Project Proposal: Facial Expression Recognition

Team
Our group, *Emo for Short*, consists of Deepak Kumar Srivastava, Marcelino Bautista, and Xinyi Zhu.

Task Objective
The goal is to perform facial expression classification, identifying emotions from facial images. This task has broad real-world applications—for example, emotion recognition can support targeted advertising strategies, enabling companies to adapt their offerings to customer sentiment.

Dataset
We will use the [FER-2013 Facial Expression Dataset] from Kaggle.

A few example images are attached for reference.

Dataset Description
- **Image size**: 48×48 pixels
- **Color channels**: 1 (grayscale)
- **Variation**: The dataset includes a wide range of facial expressions, head poses, lighting conditions, and occlusions.
- **Consistency**: The faces have been automatically registered so that the face is more or less centred and occupies about the same amount of space in each image.

Labels / Emotion Categories
The dataset provides seven emotion categories, defined as below:
- 0 = Angry 😠
- 1 = Disgust 🤢
- 2 = Fear 😨
- 3 = Happy 😀
- 4 = Sad 😢
- 5 = Surprise 😲
- 6 = Neutral 😐

Dataset Composition

Emotion	Training Set (28,709 images)	Public Test Set (3,589 images)
Angry 😠	**13.9%** (*3,995 images*)	**26.7%** (*958 images*)
Disgust 🤢	**1.5%** (*436 images*)	**3.1%** (*111 images*)
Fear 😨	**14.3%** (*4,097 images*)	**28.5%** (*1,024 images*)
Happy 😀	**25.1%** (*7,215 images*)	**49.4%** (*1,774 images*)
Sad 😢	**16.8%** (*4,830 images*)	**34.8%** (*1,247 images*)
Surprise 😲	**11.0%** (*3,171 images*)	**23.2%** (*831 images*)
Neutral 😐	**17.3%** (*4,965 images*)	**34.4%** (*1,233 images*)

Proposed Approach

1. Traditional Feature Extraction Methods
   We will begin with two classical computer vision feature descriptors to establish a strong baseline.
   a. **Histogram of Oriented Gradients (HOG)**: Captures edge orientations and overall facial structure (e.g., raised eyebrows, frowns, smiles).
   b. **Local Binary Patterns (LBP)**: Captures local texture details such as wrinkles or furrows that reflect micro-expressions and subtle emotional cues.
   c. These features will be used to train conventional classifiers such as Support Vector Machines (SVM) or Random Forests.
2. Deep Learning with Transfer Learning
   We will leverage pretrained convolutional neural networks (CNNs) for high-level feature extraction.
   a. **ResNet-18:** A neural network with 18 layers that uses skip connections to make training easier. Lightweight and works well on smaller datasets like FER-2013. *Pros: fast to train and reliable; Cons: may miss very complex features.*
   b. **EfficientNet-B0:** A modern neural network that balances layers, width, and image resolution to achieve high accuracy efficiently. *Pros: accurate and resource-efficient; Cons: slightly more complex to fine-tune.*
   c. Grayscale images will be converted to 3-channel format to match pretrained ImageNet weights.