

Facial Modalities Recognition and Manipulation

N Ramesh Babu ¹

¹Associate Professor, Dept. of CSE
Amruta Institute of Engineering and Management Sciences,
Bidadi, Karnataka, India-562109

Indumathi K ²

²UG Scholar, Dept. of CSE
Amruta Institute of Engineering and Management
Sciences, Bidadi, Karnataka, India-562109

Bhavana J ²

²UG Scholar, Dept. of CSE
Amruta Institute of Engineering and Management Science,
Bidadi, Karnataka, India-562109

Bharath A ²

²UG Scholar, Dept. of CSE
Amruta Institute of Engineering and Management
Sciences, Bidadi, Karnataka, India-562109

Dhananjaya K S ²

²UG Scholar, Dept. of CSE
Amruta Institute of Engineering and Management
Sciences, Bidadi, Karnataka, India-562109

Abstract:- Facial expression recognition has been a challenge for many years. With the recent growth in machine learning, a real-time facial expression recognition system using deep learning technology can be useful for an emotion monitoring system for Human-computer interaction (HCI). We proposed a Personal Facial Expression Monitoring System (PFEMS). We designed a custom Convolutional Neural Network model and used it to train and test different facial expression images with the Tensor Flow machine learning library. PFEMS has two parts, a recognizer for validation and data training model for data training. The recognizer contains a facial detector and a facial expression recognizer. The facial detector extracts facial images from video frames and the facial expression recognizer distinguishes the extracted images. The data training model uses the Convolutional Neural Network to train data and the recognizer also uses Convolutional Neural Network to monitor the emotional state of a user through their facial expressions. The system recognizes the seven universal emotions, angry, disgust, happy, surprise, sad and fear, along with neutral.

Keywords:- Convolutional Neural Network, Deep Learning, Emotion Recognition, Classifiers.

1. INTRODUCTION

Human emotion recognition plays an important role in the interpersonal relationship.

The automatic recognition of emotions has been an active research topic from early eras. Therefore, there are several advances made in this field. Emotions are reflected from speech, hand and gestures of the body and through facial expressions. Hence extracting and understanding of emotion has a high importance of the interaction between human and machine communication. This paper describes the advances made in this field and the various approaches used for recognition of emotions. The main objective of the paper is to propose real time implementation of emotion recognition system. Deep learning is one of the fastest-growing and most exciting areas in machine learning. With recent advancements in graphics processing unit, it is

possible to use Deep Learning for real-time applications. Emotions are an incredibly important aspect of human life, and play an important role in human interaction. Facial expressions represent the emotion of a person and it can give an indication of the emotional response of a person to the interaction with a computer. So detecting facial expressions can help create a better (i.e. less frustrating) user experience.

2. MOTIVATION

In the 21st century, HCI products, such as Siri from Apple, Echo from Amazon and Cortana from Windows, became more and more popular in the world. The recent successes of AlphaGo brought machine learning to the world. AlphaGo uses a Monte Carlo tree search algorithm to find its moves based on the knowledge gathering from a pre-train data, which trained by artificial neural network (ANN).

The successful use of machine learning in Go (game) encourages us to design a facial expression recognition system that can be used for HCI and solve facial expression recognition problem with machine learning.

3. LITERATURE REVIEW

Research in the fields of face detection and tracking has been very active and there is exhaustive literature available on the same. The major challenge that the researchers face is the non-availability of spontaneous expression data. Capturing spontaneous expressions on images and video is one of the biggest challenges ahead. Many attempts have been made to recognize facial expressions. Zhang et al investigated two types of features, the geometry-based features and Gabor wavelets based features, for facial expression recognition.

Mihai Gavrilescu developed a fully integrated neural-network based facial expression recognition system based

on FACS [13]. He used 27 out of 44 action units (AUs) from FACS to detect changes in localized facial features. His method separated face feature into 5 sub features: eye features, brow features, cheek features, lip features and wrinkle features. And then he used a classifier to classify the AU for each component. The output of the classifiers will feed to a 4-layer neural network which will take the decision of the final AU map and the corresponding recognized emotion. In terms of AU classification, the Classification Rate is 0.981 with MMI database [25] and 0.978 with Cohn-Kanade Facial Expression (CK) database [18].

