Project Proposal

COS475
Machine Learning
Miguel Guarniz

Project Title: Facial Expression Recognition

Project Idea

Facial expression recognition technology is becoming more relevant in our society. This technology is being used in different applications such as detecting the levels of pain of post-surgery patients and determining how consumers react to advertising [4].

We have decided to implement a training model that will allows us to recognize human facial expressions given an image. More specifically, our model will need to classify a facial expression into one of the seven classes which are angry, disgust, fear, happy, sad, surprise and neutral.

Solution

Our approach will be based on Convolutional Neural Networks (CNN). We will develop CNNs with different levels of depth to measure their performance for this classification task. We will be using the network architecture proposed in [1]:

[Conv-(SBN)-ReLU-(Dropout)-(Max-pool)]M - [Affine-(BN)-ReLU-(Dropout)]N - Affine - Softmax.

We will develop different models by choosing different values for M and N; this will allows us to experiment with different layers. Furthermore, the architecture allows us to control the inclusion of batch normalization, dropout and max-pool layers in our models.

In addition, before training our model, we will use a couple of preprocessing steps, such as face alignment and pose normalization, to prepare our data. This will remove variations that are not relevant to our training. Our aim is that these steps will improve accuracy.

Data

We will be using a dataset from the Challenges in Representation Learning: Facial Expression Recognition Challenge page in Kaggle.com [2]. The data was provided by Pierre-Luc Carrier and Aaron Courville as part of a research.

The data consists of face images. The sample set is made up of 35,887 examples. Each example consists of a 48x48 pixel image and a "emotion" label. The data set file is formatted in two columns, emotion and pixels. The emotion column contains a value from 0 to 6 and the pixel column contains space separated pixel values.

Out of the sample set, around 28,000 images will be used for our training, around 4,000 will be used for cross-validation and around 4,000 will be used for testing.

References

1. Alizadeh, Shima, and Fazel, Azar. (2017). Convolutional Neural Networks for Facial Expression Recognition. Retrieved from: https://arxiv.org/pdf/1704.06756.pdf.

2. Anonymous. (2013). Challenges in Representation Facial Expression Recognition data. Kaggle.com. Retrieved from: https://www.kaggle.com/c/challenges-in -representation-learning-facial-expression-recognition-challenge.

3. Li, Shan, and Deng, Weihong. (2018). Deep Facial Expression Recognition: A Survey. Retrieved from: https://arxiv.org/pdf/1804.08348.pdf.

4. Thomas, Daniel. (2018). The cameras that know if you're happy - or a threat. Retrieved from: https://www.bbc.com/news/business-44799239.