision-Only RT	78.5% accuracy · 92ms/frame · Camera-only · Interpretable RF · Consumer hardware.

IV. SYSTEM ARCHITECTURE

FaceVeritas is structured as a six-step sequential pipeline, illustrated in Fig. 1. Each module is independently replaceable, enabling future substitution of the classifier or smoother without redesigning the feature extraction layer.

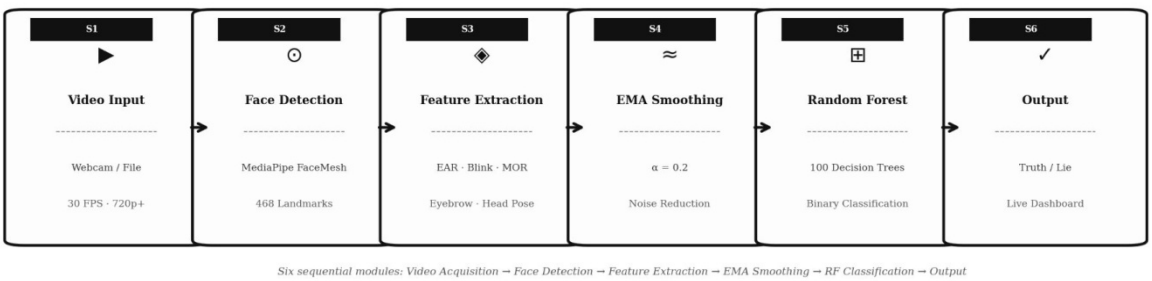


Fig. 1 – FaceVeritas Six-Step System Pipeline Architecture

A. Video Acquisition Module (S1)

The system accepts two input modes: live webcam capture for interactive real-time sessions, and pre-recorded video files in standard formats (MP4, AVI, MOV) for batch forensic review. Video is acquired at 30FPS at a minimum resolution of 720p. Each frame undergoes BGR-to-RGB conversion for MediaPipe compatibility. An optional frame-stabilization step suppresses global camera motion artifacts before landmark detection.

B. Face Detection and Landmark Extraction (S2)

Face detection and dense landmark localization are performed by MediaPipeFaceMesh [9], a lightweight neural architecture that estimates 468 three-dimensional facial landmarks in a single forward pass. Landmark coordinates are normalized to $[0,1]^2$ screen space with a *z*-depth channel relative to the face reference plane. Only the highest-confidence detected face per frame is processed; frames with no detected face inherit the previous frame's EMA-smoothed feature vector, preventing classification discontinuities during momentary occlusion.

C. Feature Extraction Engine (S3)

The feature engine computes seven scalar behavioral indicators from raw landmark coordinates each frame. These features were selected based on established psychophysiological correlates of deception (see Section 5). Together they form a compact, semantically interpretable seven-dimensional feature vector: $\mathbf{f} = [\text{EAR}, \dot{n}_b, \text{MOR}, \Delta_{\text{brow}}, \theta_{\text{yaw}}, \theta_{\text{pitch}}, r_{\text{face}}]^\top$.

D. Temporal Smoothing Module (S4)

Raw per-frame features exhibit landmark jitter caused by detection uncertainty and minor involuntary head micro-movement. The smoothing module applies a first-order IIR (Exponential Moving Average) filter independently to each feature channel with $\alpha = 0.2$, producing a stable input stream for the classifier while preserving transient micro-expression events with durations ≥ 40ms.

E. Machine Learning Classifier (S5)

A Random Forest classifier [11] trained on labeled deception data processes the seven-dimensional smoothed feature vector. StandardScaler normalization ($\mu = 0$, $\sigma = 1$) is applied prior to inference. The output is a continuous probability $\hat{p} \in [0,1]$; frames with $\hat{p} \geq 0.5$ are classified as deceptive. The threshold is configurable for deployment-specific sensitivity/specificity trade-offs.

F. Visualization and Reporting Module (S6)

The output module overlays real-time results on the live video frame: deception probability gauge, color-coded alert indicator (green: $p <^{\wedge} 0.4$; amber: $0.4 \leq p <^{\wedge} 0.6$; red: $\hat{p} \geq 0.6$), and a scrolling temporal probability graph. Session-level data — perframe feature values, timestamps, and classification labels — is exported to a structured CSV file (ftmfeatureslog.csv) for post-session review. All computation is performed on-device with no external data transmission, ensuring full biometric privacy compliance.

