Report on Real-Time Analysis of Facial Emotion using Convolutional Neural Network

Team Members:

1. Aryan Baranwal (M25CSE035)
2. Jay Mehta (M25CSE034)
3. Kshitiz Budhathoki (M25CSE036)

Table of Contents

1. Problem Statement
2. Datasets
3. Methodology
4. Results and Analysis
5. Training and Testing
6. Model Performance
7. Contributions

1. Problem Statement:

Facial Emotion Recognition (FER) assists the computers in understanding human emotions through the use of facial expressions. Nevertheless, there are some obstacles to the existing systems. When the background conditions or the lighting change, they usually find it difficult to identify the emotions. Numerous models fail to work with real-life images or other datasets. They are also unable to work with grayscale images or snap tiny details on the faces that can help reveal subtle emotions. The identification of complicated or ambivalent feelings is still a significant difficulty. Besides, these systems are usually biased towards some age groups, genders, or ethnicities, thereby lowering their accuracy and dependability. Due to such problems, FER systems do not tend to provide the same results in practice. Thus, it is very much required to have a smarter, more dynamic and efficient FER system that is able to recognize emotions of anyone in real-time regardless of background and setting.

2. Datasets:

In this project, Kaggle FER2013 dataset was taken as a training and testing dataset of Facial Emotion Recognition (FER) model. FER2013 is a common benchmark dataset, which consists of grayscale pictures of human faces with various emotions. The dataset contains 35,887 images in seven categories of emotions, among them Happy, Sad, Angry, Surprise, Disgust, Fear, and Neutral. All images are 48x48 pixels in size, which is why they are suitable to be used in the deep learning and real-time application. The pictures of FER2013 were taken in different real-life situations, which is useful in testing the validity of the model in other lighting and environmental conditions. The data is sufficiently balanced in terms of faces expressions in terms of age, genders, and ethnicities, which contribute to the minimization of biases in training. It comes in handy especially to determine the performance of the model to difficult and natural images. The application of FER2013 allows the given model to learn more effectively and to generalize more to the actual task in emotion recognition.

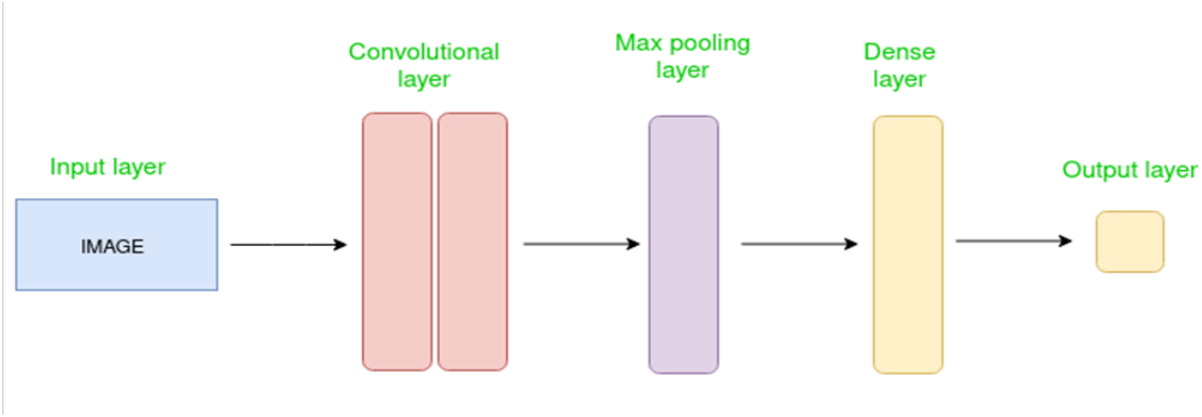
3. Methodology:

Input : We have taken 48x48 grayscale facial images FER-2013 dataset to train the model.

