ACMod Thai Smile Problem: A Browser-Based Demonstration of Cultural Bias in Facial Expression Recognition

Technical Report & System Description

Ratchapon Lungmart (Poom)

College of Arts, Media and Technology (CAMT), Chiang Mai University, Thailand

NDF Dev & Fortal Interactive

February 2026

Abstract.

This report describes a real-time, browser-based demonstration system built to visualize cultural bias in facial expression recognition (FER). The demo uses a pre-trained deep learning pipeline (TinyFace Detector + Expression Net) running entirely client-side via TensorFlow.js to classify seven universal emotions from webcam input. It then exposes the fundamental limitation of Western-trained models when confronted with culturally-specific expressions — specifically the 13 distinct smile types in Thai culture, all of which current FER systems collapse into a single "Happy" label. The system includes five interactive modules: live emotion detection with user self-report mismatch tracking, temporal emotion timeline, snapshot comparison, a Thai smile taxonomy, and a global training data bias map. The demo is designed as a research communication tool for the ACMod (Affective Computing Models: from Facial Expression to Mind-Reading) project, a Horizon Europe MSCA Staff Exchange programme (Project No. 101130271).

1. Introduction

Facial expression recognition (FER) systems have achieved impressive accuracy on benchmark datasets such as FER2013, AffectNet, and RAF-DB. However, these benchmarks are overwhelmingly composed of Western, Educated, Industrialized, Rich, and Democratic (WEIRD) participants, creating a systematic cultural bias in deployed models [1, 2].

The ACMod project (Affective Computing Models: from Facial Expression to Mind-Reading) is a Horizon Europe MSCA Staff Exchange programme involving 10 institutions across 6 countries. Its objective is to build cross-cultural facial expression databases and develop culturally-aware emotion AI, with particular attention to underrepresented Southeast Asian populations [3].

This demo was built to communicate the core research problem of ACMod in an interactive, tangible way: *current AI cannot distinguish between 13 culturally-distinct Thai smile types because its training data lacks Southeast Asian representation.*

1.1 The Thai Smile Problem

Thai culture is known for the "Siamese Smile" — a rich vocabulary of smiles that convey meanings far beyond happiness. Where Western emotion models recognize a single "happy" expression, Thai social interaction distinguishes at least 13 smile types, each carrying distinct social meaning (see Section 4). A shy smile, an apologetic smile, and a mocking smile all activate similar facial muscle groups (primarily AU6 + AU12 in FACS notation) but differ in subtle intensity, asymmetry, eye involvement, and social context.

2. System Architecture

2.1 Design Principles

- **Zero-install** — Runs entirely in the browser; no backend server, no API keys
- **Real-time** — Detection loop at ~15–30 FPS depending on hardware
- **Privacy-first** — All processing is client-side; no video data leaves the device
- **Educational** — Each module is designed to demonstrate a specific aspect of FER bias

2.2 Technology Stack

Table 1. Technology stack and dependencies.

Component	Technology	Version / Source
ML Runtime	TensorFlow.js (WebGL backend)	Bundled with face-api.js
Face Detection & FER	face-api.js	@vladmandic/face-api v1.7.12
Model Hosting	jsDelivr CDN	cdn.jsdelivr.net
Visualization	HTML5 Canvas API	Native browser
UI Framework	Vanilla JavaScript + CSS	No external dependencies
Hosting	GitHub Pages	poomxchapon.github.io

2.3 Detection Pipeline

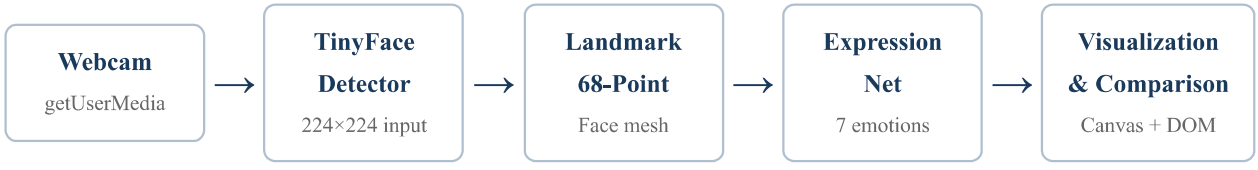

Figure 1. End-to-end detection pipeline. All stages execute client-side via TensorFlow.js WebGL backend.