James Lien, et al. developed a computer vision system that automatically recognizes individual action units or action unit combinations in the upper face using Hidden Markov Models (HMMs) in 1998 [20]. Their approach to facial expression recognition is based on FACS. They separate expressions into upper and lower face action and use the upper face for the recognition. They use three approaches to extract facial expression information on the upper face: (1) facial feature point tracking, (2) dense flow tracking with principal component analysis (PCA), and (3) high gradient component detection. The recognition results of the upper face expressions using feature point tracking, dense flow tracking, and high gradient component detection are 85%, 93%, and 85%, respectively.

4. PROPOSED SYSTEM

We develop a system that recognizes facial expression can be used to determine the emotion state of a user there by creating a personal facial expression monitoring system (PFEMS). We accomplish this by design a CNN model for recognized facial expression of the seven universal emotions including neutral using Tensor Flow. We designed and run experiments to determine CNN depth and image resolution required for effective recognition. PFEMS is able to train images, detect face from real time video and recognize facial expression. Here the facial images from the dataset is used and is kept in separate folder because we can easily train the model and predict the model if it is in this type and the trained data is separated into two ways: one for validation and another for verification that is into training and testing data that to in the 80:20 ratio. We use a openCv library for the recognition of emotion in real time.

5. OBJECTIVE

In this work, the main goal is to design a python-based Personal Facial Expression Monitoring System (PFEMS). PFEMS uses a custom Convolutional Neural Network (CNN) model which is used to train facial expression images with the TensorFlow machine learning library.

PFEMS can be used to detect the change of human facial expression during the interaction to allow changing how a system responds to the user. There are two parts in PFEMS. One is a training program. Since Convolutional Neural Network technique requires at least hours for

training, training phase needs to be separate from the front end monitoring system. We design a proper Convolutional Neural Network for PFEMS. The other one is facial detection and recognition program. This program accesses an HD/FHD camera for input video and applies facial detection on video frames. This tests and identifies the emotion from images captured in real time.

6. BACKGROUND

Dataset FER 2013:

The both training and evaluation operations would be handled with Fer2013 dataset. Compressed version of the dataset takes 92 MB space whereas uncompressed version takes 295 MB space. There are 28K training and 7K testing images in the dataset. Each image was stored as 48 X 48 pixels.



Fig 6.1: FER 2013 images

CNN (Convolutional Neural Networks)

CNN is a type of feed-forward ANN in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. It is attractive for many deep learning tasks like image classification, scene recognition, and natural language processing.

1. Covolution 2D Layer:

The 2D convolution is a fairly simple operation we start with a kernel, which is simply a small matrix of weights.

This kernel “slides” over the 2D input data, performing an element wise multiplication with the part of the input it is currently on, and then summing up the results into a single output pixel. Convolution is the first layer which is used to extract features from an input image. It preserves the relationship between pixels by learning image features using small squares of input data

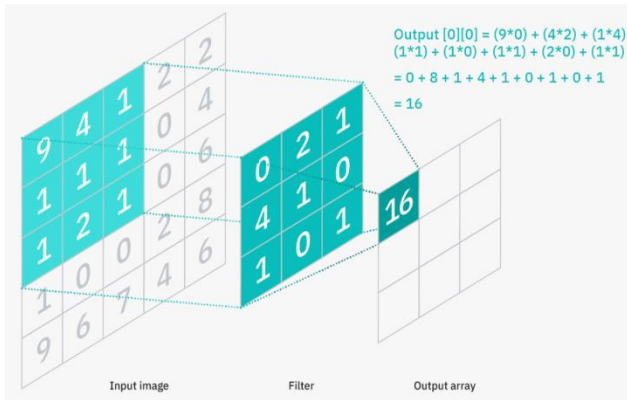


Fig 6.2: Convolution layer

2. Pooling layer:

Maximum value is taken from each stride and stored in a new matrix. Depending on the stride Pooling is of two types Max Pooling and Minimum Pooling. When the stride is large such Pooling is known as Max Pooling whereas small stride is known as Minimum Pooling. In case of Max Pooling, a kernel of size n*n is moved across the matrix and for each position the max value is taken and put in the corresponding position of the output matrix. It considers the highest values and marks the lowest values as 0.