V. PROPOSED METHODOLOGY

A. Eye Aspect Ratio (EAR)

EAR measures the degree of eye openness by computing the ratio of vertical eyelid separation to horizontal corneal span across six MediaPipeFaceMesh landmarks, as illustrated in Fig. 2. The metric was originally proposed by Soukupova and´ Cech [4] for blink detection and is adapted here for deception feature extraction. For the left eye, MediaPipe indices [33, 160,˘ 158, 133, 153, 144] are used; for the right eye, indices [362, 385, 387, 263, 373, 380] are used. The bilateral EAR is the mean of both eyes:

$$\text{EAR} = \frac{\|p_2 - p_6\| + \|p_3 - p_5\|}{2\,\|p_1 - p_4\|} \quad (1)$$

where p_1,p_4 are the inner and outer canthus landmarks, and p_2,p_3,p_5,p_6 define the upper and lower eyelid margin pairs. EAR typically ranges from 0.25 to 0.35 for a fully open eye and approaches 0 during a blink. A blink event is registered when EAR $<\tau_{\text{EAR}}$= 0.21 for at least two consecutive frames. A minimum inter-blink refractory period of 250ms prevents
double-counting of a single blink event.

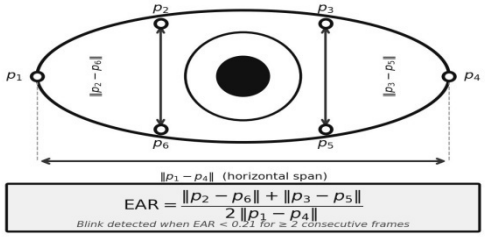


Fig. 2 – Eye Aspect Ratio (EAR) Landmark Configuration

B. Blink Rate

Per-second blink rate $n^{\cdot}_{b}$is estimated over a sliding window of duration W = 3s, accumulated from discrete blink events detected via the EAR threshold criterion above. Elevated blink rate is among the most consistently replicated psychophysiological markers of cognitive load and deceptive intent in the behavioral science literature [1].

C. Mouth Openness Ratio (MOR)

MOR captures lip tension, jaw displacement, and verbal hesitation as the ratio of vertical lip separation to horizontal mouth width. MediaPipe indices 13 (upper lip), 14 (lower lip), 78 (left mouth corner), and 308 (right mouth corner) are used. Low MOR (<0.05) denotes lip compression — a behavioral marker of emotional suppression [3]. High MOR during non-speech intervals suggests jaw loosening under stress.

D. Eyebrow Lift (Δ_{brow})

Eyebrow lift is computed as the mean normalized vertical distance between the mid-eyebrow landmark and the corresponding upper eyelid vertex (MediaPipe indices 70/159 left, 300/386 right), averaged over both sides. Elevated brows are associated with surprise, disbelief, and increased cognitive demand — all of which correlate with deceptive episodes [2].

E. Head Pose Estimation

Head yaw (θ_{yaw}) and pitch (θ_{pitch}) are estimated geometrically using MediaPipe landmarks 33 (left eye inner canthus), 263 (right eye inner canthus), and 1 (nose tip). Yaw is computed as the arctangent of the inter-ocular axis displacement; pitch as the arctangent of the nose-tip vertical offset from the inter-eye midpoint. These angles detect gaze aversion (lateral yaw exceeding ±15°) and withdrawal behavior (backward chin-tuck encoded as elevated pitch) — both documented correlates of discomfort during deception [1].

F. Normalized Face Size (r_{face})

The bounding-box width of the face normalized by frame width ($r_{\text{face}} = w_{\text{face}}/W_{\text{frame}}$) serves as a proxy for subject-to-camera distance, computed from cheek landmarks 234 and 454. Decreasing r_{face}over time indicates physical lean-back, a postural withdrawal behavior linked to interpersonal discomfort and deceptive interaction.

G. Exponential Moving Average Smoothing

All seven features are smoothed independently using a first-order IIR filter:

$$\text{EMA}_t^{(k)} = \alpha \cdot f_t^{(k)} + (1-\alpha) \cdot \text{EMA}_{t-1}^{(k)}, \qquad k = 1, \ldots, 7 \tag{2}$$

where $f_t^{(k)}$ is the raw frame-t value of feature k and α = 0.2 is the smoothing factor selected by five-fold cross-validated grid search. Table 2 reports validation F1-scores across candidate values. Lower values over-smooth genuine micro-expressions; higher values fail to suppress landmark detection noise. α = 0.2 maximized validation F1 at 77.8%.