The pipeline executes three neural networks in sequence per frame:

```
// Core detection call — executed every animation frame
const result = await faceapi
  .detectSingleFace(video, new faceapi.TinyFaceDetectorOptions({
    inputSize: 224,
    scoreThreshold: 0.5
  }))
  .withFaceLandmarks()      // 68-point mesh
  .withFaceExpressions();   // 7-class softmax
```

3. Models and Training Data

3.1 Face Detection: TinyFace Detector

The TinyFace Detector is a lightweight, MobileNet-based single-shot face detector optimized for real-time browser execution. It processes input at 224×224 pixels with a confidence threshold of 0.5. The model outputs bounding box coordinates and a detection confidence score.

Table 2. TinyFace Detector configuration.

Parameter	Value	Purpose
Input Size	224 × 224 px	Balance between accuracy and speed
Score Threshold	0.5	Minimum confidence to accept detection
Architecture	MobileNet V1 backbone	Depthwise separable convolutions for speed
Output	Bounding box + confidence	Single face per frame

3.2 Facial Landmark Detection: 68-Point Model

The landmark model predicts 68 facial keypoints following the Multi-PIE annotation scheme [4], covering the jawline (17 points), eyebrows (10), nose (9), eyes (12), and mouth (20). These landmarks are used for visualization (face mesh overlay) and provide spatial context for expression analysis.

3.3 Expression Classification: FER Model

The expression network is a lightweight CNN trained on the **FER2013** dataset [5], producing a 7-class softmax distribution over universal emotions: Happy, Sad, Angry, Disgusted, Surprised, Fearful, and Neutral.

Table 3. FER2013 dataset composition and known biases.

Property	Value
Total Images	35,887
Image Size	48 × 48 px (grayscale)
Classes	7 (Happy, Sad, Angry, Disgusted, Surprised, Fearful, Neutral)
Collection Method	Google image search with emotion keywords
Reported Accuracy	~65–72% (human agreement ~65%)
Demographic Bias	Predominantly Western/Caucasian faces
Cultural Limitation	No Southeast Asian-specific expressions
Label Taxonomy	Based on Ekman's 6 basic emotions + Neutral

3.4 The Bias Problem: Global Training Data Distribution

The demo includes an interactive world map visualizing the geographic distribution of FER training data. The data presented is aggregated from published dataset demographics across major FER benchmarks (FER2013, AffectNet, RAF-DB, SFEW, CK+, and JAFFE).

Table 4. Estimated global distribution of FER training data.

Region	Representation	Primary Datasets
Western / WEIRD	**~80%**	FER2013, AffectNet, CK+, SFEW, DISFA
East Asian	~15%	JAFFE, CFEE, RAF-DB (partial)
Southeast Asian	**~3%**	Virtually none at scale
Other Regions	~2%	Scattered across minor datasets

The map highlights 8 high-data countries (USA, Canada, UK, Germany, France, Italy, Netherlands, Australia), 5 medium-data countries (China, Japan, South Korea, Turkey, Iran), and 9 low-data countries including Thailand, Vietnam, and Indonesia — the three Southeast Asian nations participating in ACMod.

4. The 13 Thai Smiles: A Cultural Taxonomy

The demo presents 13 culturally-recognized smile types in Thai society. Each type activates overlapping facial muscle groups (primarily the zygomatic major and orbicularis oculi muscles, corresponding to FACS Action Units AU6 and AU12), yet carries fundamentally different social meanings. Current FER systems classify all 13 as "Happy."

Table 5. The 13 Thai smile types and their social functions.

#	Thai Name	English Name	Social Function	AI Label
1	ยิ้มทักทาย	Greeting Smile	Polite acknowledgment of another person	Happy (all 13)
2	ยิ้มเกรงใจ	Deferential Smile	Showing respect; not wanting to impose	
3	ยิ้มขอโทษ	Apologetic Smile	Expressing regret or mild embarrassment	
4	ยิ้มอาย	Shy Smile	Modesty or social embarrassment	
5	ยิ้มแห้ง	Dry / Forced Smile	Social obligation without genuine emotion	
6	ยิ้มเยาะ	Mocking Smile	Sarcasm, irony, or contempt	
7	ยิ้มทั้งน้ำตา	Tearful Smile	Smiling through pain or sadness	
8	ยิ้มพอใจ	Contented Smile	Inner satisfaction, quiet pleasure	
9	ยิ้มดีใจ	Happy Smile	Genuine joy or excitement	
10	ยิ้มหวาน	Sweet Smile	Affection, flirtation, or tenderness	
11	ยิ้มรู้ทัน	Knowing Smile	Awareness of hidden information	
12	ยิ้มปลง	Acceptance Smile	Letting go; Buddhist equanimity	
13	ยิ้มฝืน	Nervous Smile	Anxiety or discomfort masked by smile	