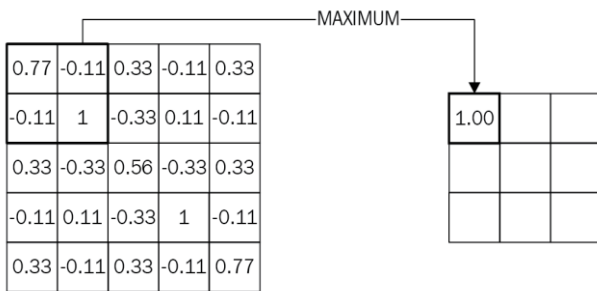


Fig 6.3: Pooling Layer

3. RELU(RECTIFIED LINEAR UNIT LAYER)

ReLU refers to the Rectifier Unit, the most commonly deployed activation function for the outputs of the CNN neurons. In ReLU layer the pixels which are not necessary are deactivated and only the important pixels are kept.

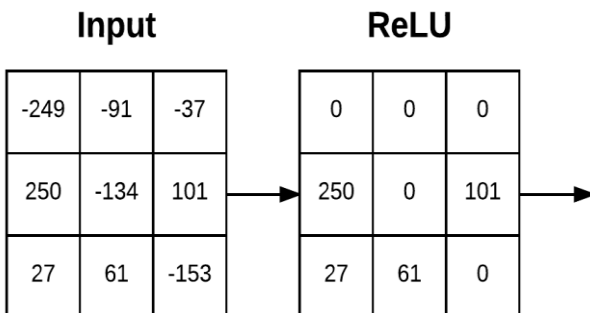


Fig 6.4: Relu Layer

4. FULLY CONNECTED LAYER:

The Fully Connected Layer in CNN has neurons that are fully connected to all the neurons of the previous layer. Multiple FC layers are stacked as per the architecture used. It is often the last layer used in CNN which is responsible to predict the output or the label of the input class. Different activation functions are used like SOFTMAX which is used to classify multi-class problems.

5. NORMALIZATION LAYER:

Batch normalization is a layer that allows every layer of the network to do learning more independently. Using batch normalization learning becomes efficient also it can be used as regularization to avoid over fitting of the model. The layer is added to the sequential model to standardize the input or the outputs.

7. SYSTEM DESCRIPTION

PFEMS contains two parts:

1. Pre-train program: The pre-train program uses a CNN deep learning training model.
2. Recognizer: a recognizer for validation and a pre-train program for data training. The recognizer contains a facial detector and a facial expression recognizer.

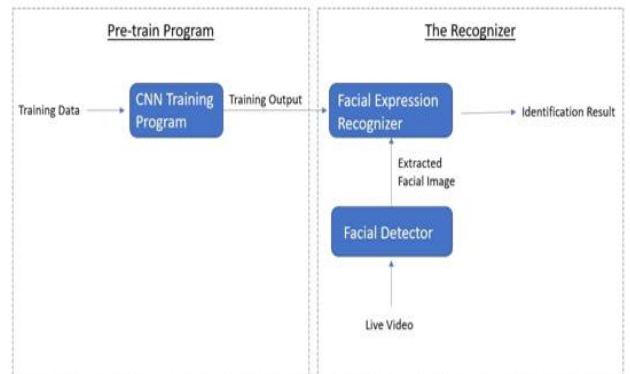


Fig 7.1: System Structure

The pre-train program uses a set of labeled facial images to train on a CNN deep learning training model. The recognizer uses an RGB camera to obtain real time video. The facial detector detects the largest face and crops a facial image from the frame. The facial image is used as input for the facial expression recognizer. The inputs of the recognizer are an RGB real time video from a camera and checkpoint file from the pre-train program. The output of the facial detector is a fixed size image which is an input for the facial expression recognizer that uses it for the identification of the facial expression. The facial detector reads frames from the live video and uses OpenCV's cascade classifier to detect faces from the frames. While faces are being detected, a square facial image of the largest face is extracted from frame.