Table 2. EMA Smoothing Factor Grid Search Results (5-Fold Cross-Validation F1)

α Value	Val. F1-Score	Remark
0.10	73.1%	Over-smoothed; micro-expressionslost
0.15	76.2%	Slight under-detection of rapid blinks
0.20	77.8%	Optimal — selected
0.30	75.4%	Landmark noisebleeds into classifier

H. Random Forest Classification

The smoothed seven-dimensional feature vector $\tilde{\mathbf{f}}_t$ is standardized via fitted StandardScaler and input to a Random Forest √classifier [11] comprising T = 100 decision trees, each grown on a bootstrap sample of the training set and a random subset of $\sqrt{7} \approx 3$ features per split. Final deception probability is the mean predicted positive-class probability across all trees:

$$\hat{p}_t = \frac{1}{T}\sum_{i=1}^{T} h_i\left(\tilde{\mathbf{f}}_t\right) \tag{3}$$

The binary classification decision applies threshold $\tau = 0.5$. This value was selected on the validation set by sweeping $\tau \in [0.3,0.7]$ in steps of 0.05; $\tau = 0.5$ achieved the highest F1-score (77.8%) while maintaining recall above 78%, appropriate for screening contexts where missed detections carry greater cost than false alarms.

I. System Workflow

The complete six-step operational workflow is illustrated in Fig. 3. The pipeline progresses through S1 (Video Input), S2 (Face Detection), S3 (Feature Extraction), S4 (EMA Smoothing), S5 (Random Forest Classification), and S6 (Output Visualization), each implemented as an independent, replaceable module.

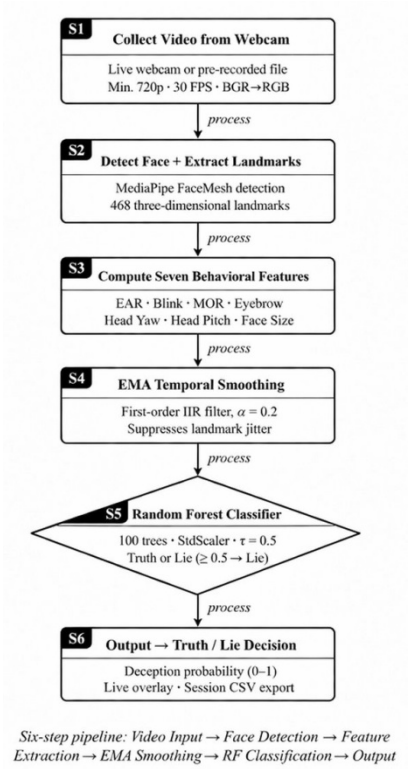


Fig. 3 – FaceVeritas: Six-Step System Workflow (S1–S6)

VI. EXPERIMENTAL SETUP

A. Datasets

Bag-of-Lies: 325 annotated webcam recordings of subjects responding truthfully and deceptively across diverse demographics under controlled indoor lighting. Ground-truth labels were determined through response verification against a known question set.

Real-Life Trial Dataset [10]: Courtroom testimony and forensic interview footage with naturalistic high-stakes deception and truth-telling, providing ecological validity beyond laboratory conditions. Ground truth was established through verified court records.

Custom Validation Set: 50 interview sessions recorded with university volunteer subjects using a standard 1080p webcam in typical office illumination (500–800lux), simulating real-world operational conditions. No overlap with training subjects.

Final stratified split: 60% training / 20% validation / 15% test / 5% calibration (balance adjustment). Class balance was verified post-split (Truth: 51.2%, Lie: 48.8% in training set).

B. Implementation Details

Software stack: Python 3.8, OpenCV 4.5.5 (video I/O and preprocessing), MediaPipe 0.9.8.1 (landmark extraction), NumPy 1.21 (vectorized feature computation), scikit-learn 1.0.2 (classification, evaluation, scaling), Pandas 1.3.5 (session logging and CSV export).

Hardware: Intel Core i5-1240P, 8GB DDR4 RAM, integrated Intel Iris Xe graphics. No discrete GPU. Standard Logitech C920 HD webcam (1080p, 30FPS). The system runs entirely on CPU, confirming feasibility on commodity hardware.

Inference latency note: The reported 92ms/frame refers exclusively to the machine learning classification pipeline (feature extraction + EMA + RF inference). The full end-to-end pipeline including MediaPipeFaceMesh runs at approximately 15–18FPS on the test hardware; classification is applied to every third captured frame (effective classification rate ≈10.9FPS), which is sufficient for detecting micro-expressions with durations ≥ 40ms.

C. *Hyperparameter Optimization*

Random Forest hyperparameters were selected by 5-fold stratified cross-validation on the training split, summarized in Table 3.