This taxonomy demonstrates that the Ekman 6+1 model of "universal" emotions fails to capture the functional diversity of facial expressions in non-Western cultures. The Thai smile vocabulary is particularly rich because Thai social interaction relies heavily on indirect communication, where the smile serves as a multi-purpose social signal rather than a simple emotion indicator.

5. Interactive Demonstration Modules

5.1 Module 1: Live Emotion Detection with Mismatch Tracking

The primary module displays real-time webcam input with overlaid face detection (cyan bounding box with corner indicators) and optional 68-point facial landmark mesh (purple). Seven emotion bars display confidence percentages updated via linear interpolation (LERP factor = 0.3) for smooth animation.

Self-Report Mechanism: A grid of 9 feeling buttons (Shy, Polite, Sorry, Confused, Nervous, Happy, Sad, Grateful, Neutral) allows users to indicate their actual emotional state. When the AI detects "Happy" with >60% confidence while the user has selected a non-happy feeling, a red mismatch alert is triggered and a running mismatch counter is incremented. This creates a tangible, quantified demonstration of cultural misclassification.

Table 6. Self-report feelings vs. AI classification mapping.

User Feeling	Culturally Valid?	AI Classification	Mismatch?
Shy	Yes (ยิ้มอาย)	Happy	Yes
Polite	Yes (ยิ้มเกรงใจ)	Happy	Yes
Sorry	Yes (ยิ้มขอโทษ)	Happy	Yes
Confused	Yes (context-dependent)	Happy / Neutral	Yes
Nervous	Yes (ยิ้มฝืน)	Happy	Yes
Happy	Yes (ยิ้มดีใจ)	Happy	No
Sad	Yes (ยิ้มทั้งน้ำตา)	Happy	Yes
Grateful	Yes (context-dependent)	Happy	Yes
Neutral	Yes	Happy / Neutral	Varies

5.2 Module 2: Emotion Timeline

A real-time line chart rendered on an HTML5 Canvas tracks all 7 emotion confidence scores over the most recent 80 detection frames (approximately 3–5 seconds of data). The currently dominant emotion is rendered with a thicker, brighter line while non-dominant emotions fade to lower opacity. This visualization reveals the temporal instability of frame-by-frame FER and the tendency for "Happy" to dominate when a Thai smile is present.

5.3 Module 3: Snapshot Comparison

Users can capture two distinct facial expressions and compare the AI's interpretation side-by-side. The comparison logic evaluates:

- **Same emotion + score difference < 0.15:** AI sees both as nearly identical
- **Same emotion + larger difference:** AI detects intensity variance only
- **Different emotions:** AI successfully differentiates

This module demonstrates that two culturally-distinct Thai smiles (e.g., **ยิ้มเกรงใจ** vs. **ยิ้มเยาะ**) are classified identically by the model — the AI measures intensity, not cultural meaning.

5.4 Module 4: Thai Smile Gallery

A visual grid displays all 13 Thai smile types with emoji representations, Thai names, English translations, and contextual descriptions. Each card carries the annotation "AI sees: Happy" to reinforce the classification collapse.

5.5 Module 5: Data Bias World Map

An interactive Canvas-based world map plots the geographic distribution of FER training data using three tiers (high/medium/low) with distinct visual encodings. Thailand is highlighted with a pulsing animation

to indicate the ACMod target region. Statistical overlays display the ~80/15/3/2% distribution breakdown.

6. Rendering and Performance

6.1 Smooth Rendering via LERP

To prevent visual jitter from frame-to-frame detection variance, all rendered values (bounding box position, emotion bar widths, confidence percentages) are interpolated using linear interpolation with a factor of 0.3:

```
// Smooth interpolation between frames
rendered_value = rendered_value + 0.3 * (detected_value - rendered_value)
```

This creates visually smooth transitions while maintaining responsiveness to genuine expression changes.

6.2 Face Visualization

Table 7. Face detection visualization parameters.