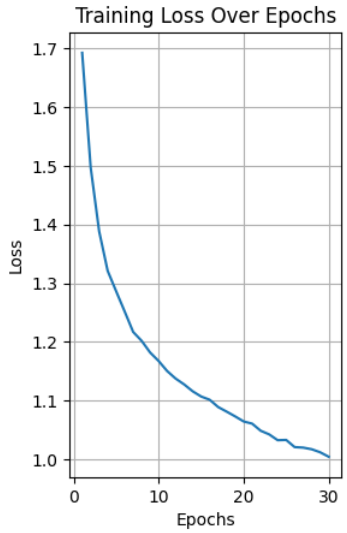
Layers :

- Conv1: 32 filters, 3x3 kernel, BatchNorm, LeakyReLU, MaxPooling.
- Conv2: 64 filters, 3x3 kernel, BatchNorm, LeakyReLU, MaxPooling.
- Conv3: 128 filters, 3x3 kernel, BatchNorm, LeakyReLU, MaxPooling.
- Conv4: 256 filters, 3x3 kernel, BatchNorm, LeakyReLU, MaxPooling.
- Fully Connected: 256 neurons, Dropout(0.5), Softmax (7 classes).

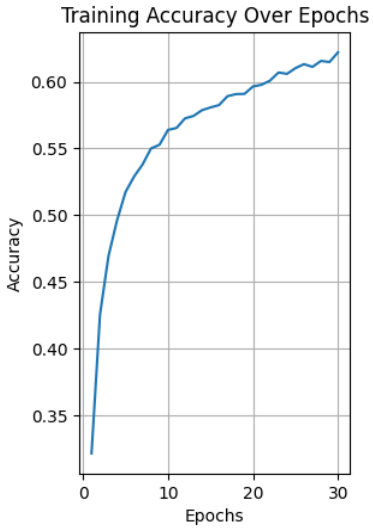
Output : We display the Current Facial Emotion above the bounded border

