Designing Non-Distracting Affective Interfaces for the Professional Workspace

Stabilizing Real-Time Facial Inference into Calm 4-DOF Robotic Expression

Academic Research Poster | ICCA 2026 | Human-Computer Interaction & Social Robotics

Group Members: Anjir Jaman (22-49235-3), Md. Towfiq Bin Hasan (22-48732-3), Hasibul Islam Hasib (22-48507-3), Hasibul Islam (22-48498-3) | Department of Computer Science, AIUB

Introduction

Affective desktop companions can support focus, productivity, and mental well-being by sensing user states and giving calm peripheral feedback.
Problem: direct frame-by-frame actuation creates behavioral jitter - abrupt movements caused by blinks, sneezes, rapid head motion, or sensor noise.
Goal: convert noisy real-time facial signals into stable, intentional robotic postures that feel observant rather than anxious.

Research Contributions

- 17-feature sensory layer from MediaPipe Face Mesh, including PERCLOS, blink rate, gaze variance, EAR, brow tension, and mouth/yawn metrics.
- Two-stage temporal filter: 85% supermajority vote buffer over 90 frames plus state-specific hysteresis timers.
- 4-DOF 3D-printed robotic companion with synchronized RGB chromatic feedback.

Methods: Sense–Think–Act Pipeline

- Sense: 30 FPS video stream -> 468-point facial mesh with iris refinement.
- Think: Random Forest classifier estimates six cognitive/affective states.
- Filter: confidence gate, vote buffer, and hysteresis stabilize the output.
- Act: Python sends serial commands to Arduino for servo posture and LED cues.

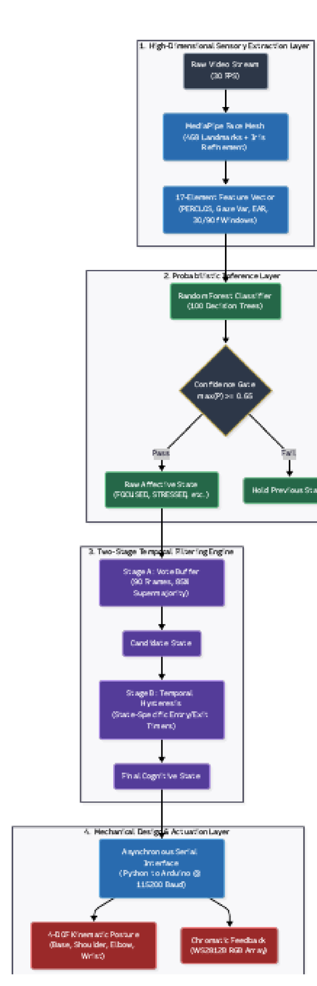


Figure 1. System architecture from MediaPipe extraction to Arduino actuation.

Feature Engineering

Facial measurements are normalized by face height to reduce camera-distance effects. Short windows (30 frames) capture micro-expressions; long windows (90 frames) track baselines.
Key fatigue and attention indicators include Eye Aspect Ratio (EAR), PERCLOS, blink frequency, yawn rate, and gaze iris offset variance.
The final 17-element vector balances interpretability, low latency, and robust affective awareness.

Ethics & Privacy

- All facial landmark extraction and Random Forest inference run locally.
- No video frames or biometric identifiers are sent to cloud storage.
- Participants provided informed consent before data collection.

Dataset & Inference Layer

Simulated interaction dataset: 12 participants, 5 hours of desktop usage, approximately 540,000 telemetry frames.
Target states: FOCUSED, STRESSED, HAPPY, DROWSY, DISTRACTED, and IDLE. The Random Forest uses 100 decision trees and an 80/20 train-validation split.

State	Kinematic cue	Color cue
FOCUSED	Leaning-in, wrist level	Cyan
STRESSED	Retracted, head down	Red
HAPPY	Base oscillation, wave	Yellow
DROWSY	Sagging posture	Purple
DISTRACTED	Rotated base, head tilt	White
IDLE	Neutral posture	Soft idle