Element	Color	Properties
Bounding Box	`#00d4ff44` (semi-transparent cyan)	1px line width
Corner Indicators	`#00d4ff` (solid cyan)	3px line width, 14px length
Landmark Mesh	Purple	68 points, toggleable
Emotion Label	Per-emotion color	Positioned above bounding box

6.3 Performance Characteristics

TYPICAL FPS
15–30
Hardware-dependent (WebGL)

MODELS LOADED
3
Face + Landmark + Expression

TIMELINE BUFFER
80 frames
~3–5 sec history

LERP FACTOR
0.3
Smooth animation interpolation

7. Discussion

7.1 What the Demo Reveals

The demo provides three levels of evidence for cultural bias in FER:

1. **Quantitative mismatch:** The self-report mechanism produces a concrete mismatch rate, demonstrating that 7 out of 9 feeling options (78%) can trigger misclassification when accompanied by a culturally-appropriate Thai smile.
2. **Temporal instability:** The emotion timeline reveals that even within a stable expression, the AI's confidence fluctuates significantly — suggesting the model lacks robust representations for non-prototypical expressions.
3. **Structural limitation:** The 7-class taxonomy itself is insufficient. Even perfect accuracy within these 7 classes would not capture the 13 Thai smile types, because the label space was designed from a Western psychological framework [6].

7.2 Relevance to ACMod Research Objectives

This demo directly supports three ACMod work packages:

- **WP1 (Facial Expression Reconstruction):** Demonstrates the need for culturally-aware expression models beyond Ekman's universal framework.
- **WP2 (4D Facial Expression Database):** Motivates the creation of a Thai/Vietnamese/Indonesian facial expression database to address the ~3% representation gap.
- **WP3 (Commercialisation):** The browser-based architecture shows a viable deployment path for culturally-aware FER in consumer applications.

7.3 Limitations

- The demo uses a lightweight model (TinyFace + FER2013-trained CNN) rather than state-of-the-art architectures, which may exaggerate classification errors.
- The 13 Thai smile taxonomy is culturally documented but not yet validated through controlled experimental study with FACS coding.
- Self-report feelings are subjective and may not precisely correspond to displayed expressions.
- Browser-based WebGL performance varies significantly across devices.

8. Conclusion

This demonstration system provides an accessible, interactive proof-of-concept that current facial expression recognition technology carries significant cultural bias. By running entirely in the browser with zero installation requirements, it enables researchers, stakeholders, and the public to experience the "Thai Smile Problem" firsthand.

The demo serves as both a research communication tool for the ACMod project and a motivation for building culturally-inclusive facial expression databases — the core mission of the ACMod Staff Exchange programme.

References

[1] Freire-Obregón, D. et al. (2025). "Modeling Cultural Bias in Facial Expression Recognition with Adaptive Agents." *arXiv:2510.13557*.

[2] Wang, Y. & Yan, J. et al. (2024). "A Survey on Facial Expression Recognition of Static and Dynamic Emotions." *arXiv:2408.15777*.

[3] ACMod Consortium (2024). "Affective Computing Models: from Facial Expression to Mind-Reading." Horizon Europe MSCA Staff Exchange, Project No. 101130271. *CORDIS*.

[4] Gross, R. et al. (2010). "Multi-PIE." *Image and Vision Computing*, 28(5), 807–813.

[5] Goodfellow, I. et al. (2013). "Challenges in Representation Learning: A Report on Three Machine Learning Contests." *ICML Workshop on Representation Learning*.

[6] Ekman, P. & Friesen, W.V. (1978). *Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Consulting Psychologists Press.

[7] Behzad, M. (2025). "Facial Emotion Learning with Text-Guided Multiview Fusion via VLM for 3D/4D FER." *arXiv:2507.01673*.

[8] Zhang, Y. et al. (2024). "Affective Computing in the Era of Large Language Models: A Survey from the NLP Perspective." *arXiv:2408.04638*.

[9] Zhao, G. et al. (2023). "4D Spontaneous Micro-Expression Database." *IEEE Transactions on Affective Computing*.

[10] Vladmandic (2023). "face-api.js: JavaScript API for Face Detection and Recognition." GitHub. v1.7.12.

ACMod Thai Smile Problem Demo — Technical Report

Ratchapon Lungmart (Poom) • CAMT, Chiang Mai University • February 2026

Live Demo: *poomxchapon.github.io/acmod-demo* • ACMod Project No. 101130271

Prepared with Friday Oracle — "The Oracle Keeps the Human Human"