Emotional Recognition Using Convolutional Neural Networks

Zaris Ovidiu Neamtiu

January 2023

1 Introduction

1.1 Context

This project attempts to provide a usable piece of software that uses a trained Convolutional Neural Network which performs emotional recognition on the end user based on facial expressions, and then displays the results back to the user.

1.2 Motivation

Thought and Emotional Recognition is one of the next steps in integrating programmed machines in our daily lives. Accurate emotional reading performed by devices can enable them to provide a much more personalised service, tending to users in exactly the way that they need in a seamless manner.

1.3 Objectives

The main objective of this project is to bring forth a possible solution to the need of emotional integration between the end user and a piece of software.

2 Proposed solution

Since this is an image classification issue, the deep-learning model chosen is a CNN. The chosen library for creating the model is the Keras library from Python.

The types of layers comprising the model are the following:

1.) Conv2D: The number of filters used ranges from 32 to 128, and the activation function is ReLU.

2.) MaxPooling2D: Pool size used is (2, 2). This layer is used in order to reduce the spatial dimensions of the input, which in turn reduces the number of parameters in the model and the amound of computation requred.

3.) Dropout: Used in order to prevent overfitting. There are three layers of this type used in the model, the first two with a rate of 0.25, and the last one with a rate of 0.5.

4.) Flatten: Used to flatten the output of the previous layers, so that it can be fed into a dense layer.

5.) Dense: There are two layers of this type in the model. The first one, with 1024 neurons and a ReLU activation function, enables the model to learn more complex representations of the input image and also to reduce the dimensionality of the input, in order to make it more manageable for the final output layer. The second and last

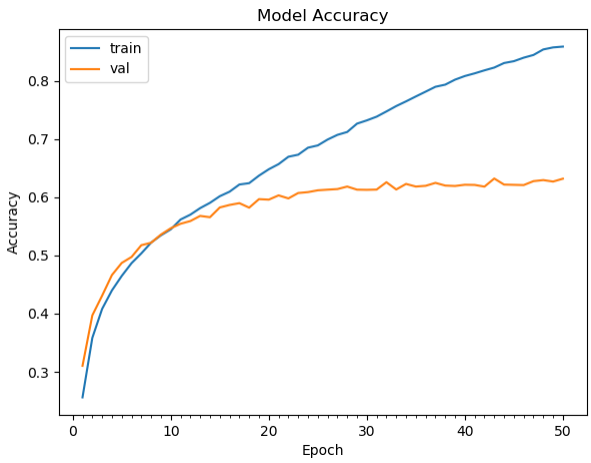
one, with seven neurons, is used as the output layer. Since the model preforms image classifcation, the activation function used is "softmax". This enables the last layer to provide probabilities that an input photo belongs to one of seven classes.

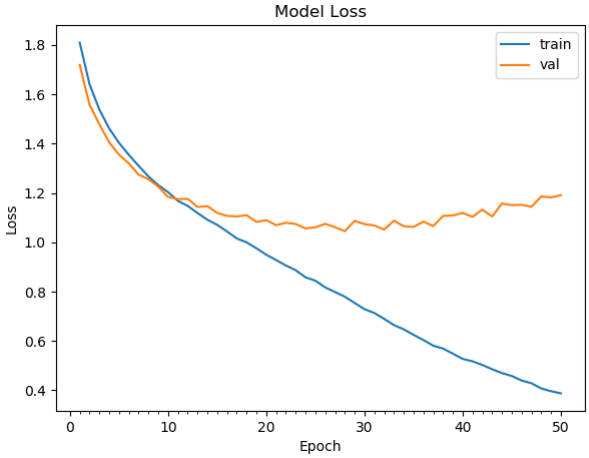
The data set used is the FER-2013 data set. The data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centred and occupies about the same amount of space in each image. The training set consists of 28,709 examples and the public test set consists of 3,589 examples.

The haar cascade method is used in order to detect faces in each frame of a supplied video or coming straight from the webcam feed.

3 Results

The model performance can be observed on the following plots.





This implementation by default detects emotions on all faces in the webcam feed. With a simple 4-layer CNN, the test accuracy reached 63.2 percent in 50 epochs.

4 Conclusions

In the end, the proposed model can somewhat read human emotions by interpreting facial expressions. Of course, the accuracy of the results depends strongly on the quality of the video input it gets, and is also closely tied to how the analyzed person looks based on ethnicity and other genetic physical traits. As far as the "seamless integration between end user and software based on emotions" goes, there are definitely improvements to be made before considering using this solution in any sort of software that flows based on the emotions of the end user.

4.1 Extension ideas

A fun idea to consider when thinking about how such a model as this one could be used is to integrate it in video games. Players would be delighted to hear in-game characters switch dialogue lines, or even witness a switch in the

entire story line of the game, based on the facial expression of the player at a certain moment.

Another, more serious application could be training a model that looks for fluctuations in emotions that are indicative of mental illnesses. Such a model would still need a specialized official closely monitoring the situation, as diagnosing mental illnesses is a rather ambiguous field, but AI software can be of great assistance.

4.2 References

Github - petercunha - Emotion

Github - akmadan - Emotion Detection CNN

Github - Tomislav-Troha - Facial-Emotional-Recognition-Using-Keras

Real-Time CNN for Emotion and Gender Classification by Octavio Arriaga, Paul G. Ploger and Matias Valdenegro

Github - datamagic2020 - Emotion detection with CNN

Github - atulapra - Emotion detection