

**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: Ammar Ali**

**Email Address:** ammaralidayal51214@gmail.com

**GitHub: https://github.com/DayalAli**

**Linkedln:** https://www.linkedin.com/in/ammar-ali-ds/

Expression Classification Report

# 1. Introduction

This report outlines the project on 'Expression Classification' using the Expression in the Wild (ExpW) Dataset. The dataset comprises 106,000 facial images, and the goal of the project is to train a machine learning model capable of recognizing different facial expressions in real-time. Facial expressions are an important modality for human interaction, and an accurate recognition system can have applications in various fields such as emotion analysis, security, entertainment, and more.

# 2. Dataset Details

The Expression in the Wild (ExpW) dataset consists of 106,000 images annotated with 7 facial expressions:

* **Angry (0)**
* **Disgust (1)**
* **Fear (2)**
* **Happy (3)**
* **Sad (4)**
* **Surprise (5)**
* **Neutral (6)**  
    
  Each image contains metadata with bounding boxes that help locate the face within the image. These annotations are crucial for pre-processing, such as cropping and resizing the facial region before feeding it into a model.

# 3. Data Preprocessing

The preprocessing of the dataset involves several key steps, which include cropping, resizing, and normalization:

* **Cropping**: Using bounding box coordinates provided in the dataset to crop the facial region.
* **Resizing**: The cropped faces are resized to 64x64 pixel dimensions to ensure uniformity and make the dataset suitable for feeding into Convolutional Neural Networks (CNN).
* **Normalization**: Each pixel's value is scaled to be between 0 and 1 by dividing by 255. This step helps the model converge faster by standardizing the input.

# 4. Model Architecture

For this task, we utilized a Convolutional Neural Network (CNN) architecture, which is well-suited for image classification. The architecture includes multiple layers of convolution, max-pooling, and fully connected layers.  
  
The general flow of the model is as follows:

* **Input Layer**: Input size of (64, 64, 3) for RGB images.
* **Convolutional Layers**: Several layers to extract features from the images.
* **Max Pooling Layers**: To reduce the spatial dimensions.
* **Fully Connected Layers**: To map the features to the final output.  
  **Output Layer**: A softmax layer for classifying the images into one of the 7 categories (angry, disgust, fear, happy, sad, surprise, neutral).

# 5. Model Training

The training process was conducted using the following techniques and tools:

* **Data Augmentation**: Applied techniques such as rotation, width and height shifts, zooming, and horizontal flips to increase the variety of the training data and prevent overfitting.
* **TensorBoard**: Used to visualize training progress and track accuracy and loss over time.
* **Optimizer**: Adam optimizer was employed to minimize the loss.
* **Loss Function**: Categorical Crossentropy was used, as the task involves multi-class classification.
* **Metrics**: Accuracy and validation loss were used as key metrics to evaluate the model's performance.

The model was trained for 20 epochs, with the validation accuracy stabilizing around 40%.

# 6. Results

After training the model, the following results were observed:

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

The validation accuracy indicates that the model is learning the general patterns of facial expressions but may require further fine-tuning and more data to reach higher levels of accuracy.

# 7. Confusion Matrix

A confusion matrix was used to evaluate the model's performance across different expression categories. It provides insights into how well the model distinguishes between similar expressions and where it may be misclassifying.  
Below is an example confusion matrix that shows the number of correct and incorrect predictions for each expression class.

# 8. Conclusion

This project demonstrated a successful implementation of facial expression recognition using a deep learning model. While the initial results show potential, there is room for improvement in terms of accuracy and generalization. Future work may include experimenting with more complex architectures, tuning hyperparameters, and using more advanced techniques like transfer learning or ensemble models.

**THE END**