Emotional Recognition from Facial Expression using Siamese Network

Introduction

In this project, we use one of the meta-learning techniques to recognize emotions using our own in-house dataset. The name of the dataset is “Amrita Emotion Database-2” (AED-2). The age group of the subjects is 20-25. It contains 7 basic emotions of around 8 images of each emotion with a total of 56 images of students showing various emotions like anger, disgust, fear, happy, neutral, sad, and surprise.

Dependencies

- Python 3, Keras, Tensorflow

Basic usage

The repository is currently compatible with TensorFlow-2.2.0 and makes use of the Keras API using the tensorflow.keras library.

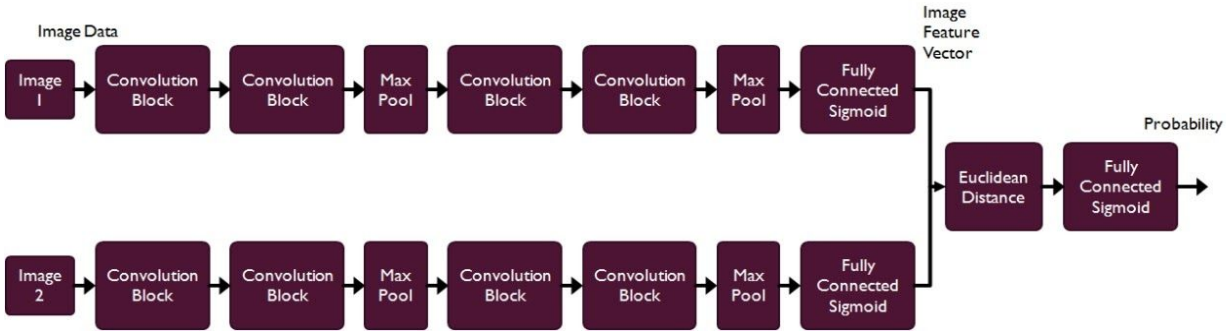
- clone the repository and enter the folder

Data preparation

- Since the sizes of the images are different from one another, we converted each image into a 256x256 grayscale format i.e the portable gray map(.pgm). Thus, the model needs to be trained only on these images to classify and identify the emotion.

- In case you are looking to experiment with new datasets, you may have to deal with data in the .pgm format. We have provided the code we wrote for data preprocessing in the repository which can be used for reference.

Algorithm-siamese network



The model, the siamese network takes data in the form pairs(genuine and imposter). Initially, the function get_data reads the images (img1, img2) from the same directory and stores them in the x_genuine_pair array and assigns y_genuine to 1. Subsequently, the same function reads the images (img1, img2) from the different directory and stores them in the x_imposite pair and assigns y_imposite to 0. Finally, we concatenate both x_genuine_pair, x_imposite to X and y_genuine, y_imposite to Y.

Model building-MLARE

The build which we are using for this project is termed as Meta-Learning Approach for Recognising Emotions(MLARE). We build two convolutional layers with 32 number of filters and kernel_size 3. We assigned border_mode as ‘valid’ to get an output that is smaller than the input because the convolution is only computed where the input and the filter fully overlap. The parameter, dim_ordering we have given as th, since it represents the channel dimension(the depth). input_shape, which represents the shape of the input data or tensor that we feed into the

first convolutional layer. We added rectified linear unit (ReLU) activations in hidden layers and max-pooling of size 2X2 followed by a flat layer with an activation function of the sigmoid.

References

- Wu, Haoran & Xu, Zhiyong & Zhang, Jianlin & Yan, Wei & Ma, Xiao. (2017). Face recognition based on convolution siamese networks. 1-5. 10.1109/CISP-BMEI.2017.8302003.
- H. Wu, Z. Xu, J. Zhang, W. Yan, and X. Ma, "Face recognition based on convolution siamese networks," 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, 2017, pp. 1-5.