8. METHODOLOGY

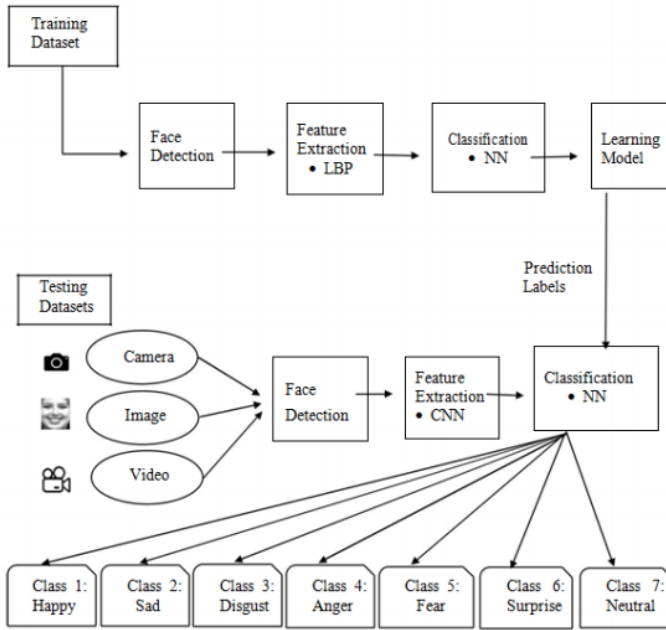


Fig 8.1: System design

The system diagram describes developing training and testing model. The training datasets is given and undergoes detection, extraction and classification and builds a training model. Next testing datasets (either from camera, image or captured video) undergoes through the same phases and training model is incorporated, is passed to neural networks classifies the emotion to either of the classes.

8.1 Training Model

Training flowchart describes the flow of the designing and implementation.

The image is preprocessed then the features are extracted and these images are processed through the neural network and a training model is built.

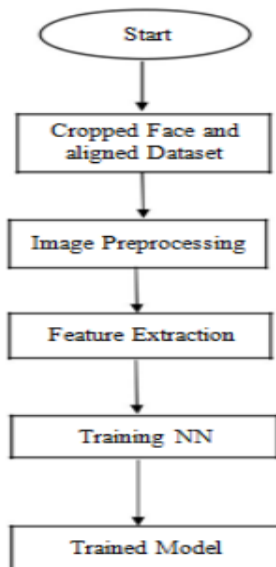


Fig 8.2: Flowchart of Training

8.3 Testing Model

The images are pre-processed and features are extracted a built training model is included to neural network to increase the stability and compatibility of the model. Then the system classifies the emotions to the specified classes.

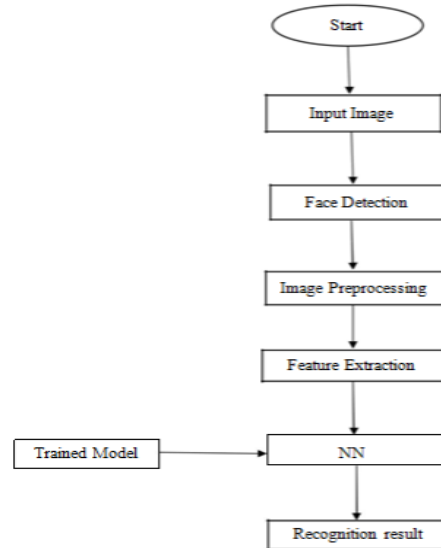


Fig 8.3: Flowchart of Testing/Prediction

Image Acquisition:

Images used for facial expression recognition are static images or image sequences. Images of face can be captured using camera.

Face detection

Face Detection is useful in detection of facial image. Face Detection is carried out in training dataset using Haar classifier called Viola-Jones face detector and implemented through Opencv. Haar like features encodes the difference in average intensity in different parts of the image and consists of black and white connected rectangles in which the value of the feature is the difference of sum of pixel values in black and white regions.

Image Pre-processing

Image pre-processing includes the removal of noise and normalization against the variation of pixel position or brightness.