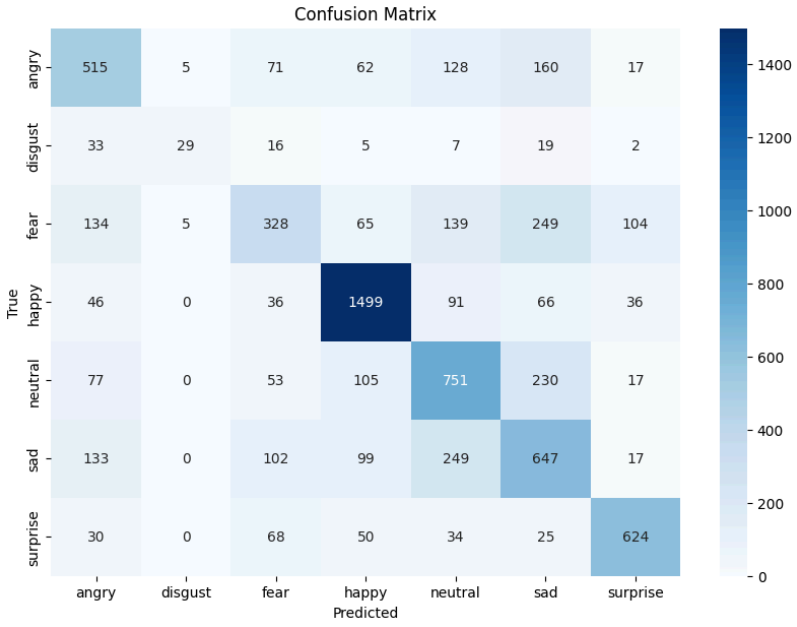


4. Result and Analysis:

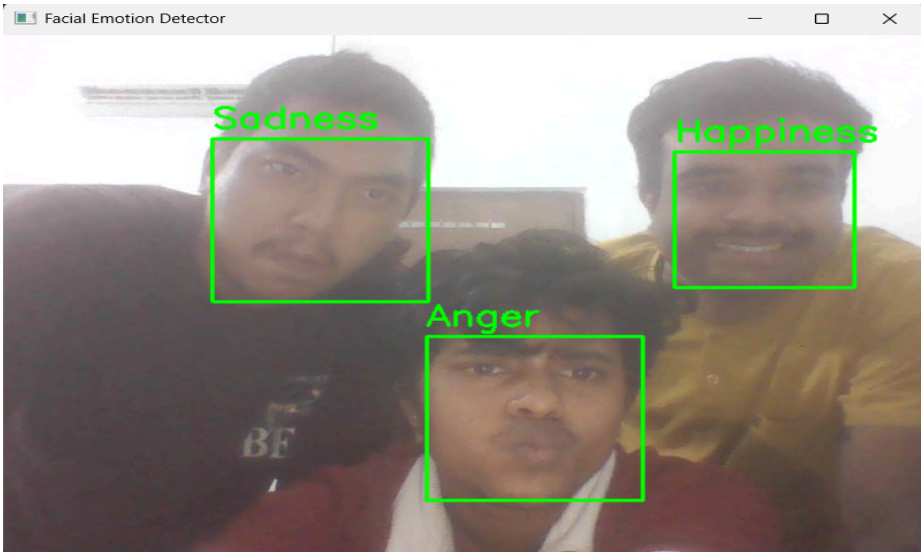



```
Accuracy: 0.6120
Precision: 0.6096
Recall: 0.6120
F1 Score: 0.6057
```





Epoch [1/30], Loss: 1.6925, Accuracy: 0.3214
Epoch [2/30], Loss: 1.4972, Accuracy: 0.4257
Epoch [3/30], Loss: 1.3883, Accuracy: 0.4695
Epoch [4/30], Loss: 1.3210, Accuracy: 0.4961
Epoch [5/30], Loss: 1.2862, Accuracy: 0.5173
Epoch [6/30], Loss: 1.2518, Accuracy: 0.5289
Epoch [7/30], Loss: 1.2169, Accuracy: 0.5379
Epoch [8/30], Loss: 1.2017, Accuracy: 0.5500
Epoch [9/30], Loss: 1.1818, Accuracy: 0.5528
Epoch [10/30], Loss: 1.1674, Accuracy: 0.5639
Epoch [11/30], Loss: 1.1503, Accuracy: 0.5654
Epoch [12/30], Loss: 1.1374, Accuracy: 0.5726
Epoch [13/30], Loss: 1.1275, Accuracy: 0.5744
Epoch [14/30], Loss: 1.1158, Accuracy: 0.5787
Epoch [15/30], Loss: 1.1069, Accuracy: 0.5808
Epoch [16/30], Loss: 1.1015, Accuracy: 0.5825
Epoch [17/30], Loss: 1.0890, Accuracy: 0.5892
Epoch [18/30], Loss: 1.0812, Accuracy: 0.5908
Epoch [19/30], Loss: 1.0731, Accuracy: 0.5910
Epoch [20/30], Loss: 1.0644, Accuracy: 0.5965
Epoch [21/30], Loss: 1.0607, Accuracy: 0.5978
Epoch [22/30], Loss: 1.0486, Accuracy: 0.6009
Epoch [23/30], Loss: 1.0424, Accuracy: 0.6070
Epoch [24/30], Loss: 1.0325, Accuracy: 0.6060
Epoch [25/30], Loss: 1.0330, Accuracy: 0.6103
Epoch [27/30], Loss: 1.0199, Accuracy: 0.6112
Epoch [28/30], Loss: 1.0171, Accuracy: 0.6157
Epoch [29/30], Loss: 1.0119, Accuracy: 0.6148
Epoch [30/30], Loss: 1.0043, Accuracy: 0.6221



5. Training and Testing

- Dataset: FER-2013 (Facial Expression Recognition dataset).
- Training Parameters:
  - Optimizer: Adam (lr = 0.001, weight decay = 1e-4).
  - Loss Function: Cross-Entropy Loss.
  - Batch Size: 64.
  - Epochs: 30.
- Evaluation Metrics: Accuracy, Precision, Recall, F1-Score.

6. Model Performance

Model	Accuracy
Custom CNN	61.20%

7. Contributions

1. Problem Statement and Research + Application Dashboard Creation + Coding : Kshitiz Budhathoki (M25CSE036)
2. Model Design and Implementation + Coding : Aryan Baranwal (M25CSE035)
3. Executing Training and Analysis + Coding + Report : Jay Mehta (M25CSE034)