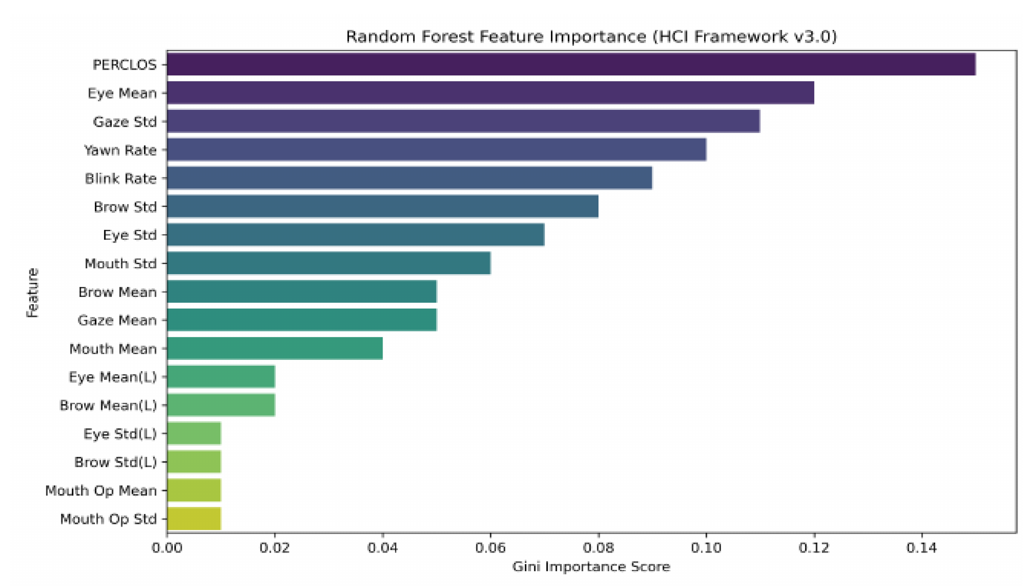
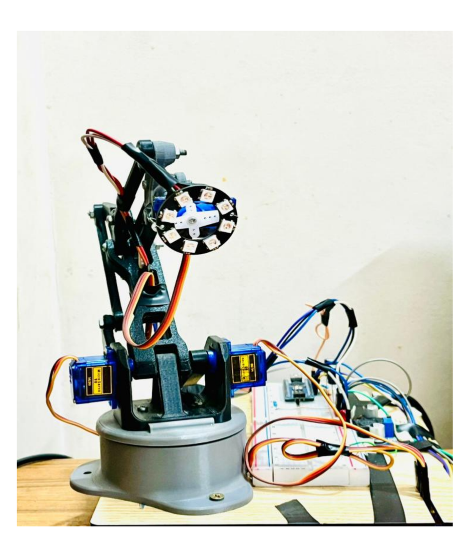


Figure 2. Feature importance: PERCLOS, EAR, and gaze variance dominate the RF model.

Mechanical Design & Actuation



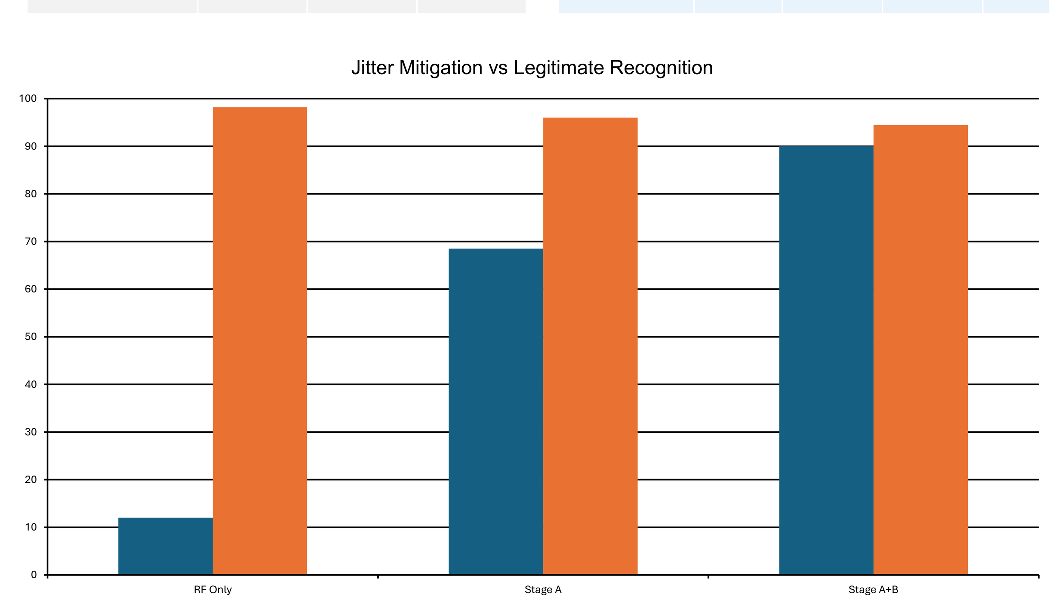
A rapid-prototyped PLA arm provides four expressive joints: base yaw, shoulder elevation, elbow extension, and wrist articulation.
A WS2812B RGB array adds peripheral chromatic feedback so state changes can be noticed without demanding focal attention.

Figure 3. 4-DOF affective desktop companion prototype.

Evaluation Results

State	Prec.	Recall	F1
Focused	0.85	0.87	0.86
Stressed	0.81	0.79	0.80
Happy	0.88	0.84	0.86
Drowsy	0.92	0.94	0.93
Distracted	0.78	0.82	0.80
Macro Avg	**0.86**	**0.86**	**0.86**

Filter	JMR	LSRR	SSR	Latency
RF Only	12.0%	98.2%	12.3/m	0.2s
Stage A	68.5%	96.0%	4.2/m	2.8s
Stage A+B	**90.0%**	**94.5%**	**0.3/m**	**3.8s**



Stage A+B suppresses 90% of transient jitter while preserving 94.5% sustained-state recognition.

Interpretation

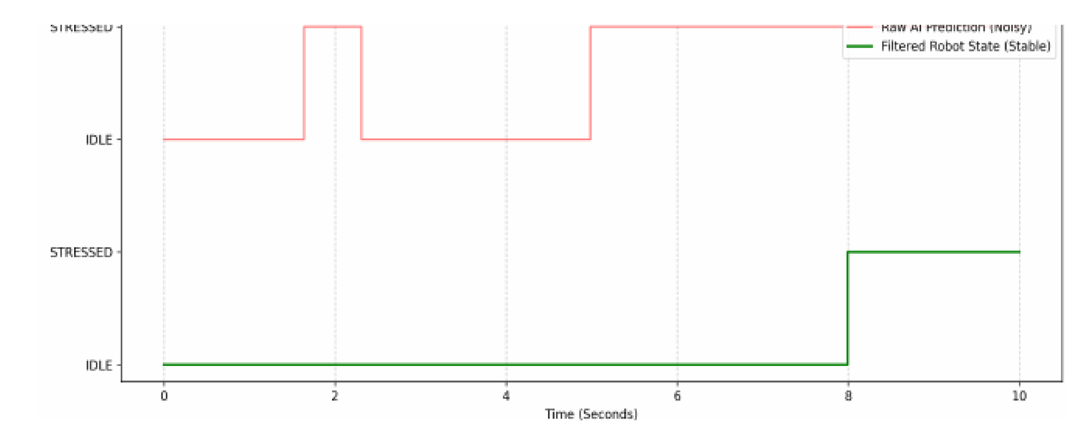
The RF-only baseline is highly responsive but physically unstable, producing 12.3 transitions per minute. Stage A reduces flicker by requiring a 3-second majority, while Stage B makes transitions deliberate through entry/exit timing.
The 3.8-second latency is an intentional design trade-off: the robot becomes calmer and less distracting, which is more suitable for peripheral HCI than instant reaction.

Results at a Glance

90.0% Jitter Mitigation Rate	**94.5%** Legitimate State Recognition	**0.3/m** State Stability Ratio
3.8s Intentional Delay	**0.86** Macro F1 Pre-Filter	**12** Study Participants

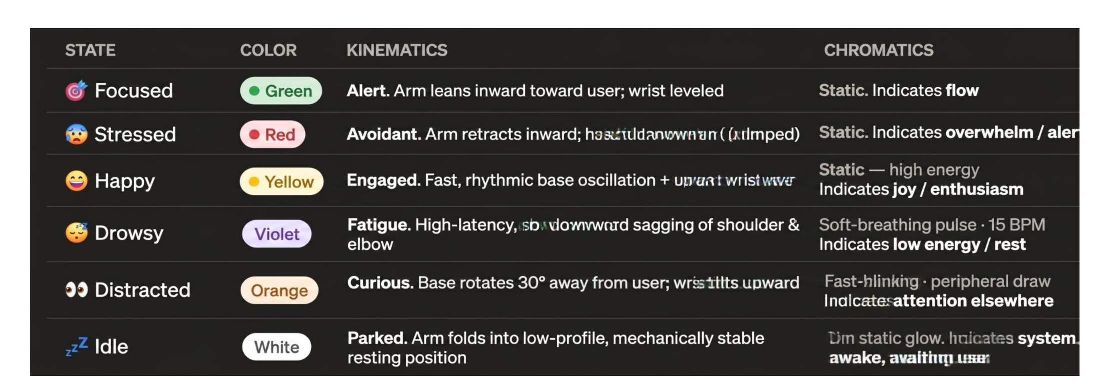
Participants rated the fully filtered system as significantly less distracting than the unfiltered baseline ($p < 0.01$), describing its behavior as more observant and less anxious.

