

**NATIONAL UNIVERSITY SCIENCE AND TECHNOLOGY (NUST)**

(**High Impact Skills Development Program for Gilgit Baltistan**)

**Project Title:** Expression Classification from Facial Images

A Project Report

**Name: ASHFAQ JAVED**

**Email Address:** ashfaqahmedhunzai@gmail.com

**GitHub: https://github.com/ashfaqde/Expression-Classification-from-Facial-ImagesLinkedln:** https://www.linkedin.com/in/ashfaq-javed-b507541b5/

Expression Classification Report

### 1. Introduction

This report presents a project on 'Expression Classification' utilizing the Expression in the Wild (ExpW) Dataset, which contains 106,000 facial images. The objective of this project is to develop a machine learning model capable of identifying various facial expressions in real-time. Since facial expressions play a crucial role in human interaction, an accurate recognition system has significant applications in areas like emotion detection, security systems, entertainment, and beyond.

### 2. Dataset Overview

The Expression in the Wild (ExpW) dataset comprises 106,000 images, each labeled with one of seven facial expressions:

* Angry (0)
* Disgust (1)
* Fear (2)
* Happy (3)
* Sad (4)
* Surprise (5)
* Neutral (6)

Additionally, the dataset includes bounding box annotations, which are essential for locating the facial region within each image. These annotations are especially useful for pre-processing tasks such as cropping and resizing the face area before feeding the images into the model.

### 3. Data Preprocessing

To prepare the dataset for model training, several preprocessing steps were applied, including:

* **Cropping:** Using the provided bounding box coordinates, the facial region of each image was cropped to focus on the face.
* **Resizing:** Each cropped face was resized to 64x64 pixels to ensure uniformity, making the images suitable for input into a Convolutional Neural Network (CNN).
* **Normalization:** The pixel values were scaled between 0 and 1 by dividing by 255. This standardization step helps accelerate model convergence and enhances training performance.

### 4. Model Architecture

For facial expression classification, a Convolutional Neural Network (CNN) was selected, given its effectiveness in image-based tasks. The architecture includes a combination of convolutional layers, max-pooling layers, and fully connected layers to extract and process image features.

The key components of the model are:

* **Input Layer:** The input size is set to (64, 64, 3) to accommodate the RGB images.
* **Convolutional Layers:** Several layers of convolutions are used to extract hierarchical features from the images.
* **Max Pooling Layers:** Pooling layers are used to downsample the feature maps and reduce spatial dimensions.
* **Fully Connected Layers:** These layers map the extracted features to the final output.
* **Output Layer:** A softmax activation function is applied in the output layer to classify each image into one of the seven expression categories.

### 5. Model Training

The training process employed several strategies and tools to optimize model performance:

* **Data Augmentation:** Techniques such as random rotations, shifts in width and height, zooming, and horizontal flips were applied to the training data. This enhanced the diversity of the data and helped prevent overfitting.
* **TensorBoard:** Used for real-time visualization of the training process, including tracking accuracy and loss over epochs.
* **Optimizer:** The Adam optimizer was used to efficiently minimize the loss function.
* **Loss Function:** Categorical Crossentropy was used, which is appropriate for multi-class classification tasks.
* **Evaluation Metrics:** Accuracy and validation loss were the primary metrics for assessing the model's performance during training.

The model was trained for 20 epochs, and the validation accuracy plateaued at around 40%.

### 6. Results

After training the model, the following outcomes were observed:

* **Training Accuracy:** Approximately 58%
* **Validation Accuracy:** Stabilized at 57%
* **Training Loss:** 1.16
* **Validation Loss:** 1.14

Although the model is capturing general patterns of facial expressions, these results indicate that further tuning and possibly additional data are required to improve performance.

### 7. Confusion Matrix

A confusion matrix was generated to evaluate the model's performance across the different facial expression categories. This matrix provides a detailed view of where the model is correctly classifying expressions and where it struggles, particularly in distinguishing between similar expressions.

### 8. Conclusion

This project successfully implemented a facial expression recognition system using a deep learning approach. While the model shows promise, particularly in capturing general facial expression patterns, there is considerable room for improvement in terms of accuracy and generalization. Future enhancements could involve experimenting with more advanced model architectures, fine-tuning hyperparameters, and exploring techniques such as transfer learning or ensemble models to further boost performance.

**THE END**