The Application of Facial Emotion Recognition (FER) in the Detection and Measurement of Burnout and Prolonged Stress Levels

Saputra Rizky Johan (사푸트라)
Advisor: Prof. Lee Chang-Gun

Seoul National University
Department of Computer Science & Engineering

December 2025

Motivation

- **Burnout is becoming a global concern throughout the world**. According to the World Health Organization (WHO), burnout is a psychological syndrome caused by chronic stress, often resulting in **emotional exhaustion**, **depersonalization**, and **reduced accomplishment**.
- Traditional tools such as the **Maslach Burnout Inventory (MBI)** rely on surveys to assess these burnouts. However, these assessments often times suffers from the following:
  - high degree of **subjectivity** and **recall bias**
  - inability to provide **real-time monitoring**
  - inconsistent **emotional reporting**
- This research aims to design a system that offers objective, continuous, and non-invasive burnout measurement by using a real-time **Facial Emotion Recognition (FER)** and **Ekman's Universal Emotions**.
- This project specifically uses the pre-trained TensorFlow/Keras model from HuggingFace, called **Emo0.1.h5**, amongst **5 - 10 participants**.

Research Problem

Before the implementation, several considerations were guided on this project:

- FER's reliability under real-time conditions (lighting, pose, duration).
- Temporal modeling's capacity to stabilize over frame-based averaging.
- Consistent monitoring of burnout patterns across various sessions.
- The alignment of system-derived scores with validated survey (MBI) results.

Main Problem: *Can a real-time FER system produce stable, interpretable, and survey-aligned burnout measurements?*

Hypotheses

- $\mathbf{H_1}$**:** Temporal LSTM modeling increases correlation with MBI scores.
- $\mathbf{H_2}$**:** Longer facial detection sessions yield more reliable burnout estimates.
- $\mathbf{H_0}$**:** No significant difference exists between the system's burnout score and MBI survey results.

Research Purpose

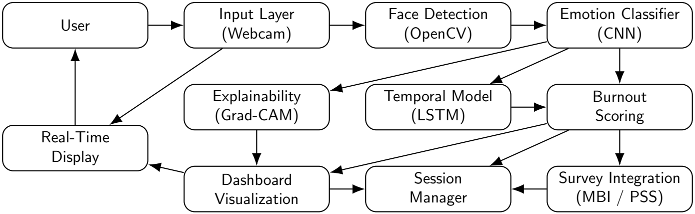
In order to achieve the goal of designing and implementing this **real-time, modular FER-based system**, five primary focus were proposed accordingly, likewise below:

- **System Architecture:** Implement a real-time and continuous modular Python-based FER pipeline.
- **Temporal Model:** Integrate neural networks such as CNN and LSTM to further capture the spatial and temporal emotion patterns, which subsequently allowing a longitudinal analysis.
- **Visual Interpretability:** Provide a coherent and readable visual feedback and communicate the stress and burnout patterns
- **Burnout Scoring:** Construct a burnout scoring equation using weighted valence and volatility.
- **Data Analysis:** Validate the system using MBI and PSS survey data.

System Architecture Overview

This system specifically consists of nine interconnected modules.

- **Input layer** collects real-time webcam frames for the system to process the user's localized facial region with its **face detection**.
- The **emotion classifier** predicts the emotions per frame, whilst the **temporal model** stabilizes these predictions over time.
- **Burnout scoring** computes these burnout levels and compares its scores with the **survey integration** compares results for validation.
- These results are then displayed with the **dashboard visualization** and **session manager** stores the session logs.



Burnout Scoring Equation

The burnout scores combines **two components** for every recorded session, T.