Temporal Stability Analysis



Temporal stability: transient noise ignored; sustained change detected.

The filter blocks short-lived false positives while responding to genuine sustained state changes after a controlled delay.



STATE	COLOR	KINEMATICS	CHROMATICS
Focused	Green	**Alert.** Arm leans inward toward user; wrist leveled	Static. Indicates **flow**
Stressed	Red	**Avoidant.** Arm retracts inward;	Static. Indicates **overwhelm / aler**
Happy	Yellow	**Engaged.** Fast, rhythmic base oscillation +	Static — high energy Indicates **joy / enthusiasm**
Drowsy	Violet	**Fatigue.** High-latency, downward sagging of shoulder & elbow	Soft-breathing pulse · 15 BPM Indicates **low energy / rest**
Distracted	Orange	**Curious.** Base rotates 30° away from user; tilts upward	Fast-blinking · peripheral draw **attention elsewhere**
Idle	White	**Parked.** Arm folds into low-profile, mechanically stable resting position	Dim static glow. **system. awake,**

Conclusion

This poster presents a temporal mediation layer between affective inference and physical robot expression.
Core finding: classification accuracy alone is not enough for believable embodied HCI; temporal stability is essential for calm, trustworthy robotic companions.
The two-stage filter translates noisy facial telemetry into deliberate 4-DOF posture and RGB feedback.

Future Work

- Add audio prosody and keyboard telemetry for multimodal stress detection.
- Conduct longitudinal in-the-wild studies in real workplaces and study rooms.
- Explore personalization of hysteresis delays for different users and contexts.

References

[1] R. W. Picard, *Affective Computing*. Cambridge, MA, USA: MIT Press, 1997.
[2] M. Weiser and J. S. Brown, "The coming age of calm technology," Xerox PARC, Palo Alto, CA, USA, Tech. Rep., Oct. 1996. [Online]. Available: https://calmtech.com/papers/coming-age-calm-technology
[3] W. W. Wierwille, S. S. Wreggit, C. L. Kirn, L. A. Ellsworth, and R. J. Fairbanks, "Research on vehicle-based driver status/performance monitoring: Development, validation, and refinement of algorithms for detection of driver drowsiness," National Highway Traffic Safety Administration, Washington, DC, USA, Final Rep. DOT HS 808 247, Dec. 1994.
[4] T. Abe, "PERCLOS-based technologies for detecting drowsiness: Current evidence and future directions," *SLEEP Advances*, vol. 4, no. 1, p. zpad006, 2023. doi: 10.1093/sleepadvances/zpad006.
[5] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001. doi: 10.1023/A:1010933404324.
[6] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.
[7] Google LLC, "MediaPipe Face Mesh," Google AI Edge, 2020. [Online]. Available: https://ai.google.dev/edge/mediapipe/solutions/vision/face_landmarker
[8] G. Bradski, "The OpenCV library," *Dr. Dobb's Journal of Software Tools*, vol. 25, no. 11, pp. 120–125, 2000.
[9] M. Spezialetti, G. Placidi, and S. Rossi, "Emotion recognition for human-robot interaction: Recent advances and future perspectives," *Frontiers in Robotics and AI*, vol. 7, p. 532279, 2020. doi: 10.3389/frobt.2020.532279.
[10] C. S. Karunakaran *et al.*, "Advancing emotionally aware child–robot interaction with biophysical data and insight-driven affective computing," *Sensors*, vol. 25, no. 4, p. 1234, 2025. doi: 10.3390/s25041234. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC11858965/
[11] D. Abdurahman *et al.*, "Affective computing for human-robot interaction research: Four critical lessons for the hitchhiker," *arXiv preprint*, arXiv:2303.18176, 2023. [Online]. Available: https://arxiv.org/abs/2303.18176
[12] FastLED Project, "FastLED: Fast, easy LED library for Arduino," GitHub, 2013. [Online]. Available: https://github.com/FastLED/FastLED

Acknowledgement

We gratefully acknowledge Prof. Dr. Tabin Hasan, Department of Computer Science, AIUB, for his supervision, guidance, and valuable feedback throughout this HCI research work.