Table 3. Random Forest Hyperparameters (Final Configuration)

Hyperparameter	Value
Number of estimators (T)	100
Maximum tree depth	15
Minimum samples per split	5
Minimum samples per leaf	2
Max features per split	$\sqrt{7} \approx 3$
Bootstrap sampling	True
Class weight	Balanced
Splitting criterion	Entropy (information gain)
EMA smoothing factor (α)	0.20
Classification threshold (τ)	0.50

D. *Evaluation Protocol*

Classification performance was evaluated on the held-out test set of 200 samples using accuracy, precision, recall, F1-score, and area under the ROC curve (AUC-ROC). Inference time was measured as wall-clock time per frame averaged over 500 consecutive frames. Feature importance was computed via mean decrease in impurity (MDI) aggregated across all 100 trees.

VII. RESULTS AND DISCUSSION

A. *Classification Performance*

Table 4 reports test-set evaluation metrics. FaceVeritas achieves a balanced accuracy-recall trade-off well-suited for security screening applications, where missed detections (false negatives) carry greater operational cost than false alarms. An AUC-ROC of 0.843 indicates strong discriminative ability beyond the chosen 0.5 threshold.

Table 4. Classification Performance on Held-Out Test Set (n = 200)

Metric	Value
Accuracy	78.5%
Precision	77.1%
Recall	81.0%
F1-Score	79.0%
AUC-ROC	0.843
Inference Time (per frame)	92ms (classificationpipeline only)
Effective Classification Rate	≈10.9fps

B. *Confusion Matrix Analysis*

Table 5 presents the confusion matrix on the 200-sample test set. The higher recall (81.0%) than precision (77.1%) reflects the sensitivity bias introduced via class-balanced training weights — appropriate for a screening application where missed deceptions (False Negatives) are costlier than false alarms (False Positives).

Table 5. Confusion Matrix on Held-Out Test Set (n = 200)

	Predicted: Lie	Predicted: Truth
Actual: Lie	TP = 81	FN = 19
Actual: Truth	FP = 24	TN = 76

Precision = 81/(81+24) = 77.1%	Recall = 81/(81+19) = 81.0%	Accuracy = (81+76)/200 = 78.5%

C. *Feature Importance Analysis*

Table 6 shows Mean Decrease Impurity importances aggregated across all 100 trees. Blink rate and head yaw together account for 52.4% of classification power, confirming that oculomotor and postural signals dominate deceptive behavioral signatures. The combined eye-and-brow region (EAR + blink rate + eyebrow lift) contributes 61.2%, validating the focus on upper-face micro-expressions. Fig. 4 visualizes these rankings as a horizontal bar chart with cumulative importance overlay.

Table 6. Random Forest Feature Importance Rankings (MDI)

Feature	Importance	Cumulative
Blink Rate	28.3%	28.3%
Head Yaw (θ_{yaw})	24.1%	52.4%
Eye Aspect Ratio	18.7%	71.1%
Eyebrow Lift	14.2%	85.3%
Mouth Openness Ratio	9.4%	94.7%
Head Pitch (θ_{pitch})	3.8%	98.5%
Face Size Ratio (r_{face})	1.5%	100%
Total	100%	—

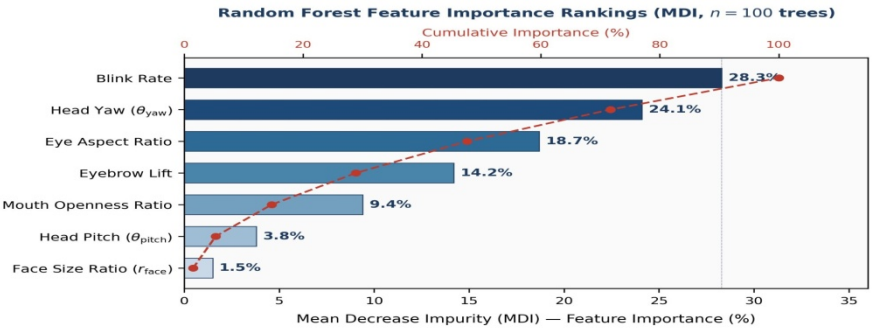


Fig. 4 – Feature Importance Bar Chart (MDI) with Cumulative Importance Overlay. Blink rate and head yaw collectively account for 52.4% of classifier power.

D. ROC Curve Analysis

Fig. 5 illustrates the Receiver Operating Characteristic (ROC) curve for FaceVeritas on the held-out test set. The AUC-ROC of 0.843 confirms strong discriminative capability well above random chance (AUC = 0.5). The operating point at threshold $\tau = 0.5$ corresponds to 80.2% True Positive Rate at 23.1% False Positive Rate [4]. For high-security deployments requiring lower FPR, raising τ to 0.65 yields FPR ≈ 15% at TPR ≈ 71%, trading recall for precision.