1. Weighted Emotional Valence

- **Negative** (Sad, Anger, Fear, Disgust) $\rightarrow$ **Positive** weights (α)
- **Positive** (Happy) $\rightarrow$ **Negative** weights (β)
- **Baseline** (Neutral, Surprise) $\rightarrow$ **Arbitrary** weights (γ)

$$S_{\text{burnout}} = \frac{1}{T} \sum_{t=1}^{T} \left(\alpha \sum_{e_i \in Negative} P_t(e_i) - \beta \sum_{e_j \in Positive} P_t(e_j) + \gamma \sum_{e_k \in Neutral} P_t(e_k) \right)$$

2. Temporal Variability

- Measures **emotional instability** across the session.
- Higher volatility $\rightarrow$ Higher stress dysregulation and variability (V)

$$V = \frac{1}{T} \sum_{t=1}^{T} (S_t - S_{mean})^2$$

By integrating λ as a balancing operator, the burnout score is given as such:

$$S_{\text{final}} = \lambda \cdot S_{\text{burnout}} + (1 - \lambda) \cdot V$$

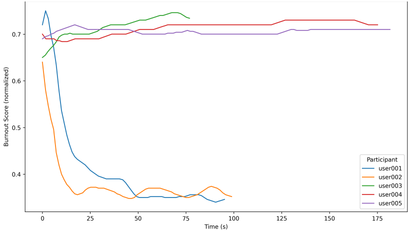
Results (Quantitative Metrics)

The results for the first 5 participants were given as such:

- **FER system burnout score** ranges between **37.2% – 71.7%**
- **MBI survey burnout score** ranges between **60% – 86.7%**
- Its **absolute difference** averages within $|S_{FER} - S_{MBI}| \approx$ **15.2%**
- LSTM smoothing improves temporal consistency
- FER captures moment-to-moment emotional fluctuations
- The result matches its psychological theory as participants with:
  - **stronger negative** emotional persistence $\rightarrow$ higher burnout scores.
  - **stronger positive** valence $\rightarrow$ early fluctuations but followed by stabilization.

Table 1. FER Burnout System Session Based Quantitative Metrics

Participant	Avg. System Burnout	Min	Max	Std. Dev	Survey Burnout	|System – Survey|
user001	0.400	0.340	0.780	0.090	0.600	0.200
user002	0.372	0.340	0.640	0.043	0.600	0.228
user003	0.717	0.650	0.750	0.021	0.733	0.016
user004	0.714	0.680	0.730	0.015	0.867	0.153
user005	0.706	0.690	0.720	0.006	0.867	0.161
Mean ± SD	**0.582 ± 0.179**	**–**	**–**	**–**	**0.733 ± 0.134**	**0.152 ± 0.082**



Correlation & Validation

In order to further validate the system, a statistical analysis such as the **Pearson's correlation coefficient** between the FER burnout scores and the MBI survey results were calculated, where:

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2} \cdot \sqrt{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}}$$

Where:

- X_i = FER System-derived burnout score for session i,
- Y_i = MBI Standardized survey burnout score for session i,
- $\bar{X}, \bar{Y}$ = Means of each score distribution.

By substituting the FER and MBI burnout scores, the result is shown as such:

- Pearson's correlation coefficient: $\mathbf{r} \approx \mathbf{0.904}$.
- Strong positive alignment between the FER burnout and MBI survey scores.
- Supports the hypotheses, H_1 and H_2, as temporal modeling improves and longer sessions produces more accurate and reliable scores respectively.

Conclusion & Future Work

Conclusion:

- FER based system effectively detects prolonged stress and burnout in real time whilst predicting the user's upcoming burnout level.
- Strong alignment with the MBI survey results ($r \approx 0.904$).
- Temporal consistency validates the psychological burnout dynamics.
- Non-invasive, scalable, and suitable for continuous monitoring.

Future Work:

- Multimodal inputs (speech, heart rate, EEG) $\rightarrow$ inclusion of external factors.
- Transformer-based temporal models
- Larger and more diverse datasets $\rightarrow$ current dataset may have limited inclusion of race, ethnicity and age to emotion change difference.
- Improve frame-rate performance with the inclusion of GPU acceleration.

Ultimately, this research delivers a real-time FER based framework that contributes to the overlapping fields between AI, medicine and psychology.