- a) Color Normalization
- b) Histogram Normalization

Feature Extraction

Selection of the feature vector is the most important part in a pattern classification problem. The image of face after pre-processing is then used for extracting the important features.

Emotion Specification

An emotion is identified after extracting the features and particular emotion is specified belonging to seven classes. OpenCV cascade classifier can process fast enough to meet our need. The facial detector reads frames from the live video and uses OpenCVs cascade classifier to detect faces from the frames.

9. RESULTS & ANALYSIS

The proposed system uses CNN model and FER 2013 dataset for training and produces 90% accuracy in training and 66% accuracy in validation in training the model. Facial expression recognizer reaches 100% accuracy on disgust expression. Happy and sad expression had over 80% accuracy with or without error images. Surprise had 78.57% accuracy without error, but only 61.1% accuracy with the error image. Neutral is the most difficult expression for the facial expression recognizer, only 45.8% with the error and 47.8% without error image. There are some confirmed factors that reduce the accuracy, such as incorrect cropped images and low quality cropped images. The incorrect cropped images definitely decrease the accuracy of the facial expression recognizer.

10. CONCLUSION

The Facial Emotion Recognition System presented in this project work contributes a resilient face expression recognition model based on the mapping of behavioral characteristics with the physiological biometric characteristics. The facial detector has an acceptable cropping success rate with front face which is enough for the system to use, but the error rate still has room for improvement. The facial expression recognizer had 73.8% accuracy on fixed user, but the facial expression recognizer had a hard time handling an expression while it is in transition.

For the future improvement, PFEMS needs a larger dataset for training and more samples for testing. The CNN model can be improved by making it deeper and wider, but the processing time will be longer. The facial detector can be improved by adding eye and mouth detector. Detecting eye and mouth from the cropped images can double check the result from the cascade classifier.

REFERENCES

- [1] Facial Emotion Recognition using Convolutional Neural Networks by Ninad Mehendale published on 18 Feb 2020.
- [2] A. DHALL, O. V. R MURTHY, R. G. . . J., ANDT.GEDEON. Video and image based emotion recognition challenges in the wild: EmotiW 2015. In the 2015 ACM on international Conference on Multimodal interaction (2015), ACM, pp. 423–426.
- [3] ABADI, M., AGARWAL, A., BARHAM, P., BREVDIO, E., CHEN, Z., CITRO, C., CORRADO, G. S., DAVIS, A., DEAN, J., DEVIN, M., ET AL. Tensor-flow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016).
- [4] A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing by Hongli Zhang, Alireza Jolfaei, and Mamoun Alazab Published in October 2019.
- [5] GUO, Y., TAO, D., YU, J., XIONG, H., LI, Y., AND TAO, D. Deep neural networks with relativity learning for facial expression recognition. In Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on (2016), IEEE, pp. 1–6.
- [6] Facial Emotion Analysis using Deep Convolution Neural Network Rajesh Kumar G A1, Ravi Kant Kumar2, Goutam Sanyal3
- [7] My first CNN project – Emotion Detection Using Convolutional Neural Network with TPU.
- [8] CHAN, K.-P., ZHAN, X., AND WANG, J. Facial expression recognition by correlated topic models and bayes modeling. In Image and Vision Computing New Zealand (IVCNZ), 2015 International Conference on (2016), IEEE, pp. 1– 5.
- [9] CHENG, F., YU, J., AND XIONG, H. Facial expression recognition in jaffedataset based on gaussian process classification. IEEE Transactions on Neural Networks 21, 10 (2010), 1685–1690
- [10] GAVRILESCU, M. Proposed architecture of a fully integrated modular neu-ral network-based automatic facial emotion recognition system based on facial action coding system. In Communications (COMM), 2014 10th International Conference on (2014), IEEE.
- [11] Real Time Facial Expression Recognition using Deep Learning by Isha Talegaonkar, Kalyani Joshi, Shreya Valunj, Rucha Kohok, Anagha Kulkarni. Published article in ICCIP 2019.
- [12] Facial Expression Recognition via Deep Learning October 2017 by Abir Fathallah, Lotfi Abdi, Ali Douik.