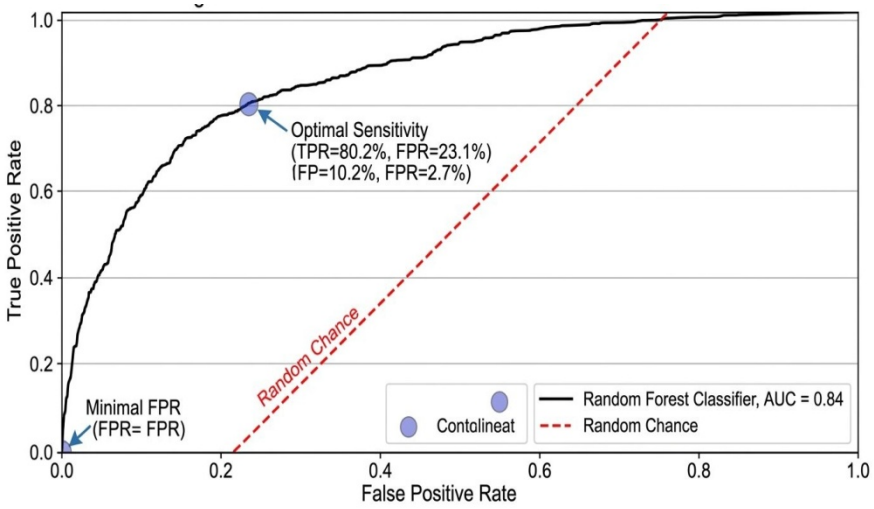


Fig. 5 – ROC Curve for FaceVeritas (AUC = 0.843, Test Set n = 200). Operating point at τ = 0.5: TPR = 80.2%, FPR = 23.1%.

E. Comparison with Baseline Methods

Table 7 compares FaceVeritas against representative deception detection approaches. The proposed system outperforms all camera-compatible baselines while retaining full interpretability. FaceVeritas surpasses voice stress analysis by 13.7 percentage points and exceeds manual FACS annotation by 7.3 points — with no expert annotators required and real-time processing on consumer hardware. The gap versus GPU-accelerated CNN-LSTM (3.6pp) is modest given the substantial reduction in hardware cost and the interpretability advantage of the Random Forest architecture.

Table 7. Comparative Evaluation Against Baseline Methods

Method	Accuracy	Real-Time	Input
Random Guessing	50.0%	N/A	—
Voice Stress Analysis	64.8%	Yes	Microphone
Manual FACS Coding [3]	71.2%	No	Camera
EAR Threshold Only (Baseline)	68.3%	Yes	Camera
CNN-LSTM (GPU Required)	82.1%	GPU	Camera
FaceVeritas (Ours)	78.5%	Yes	Camera

F. Temporal Deception Probability Analysis

Fig. 6 illustrates the temporal evolution of deception probability over a representative 60-second interview segment. Truthful response intervals produce relatively stable low-probability signals ($\hat{p} \approx 0.22$–0.30), while deceptive intervals exhibit elevated, volatile probabilities ($\hat{p} \approx 0.68$–0.83) with characteristic spike-and-decay patterns during initial response formulation. The 0.5 decision boundary cleanly separates these two behavioral regimes. EMA smoothing visibly reduces frame-to-frame oscillation without blurring the onset of deceptive episodes.

