Fontys University of Applied Sciences

Ethical Considerations in Emotion Detection Using AI

Emotion Detection Using AI (MobileNet + Transfer Learning)

Author: David Geamanu
Student ID: 5052238
Course: Minor: AI Systems
Coach: Frank van Gennip
Version: v1.1
Date: November 10, 2025

This document outlines the ethical, legal, and social implications considered during the design and development of the project, referencing GDPR [1], the EU AI Act [2], and IEEE Ethically Aligned Design [3].

Executive Summary

This statement summarizes how the project addresses privacy, consent, fairness, data security, ethical scope, and compliance. The system prioritizes on-device processing, explicit user control, and transparent communication. Risks are identified and mitigations are defined, aligning with GDPR principles [1] and current EU AI guidance [2].

Introduction

This document describes the ethical considerations for an emotion detection prototype that processes camera input to infer facial expressions. The goal is to ensure the design remains privacy-preserving, fair, transparent, and compliant with applicable regulations. The application is intended for research and educational purposes (e.g., wellbeing feedback), not surveillance or profiling, in line with principles of accountability and human agency stated in IEEE Ethically Aligned Design [3].

1. Privacy and Data Protection

Facial data constitutes sensitive biometric information. To minimize risk, the project follows data minimization and purpose limitation, as required by the GDPR [1].

- **Local processing:** All inference is performed on-device; no raw images or identifiers are transmitted.
- **No permanent storage:** By default, images/frames are processed in memory only. Any diagnostic logging excludes personal data.
- **Data deletion:** Temporary buffers are cleared immediately after inference.

2. User Consent and Transparency

The application implements explicit, informed consent in compliance with Article 6(1)(a) of the GDPR [1].

- **Opt-in:** Camera access and emotion inference are disabled until the user enables them.
- **Just-in-time notices:** A short notice explains what is processed, for what purpose, and for how long.
- **Controls:** Clear on/off toggle, session indicator (e.g., icon/LED), and a link to the privacy summary.

3. Data Storage and Security

If short-term storage is required for evaluation, the design follows the principles of integrity and confidentiality from Article 5 of the GDPR [1] and ISO-style information security best practices.

- **Scope:** Store only what is strictly necessary (e.g., aggregated metrics).
- **Protection:** Encrypt at rest; restrict access to authorized personnel.
- **Retention:** Define retention windows and automatic deletion schedules.

4. AI Bias and Fairness

The model may inherit bias from training data, which is addressed in line with the fairness principles of the EU AI Act [2]. Mitigations include:

- Inspecting class distribution and augmenting underrepresented classes.
- Reporting class-wise metrics and confusion matrices.
- Avoiding high-stakes use; keeping predictions informational, not determinative.

5. Ethical Use and Context

The main goal of this project is to use emotion detection in a positive and helpful way, such as reminding users to take breaks or encouraging them when they show positive emotions. The system is not made for surveillance, profiling, or influencing people's behaviour. All demonstrations and tests clearly explain the purpose of the system to anyone involved. Users have full control over when the application is active and can pause or stop it at any time. Any future version of the app will show a clear indicator when the camera is on and will include a short, easy-to-understand privacy explanation. By setting clear limits and responsible uses, the project aims to keep emotion detection ethical and focused on helping people, consistent with the IEEE's principles of human wellbeing and accountability [3].

6. Regulatory Alignment

The project aligns with key European data-protection and AI-ethics frameworks, including the General Data Protection Regulation (GDPR) [1] and the upcoming EU AI Act [2]. Under GDPR, it respects all major principles: lawfulness, fairness, transparency, purpose limitation, data minimisation, and integrity/confidentiality. The design treats all facial analysis as sensitive processing and ensures that it stays within a low-risk category by design—local processing, explicit user consent, and transparent operation. In terms of governance, documentation is maintained for the model's purpose, assumptions, and limitations to promote accountability. The project also draws on the IEEE Ethically Aligned Design framework [3], focusing on human agency, accountability, and the promotion of wellbeing. Together, these measures ensure that the prototype not only complies with current legal requirements but also anticipates future standards for trustworthy and responsible AI systems.

Risk & Mitigation Summary

Risk	Description	Likelihood	Mitigation
Privacy leakage	Unintended storage/transmission of frames	Low	On-device processing, no raw logs, code review
Bias	Uneven accuracy across groups/classes	Medium	Class-balance checks, per-class metrics, iterative tuning
Misuse	Use for surveillance without consent	Low/Med	Usage policy, UI indicators, opt-in only
Security	Unauthorized access to temp data	Low	Encryption, least-privilege access, short retention

Conclusion

With local processing, explicit consent, limited storage, fairness checks, and clear scope limits, the prototype adheres to responsible AI practices and prepares for compliant, user-respecting experimentation. This approach demonstrates alignment with the ethical frameworks of GDPR [1], the EU AI Act [2], and IEEE Ethically Aligned Design [3].

References

1. European Parliament and Council. *General Data Protection Regulation (EU) 2016/679.* Available at: https://eur-lex.europa.eu/eli/reg/2016/679/oj
2. European Commission. *Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act).* COM(2021) 206 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206
3. IEEE. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition.* IEEE Standards Association, 2019. Available at: https://ethicsinaction.ieee.org/