Facial Emotion Recognition Using CNN

By: Robbel Habtamu, Tibebu Asmare and Zehara Eshetu

Addis Ababa Institute of Technology

Abstract. Now a day technology grow rapidly and it affect life of people in many area.one of these emerging technology is Artificial intelligence. Our project is entitled on facial emotion recognition(FER) using CNN. Now a day frauds become increase every where.Deep learning handle these issue by processing big data. So the main Goal of our project set out to recognise emotion of human being using convolutional neural network. The data used in these project is from one of biggest public dataset repository Kaggle. First we preprocess data and avoiding data imbalance by appling data augmentation. Secondly we build four layer Convolutional neural network architecture with two fully connected layer. We used ReLU activation function for each layer and finally we apply Softmax algorithms on the last layer of the model. Despite the good performance of these project, there still seems to be great room for improvement.
https://github.com/R-hab/cvproject.git

Keywords: Deep Learning, Convolutional Neural Network,FER

1 Introduction

Our life world is drastically influenced by technology[1]. Technology has advanced into many areas of our lives, making information more accessible, improving communication, fraud detection, authentication. Deep learning is key solution for every walk of our life challenges.Facial expressions are one of the main information channels in interpersonal communication. Therefore, it is natural that facial emotion research has gained a lot of attention over the past decade with applications in perceptual and cognitive sciences.

Over the last few years, there is an increase in the desire to use deep learning tools to understand human behavior patterns. One of the most trendy challenges in this field is facial expression recognition.Facial emotions are important factors in human communication that help to understand the intentions of others. In general, people infer the emotional state of other people, such as happyness,sadness and anger, using facial expressions.

The main goal of these project is to recognise human facial emotion using convolutional neural network. we used fer2013 public dataset from kaggle repository.we concern on preprocess the data for better performance. But these dataset still needs bit more preprocessing. because the dataset set contain noisy image and also occculusion usually(hand) as well as glass, beside these it contain imbalanced data which affect model performance.

2 Related works

Earlier works on emotion recognition, rely on the traditional two-step machine learning approach, where in the first step, some features are extracted from the images, and in the second step, a classifier (such as neural network, SVM, or random forest) is used to detect the emotions. Some of the popular hand-crafted features are used for facial expression recognition. Then, a classifier would assign the best emotion to the image. These approaches seemed to work fine on simpler datasets, but with the advent of more challenging datasets (which havemore intra-class variation), they started to show their limitation.

Despite the notable success of traditional facial recognition methods through the extracted of handcrafted features,over the past decade researchers have directed to the deep learning approach due to its high automatic recognition capacity. In this context, we will present some recent studies in FER, which show proposed methods of deep learning in order to obtain better detection. Train and test on several static or sequential databases.

Nowadays Researchers are working to solve complex problems using most progressive and rapidly growing computer vision applications.First challenge in facial emotion recognition of facial expressions, held during the IEEE conference on Face and Gesture Recognition 2011[2].

Facial communication studies have been carried out for years. But for any experiment, there was still room for progress. That is why this topic is convenient. The key objective of the researchers is to enhance the precision of a basic data collection FER2013.

Lopes et al. Studied the impact of data pre-processing before the training the network in order to have a better emotion classification. Data augmentation, rotation correction, cropping, down sampling with 32x32 pixels and intensity normalisation are the steps that were applied before CNN, which consist of two convolution-pooling layers ending with two fully connected with 256 and 7 neurons. The best weight gained at the training stage are used at the test stage. This experience was evaluated in three accessible databases: CK+, JAFFE, BU-3DFE. Researchers shows that combining all of these pre-processing steps is more effective than applying them separately[3].

Ekman et al. defined seven basic emotions, irrespective of culture in which a human grows with the seven expressions anger, feared, happy, sad, contempt, disgust, and surprise[4].Most of the existing approaches are based on engineered features (e.g. HOG, LBPH, and Gabor) where the classifier's hyperparameters are tuned to give best recognition accuracy's across a single database, or a small collection of similar databases[5].

Facial emotion recognition have application on many areas like, Robotics, Medicine, Deriving assistance system, Lie detection and so on. Recently it develops in area of neuroscience and cognitive science. further more it attract researchers attention in computer vision. And day to day researcher gets great accuracy using deep learning.