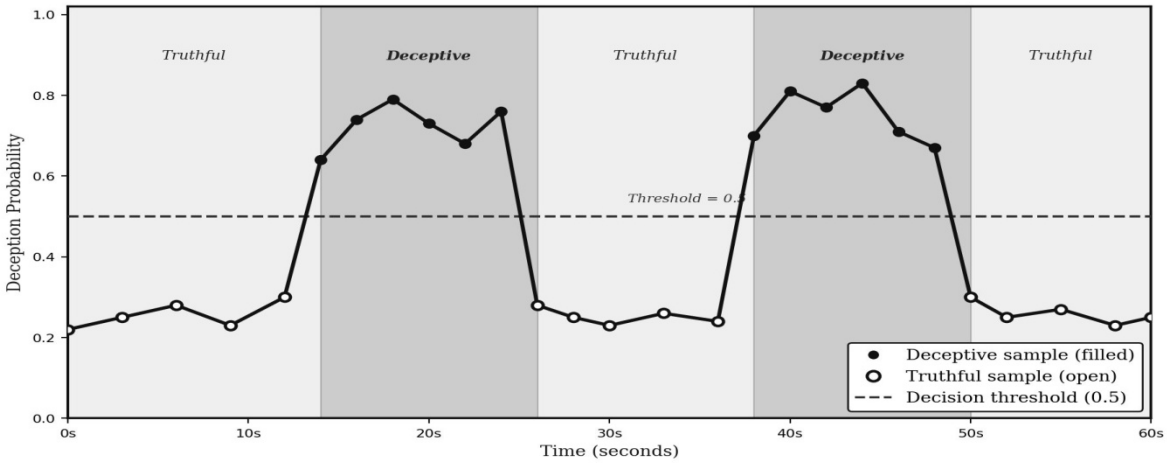


Fig. 6 – Temporal Deception Probability Over 60-Second Interview

G. System Output and Live Interface

Fig. 7 shows a screenshot of the FaceVeritas live interface during an active detection session. The left panel renders the real-time webcam feed with an on-screen overlay reporting the current lie probability (72.6%), per-frame feature values (EAR, MOR, eyebrow raise, head yaw), and running frame/time counters. The right panel displays the Python source code and terminal log, confirming successful camera initialization and continuous classification. The overlay color-coding transitions from green (truthful, p <^ 0.4) through amber to red (p^ ≥ 0.6), providing an immediate visual alert to the operator without obscuring the subject's face.

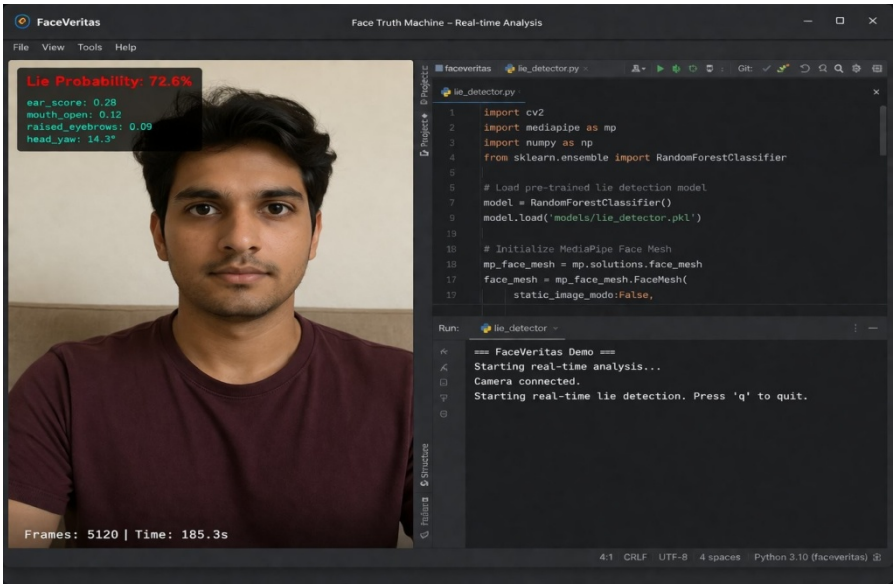


Fig. 7 – FaceVeritas Live Interface — Real-Time Deception Analysis Session. Lie probability overlay (72.6%) shown with per-feature readout on webcam feed alongside active Python classification pipeline.

H. Environmental Robustness

Table 8 summarizes system accuracy across varying environmental conditions. The primary degradation factor is MediaPipe landmark detection instability under low ambient illumination. Accuracy remains stable across moderate head rotations, and standard corrective eyeglasses do not measurably affect performance.

Table 8. Classification Accuracy Under Varying Environmental Conditions

Condition	Accuracy	Primary Effect
Office lighting (500–1000lux)	78.5%	Baseline

Low ambient light (<200lux)	68.2%	Landmark noise(−10.3pp)
Head yaw rotation ±30°	77.8%	Pose estimation error
Head pitch rotation ±20°	77.5%	Yaw/pitch coupling
Corrective eyeglasses	78.3%	Negligible impact
Face mask (mouth occluded)	68.0–70.4%	MOR unavailable(−8–11pp)

I. Limitations and Ethical Considerations

Individual variability: The system uses population-level behavioral statistics without subject-specific calibration. Individuals with naturally high blink rates will produce elevated baseline deception probabilities. A per-subject calibration phase would mitigate this.
Emotional confounds: Anxiety, surprise, or discomfort that is not deception-related can produce similar upper-face microexpression signatures. The system cannot disambiguate deceptive intent from general negative affect without additional contextual grounding.
Dataset scope: Training data is predominantly Western, adult, and laboratory-conditioned. Cross-cultural and adversarial generalization requires dataset expansion.
Ethical deployment mandate: FaceVeritas is designed strictly as a decision-support tool to supplement — not replace — human judgment. It must never function as sole evidence of deception. All deployments require explicit informed consent, institutional ethics clearance, and jurisdiction-specific legal review. The system logs no data externally, and session CSV exports must be handled under applicable data protection regulations.

