### **Project Report**

### **Project Report: Facial Expression Recognition Using VGG16**

**Author: Ehtijad Ali Shah**

**Email:** [**Here**](mailto:ehtyalee1919@gmail.com)

**GitHub Link:** [**Here**](https://github.com/Ehtijad-Ali)

**Section: 2**

**Roll Number: DSAI-GB-088**

**Date: 19 Oct 2024**

**Objective:** The project uses a custom dataset and transfer learning to build a facial expression recognition model. The goal is to classify facial expressions into seven distinct emotion categories: Angry, Disgusted, Fearful, Happy, Sad, Surprised, and Neutral.

### **Dataset and Preprocessing**

* **Dataset:**

The dataset contains images of faces with corresponding labels representing different emotions. It also includes bounding box information to crop the faces from the images.

**Data Source:**

* + - Label file path: [Here](https://www.kaggle.com/datasets/shahzadabbas/expression-in-the-wild-expw-dataset)
    - Image folder path: [Here](https://www.kaggle.com/)
* **Data Preparation:**

The labels are mapped to the following emotions:

* + - 0: Angry
    - 1: Disgust
    - 2: Fear
    - 3: Happy
    - 4: Sad
    - 5: Surprise
    - 6: Neutral

Faces with a bounding box confidence score below 60% are filtered out.

Each face is cropped, resized to 64x64 pixels, and normalized to have pixel values in the range [0, 1].

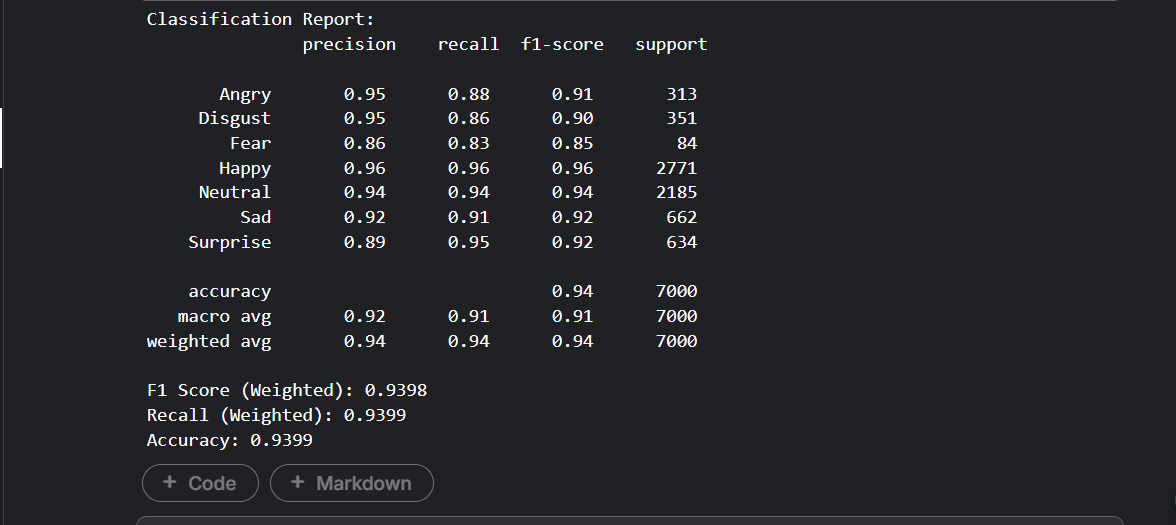
* **Class Distribution:** A class distribution plot shows the number of images per emotion category. This helps in understanding whether the dataset is balanced or not.

### **Model Architecture and Training**

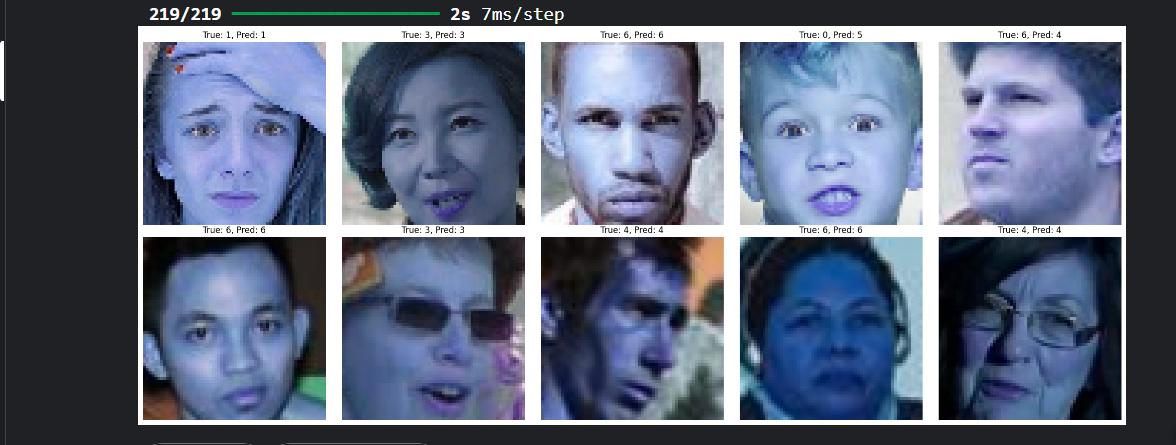
* **Transfer Learning with VGG16:**
  + **Base Model:** VGG16 (pre-trained on ImageNet) is used to extract features from the input images. The convolutional layers of VGG16 are frozen to retain the learned features.
  + **Custom Layers:**
    - Flattening the output from VGG16.
    - Two fully connected layers with 256 and 128 neurons, both using ReLU activation.
    - An output layer with softmax activation classifies the input into one of the seven emotion categories.
* **Training:**
  + The model is compiled using the Adam optimizer and sparse categorical cross-entropy as the loss function.
  + The dataset is split into 80% training and 20% testing. The model is trained for 10 epochs with a batch size of 32.
* **Evaluation:**
  + Test set accuracy: **~93%**
  + A second round of training is performed for an additional 10 epochs to further improve the model's performance.

### **Results and Analysis**

1. **Loss and Accuracy Trends:**
   * Training and validation loss, as well as accuracy trends, are plotted over the epochs. The plot shows how well the model is learning and generalizing on unseen data.
2. **Confusion Matrix:**
   * A confusion matrix is generated to visualize how well the model is performing in classifying each emotion. This provides a detailed view of misclassifications between emotions.
3. **Classification Metrics:**
   * **F1 Score (Weighted):** 0.93
   * **Recall (Weighted):** 0.93
   * **Accuracy:** 0.93



1. **Predictions on Test Data:**
   * The model's predictions on 10 random test images are displayed along with their true and predicted labels, providing a qualitative assessment of the model's performance.



### **Conclusion and Future Work**

This project demonstrates how transfer learning can be effectively used for facial expression recognition. Although the model achieves reasonable accuracy, there is room for improvement, such as addressing class imbalances or tuning the model further. Future work could also include augmenting the dataset or experimenting with other architectures.

**References:**

* VGG16 model: Simonyan, K., & Zisserman, A. (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition."