METHODOLOGY

The method by which the accuracy has been achieved is by using the dataset that has been described with an aim through which the approach moves towards selecting the data, processing the data, and training the data to select the appropriate model, which would give a good accuracy and remain appropriate to the problem and extract the features from the given image in order to make a list to help the model, using activation and generalization for the model to properly.

Face Detection

Data preparation begins with the image being collected from the dataset, followed by the conversion of the image into grayscale from RGB, face detection, and image resizing, which is a crucial step in the process. After that process is complete, the flow continues with image segmentation and feature extraction. After the features are extracted from the image, some data is sent to the training, and some data is sent to the testing, where Train/Test is used with the classifier. Then, we can obtain the recognition's final results.

Emotion Recognition

The input image is first shown to the model after data preprocessing, and then convolution layers map the given image's similar features through different layers. Next, max pooling is performed to reduce the image size while assuming that the features will be the same, and finally, we get a fully connected layer that will be useful in emotion detection.

Facial Emotion Recognition using CNN method

A family of deep neural networks called convolutional neural networks (CNNs) are frequently used to analyse visual imagery. Based on the shared weight architecture of the convolution kernels that scan the hidden layers and the properties of translation in variance, they are often referred to as shift invariant or space invariant artificial neural networks. The filters used in the majority of image processing algorithms are frequently developed using heuristics by engineers. CNNs can discover which filtering criteria are most crucial. As a result, we don't require as many parameters, which saves a lot of time and trial-and-error effort. A CNN and a standard neural network differ greatly in that a CNN uses convolutions to handle the math in the background. At least one layer of the CNN substitutes a convolution for matrix multiplication. Convolutions take two functions and produce a function as their output. The input image is initially transmitted into CNN as preprocessed data, therefore the image now goes through various layers. Convolution is the first layer in which the features from the input image are extracted; it uses mathematical operations to take as inputs from the image an image filter and a kernel. When the image is too huge, the pooling layer will then assist in reducing the number of parameters; for this, we can utilise max pooling, average pooling, or sum pooling. Max pooling has been employed in this model. The convolution layer of the image is applied once more, and this time, several methods are used to extract features. For the suggested model, ReLU, Leaky ReLU, and LeakyReLU are used. The Fully Connected Layer (FC), the final stage of the CNN architecture/structure, involves flattening our matrix into a vector and feeding it into a fully connected layer that resembles a neural network.