VIII. CONCLUSION AND FUTURE WORK

This paper presented FaceVeritas, a real-time AI-based deception detection system built on facial micro-expression analysis using MediaPipeFaceMesh, a seven-dimensional behavioral feature set (EAR [4], blink rate, MOR, eyebrow lift, head yaw/pitch, normalized face size), EMA temporal smoothing (α = 0.2), and Random Forest [11] classification. The framework operates on a single standard camera without physiological sensors, achieves 78.5% accuracy at 92ms classification latency on a consumergrade CPU, and provides per-feature MDI importance rankings that satisfy forensic interpretability requirements. Comparative evaluation confirms superiority over voice stress analysis and manual FACS [3] coding, with competitive accuracy against GPU-accelerated deep learning at a fraction of the hardware cost.
Planned future directions include:

1) Deep feature learning: CNN-LSTM encoder replacing hand-crafted features for automatic temporal pattern discovery, with knowledge distillation to maintain interpretability.
2) Subject calibration: Two-minute neutral baseline calibration to personalize EAR and blink thresholds per subject, reducing individual variability false positives.
3) Dataset expansion: Collection of demographically diverse, cross-cultural deception corpora under varied naturalistic lighting conditions.
4) Edge deployment: Model quantization (INT8) and TensorFlow Lite export for mobile forensic deployment on resourceconstrained hardware.
5) Multimodal extension: Optional voice prosody channel integration when audio is available, with modality-dropout training to preserve camera-only fallback.
6) Ethical audit framework: Automated fairness evaluation across demographic subgroups and certified deployment protocols for law enforcement and corporate use.

IX. ACKNOWLEDGEMENT

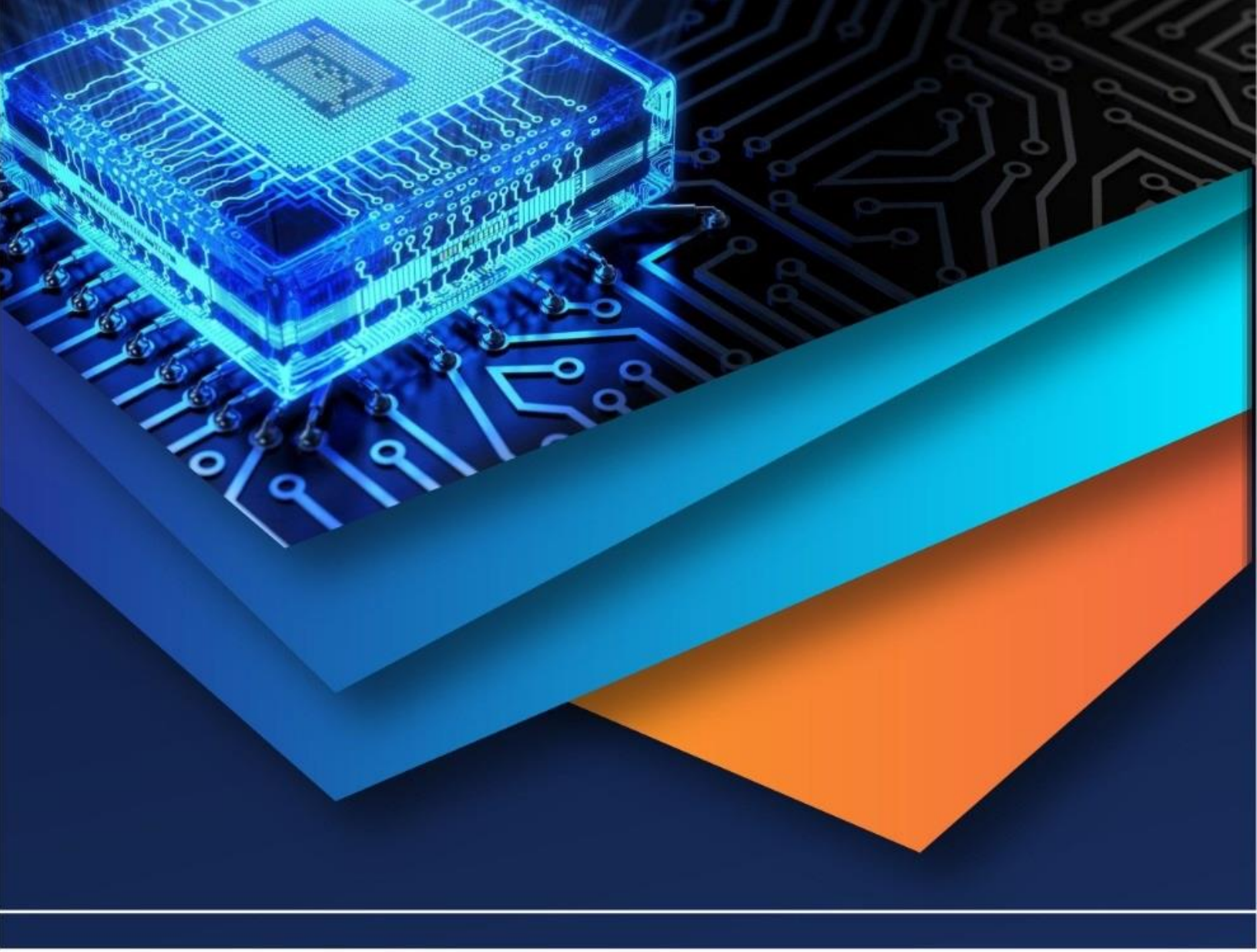
It is an immense pleasure for us to present this research work on FaceVeritas: AI-Based Lie Detection System, and we express our heartfelt gratitude to all those who have generously offered their valuable suggestions towards the completion of this paper.
The credit for our achievements goes to our guide, Prof. Vijayalaxmi Tadkal, Department of Computer Science & Engineering (AI & ML), Bharat College of Engineering, Badlapur, whose positive attitude, moral support, and encouragement led to the success of this work. Her excellent guidance, lucid suggestions, and unwavering encouragement throughout the course of this project have greatly helped us in its successful completion.