The efficiency of CNN-based image recognition methods has also been continuously improved and has gradually replaced the traditional facial expression

image recognition methods. In the 2012 Image Net Object-Oriented Recognition Challenge, some scholars used the CNN model Alex Net to win the championship. In the 2014 ILSVRC competition, the CNN model Google Net architecture won first place in the classification[6].

Recently, researchers have made extraordinary accomplishment in facial expression detection[7], which led to improvements in neuroscience and cognitive science that drive the advancement of research, in the field of facial expression.

All of the above works achieve significant improvements over the traditional works on emotion recognition, but they are using the raw data as the input. In this work, we try to modify the data, such as we only use the vital data from the image.

Generally, As we observe from many researcher in computer vision, they are interested in image processing. specially facial emotion is still open for scholar's. This is because Image pose difficulty i.e a person's different(side) position make it difficult to predict emotion of person. However researcher day to day gets improved accuracy.

3 Dataset

In these project we used Public dataset from kaggel FER-2013.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 task is to categorize each face based on the emotion shown in the facial expression into one of seven categories (0=Angry, 1=Disgust, 2=Fear, 3=Happy, 4=Sad, 5=Surprise, 6=Neutral). The training set consists of 28,709 examples and the public test set consists of 3,589 examples.

dataset problems and our solution for the challenge.The first problem is occlusion, when the image or part of image is covered by hand or glass. this kind of image must be removed from dataset because the model will not recognise the emotion.The second problem, is data imbalance so data class must be balanced or not to be exaggerated variation.for this problem we used Data augmentation.

The other problem is contrast variation these can be too dark or too light and this can be affect CNN model performance. These is solved by changing the focus of image to face. The other one is Intra-class variation, which means some image in dataset are not human face as there are drawings and animated face. so we have checked the presence of these kinda image.

Image-1 sample data from dataset

4 Methods

For these project we used deep learning, CNN architecture to extract feature's of image to classify Emotion of human. A CNN is a Deep Learning algorithm which takes an input image, assigns importance, learnable weights and biases to various aspects/objects in the image and is able to differentiate between images.

The proposed Architecture we have used is CNN architecture Model. the Public data sets we have used contains 7 class facial emotion. First we load the data set and display some sample images.In order to examine the effects of the size of the training set, the data set was divided into different sized training and testing sets. Increasing the size of the training set has positive effect on the result up to a certain level. So we split the data 80 by 20 for training and testing respectively.

We have done preprocessing opration after we load the data. since the dataset contain noise, some light and also color variation it needs preprocessing step inorder for better performance for our CNN model.

Based on our knowledge with classifying facial expressions regardless of skin color, facial expression recognition using CNN must also be conducted independently of the color information of the input image. So in preprocessing step we have normalized the image in order to avoid feature difference between image and produce improved image of face.

After we normalize, we convert images to one channel or to gray scale color. This is because of colored image are difficult to process. Then we reshape image dimension and avoid unnecessary portion of image.This reduce using much memory size and increase running speed. And also We will use Haar classifier which trained by Haar-like small features and it will used texture descriptor, and its main features are linear, edge, canter, and diagonal.finally, we build our own CNN model alongside feature extraction and pooling to down sample the feature of image.

4.1 Architecture

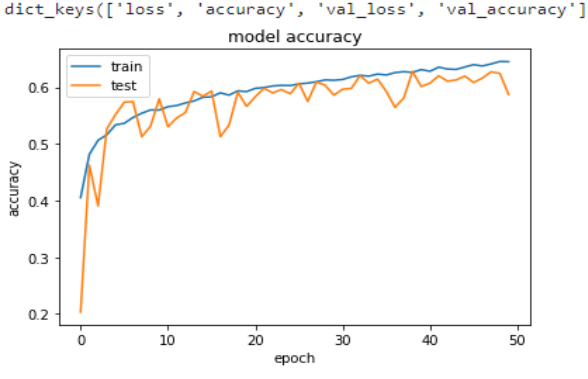
This project is performed using CNN.Feature extraction part consists of four convolutional layers, each four layer are followed by rectified linear unit (ReLU) activation function, and after the first layer, we use Batch Normalization. They are then followed by a dropout layer and two fully-connected layers, each two followed by a ReLU activation, then, another dropout layer, which ends with another fully connected layer. From that, we extract feature vectors, which we concatenate, and then send to a similar network consisting of a dropout layer and two fully-connected layers, each two followed by a ReLU activation, then, another dropout layer, which ends with another fully connected layer.

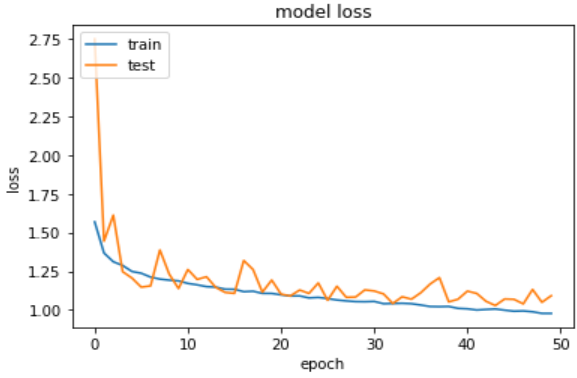
This model is then trained by optimizing a loss function using RMSprop optimizer. The loss function in this work is categorical cross entropy loss. Adding dropout enables us to train our models from scratch even on very small datasets.

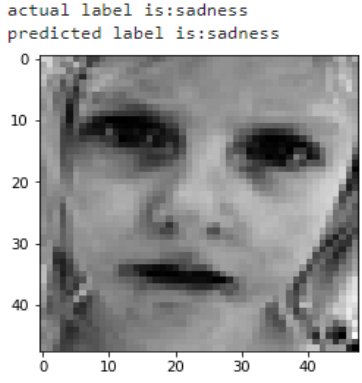
5 Experiment Result

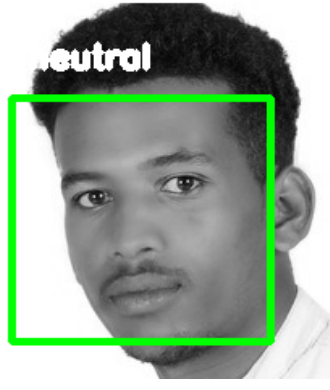
Intially we used 48x48 input image and then we resize it. For the FER2013 dataset, we used 28709 images for training, 3589 for validation, and 3589 for testing. Each image we randomly assigned to the train, validation, or test set. We were able to achieve an accuracy rate of around 65 percent. The train and validation loss and accuracy are reported in Figure below.

Some Predicted Results:




model loss
train
test
loss
epoch
2.75
2.50
2.25
2.00
1.75
1.50
1.25
1.00
0
10
20
30
40
50


actual label is:sadness
predicted label is:sadness
0
10
20
30
40


eutral

6 Conclusion

This project is about facial emotion recognition. We use CNN model to develop the project. As we observerd from this project the dataset we used has noisy which hinder or decrease model performance. so it needs much preprocessing stage. As of now we didn't cover all issue. for the future we planned to use pytorch deep learning frame work and improve performance of recognizing facial expression of human. because this area need attention now a day to overcome fraud everywhere.

References

1. Stolterman, E., Fors, A.C.: Information technology and the good life. In: Information systems research. Springer (2004) 687–692
2. Mellouk, W., Wahida, H.: Facial emotion recognition using deep learning: review and insights. Procedia Computer Science **175** (01 2020) 689–694
3. Lopes, A.T., De Aguiar, E., De Souza, A.F., Oliveira-Santos, T.: Facial expression recognition with convolutional neural networks: coping with few data and the training sample order. Pattern recognition **61** (2017) 610–628
4. Ekman, P.: Are there basic emotions? (1992)
5. Mollahosseini, A., Chan, D., Mahoor, M.H.: Going deeper in facial expression recognition using deep neural networks. In: 2016 IEEE Winter conference on applications of computer vision (WACV), IEEE (2016) 1–10
6. Pak, M., Kim, S.: A review of deep learning in image recognition. In: 2017 4th international conference on computer applications and information processing technology (CAIPT), IEEE (2017) 1–3
7. Mehendale, N.: Facial emotion recognition using convolutional neural networks (ferc). SN Applied Sciences **2**(3) (2020) 1–8