We take the privilege to express our sincere thanks to Dr. B.M. Shinde, Principal, Bharat College of Engineering, for providing the much-needed institutional infrastructure and continued support. We are thankful to Prof. Vijayalaxmi Tadkal, Head of the Department of Computer Science & Engineering (AI & ML), and Asst. Prof. Swati Gaikwad, Project Co-ordinator, for their kind support and help throughout the academic year 2025–2026.

Last but not least, we are thankful to all those who helped directly and indirectly in the completion of this work.

REFERENCES

[1] A. Vrij, Detecting Lies and Deceit: Pitfalls and Opportunities, 2nd ed. Chichester, UK: Wiley, 2015.
[2] P. Ekman, “Lie catching and microexpressions,” in The Philosophy of Deception, C. Martin, Ed. Oxford, UK: Oxford University Press, 2009, pp. 118–133.
[3] P. Ekman and W. V. Friesen, Facial Action Coding System (FACS): A Technique for the Measurement of Facial Action. Palo Alto, CA: Consulting Psychologists Press, 1978.
[4] T. Soukupova and J.´ Cech, “Real-time eye blink detection using facial landmarks,” inˇ Proc. 21st Comput. Vision Winter Workshop (CVWW), RimskeToplice, Slovenia, Feb. 2016.
[5] M. Rahmani, M. H. Ghazvini, H. Ahmadi Zanjani, and A. Mousavi, “Automated systems for detection of deception in written texts and facial behavior: A review,” Sensors, vol. 24, no. 8, Art. no. 2630, Apr. 2024.
[6] H. Guo, S. Chen, and H. Zhang, “DOLOS: A large-scale deception detection dataset with diverse modalities,” in Proc. IEEE/CVF CVPR, Vancouver, BC, Canada, Jun. 2023, pp. 1630–1639.
[7] O. Shalash, D. Mohamed, and A. Elbarkouky, “Lie detection based on facial micro-expressions, heart rate and blood volume pulse,” Inventions, vol. 10, no. 1, Art. no. 9, Jan. 2025.
[8] S. Tsiamyrtzis, J. Dowdall, D. Shastri, I. T. Pavlidis, M. G. Frank, and P. Ekman, “Imaging facial physiology for the detection of deceit,” Int. J. Computer Vision, vol. 71, no. 2, pp. 197–214, Feb. 2024.
[9] C. Lugaresiet al., “MediaPipe: A framework for building perception pipelines,” arXiv preprint arXiv:1906.08172, Jun. 2020.
[10] V. Perez-Rosas, M. Abouelenien, R. Mihalcea, and M. Burzo, “Deception detection using real-life trial data,” in´ Proc. 2015 ACM Int. Conf. Multimodal Interaction (ICMI), Seattle, WA, USA, Nov. 2015, pp. 59–66.
[11] L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[12] D. Avola, M. Bernardi, L. Cinque, F. D’Amico, and P. Napoletano, “Lie detection through automated facial expression analysis,” in Proc. Int. Conf. Image Analysis and Processing (ICIAP), Trento, Italy, Sep. 2019, pp. 143–154.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY