

Emotion Recognition from Facial Expression

Er. Harjasdeep Singh¹, Anurag Srivastva², Jaiprakash Singh²

1. Assistant professor, Department of CSE, MIMIT, Malout

2. Students, MIMIT, Malout

ABSTRACT--- Facial expression detection is comprised of Facial Recognition further it is gaining importance rapidly. There are number of methods for identifying emotions by the use of Machine Learning and Artificial Intelligence techniques, in this research we are going to use deep learning and image classification techniques to identify the emotions and does the classification of the images according to the images. Different types of datasets are used to experiment and investigate for training the machine the emotion recognition are to be explored in this research paper.

I. INTRODUCTION

Human expressions and emotions capturing has been a revolution in the field of research and it also catch the attention of researchers, the ability of computer to recognize the emotion or expressions is often called as human-computer interaction. Furthermore, after ordaining with an augmented and developing human communication, by enhancing the emotional intelligence of the beings. There are diverse types of techniques available to investigate or find the human expressions, ranging from facial expressions, human gestures, voice frequency etc. This paper will explore facial expression recognition. Facial Emotion Recognition(FER) is fast-growing research area, many advancements such as automatic translation system, interaction between machine and humans are done in the government and private sectors. The following research paper will emphasize on the survey and take review of variety facial processing features, from the emotional datasets, classifiers algorithms. In this paper is divided into many sections, Section 2 describes the supporting information about the emotion recognition system and also applications of emotion detections. Furthermore, the section 3 explains the feature selection methods and also image optimizations

techniques. Section 4 compares about variety of Facial emotional database. Section 5 will explore various classifier algorithms for differentiating images on basis of emotions and expressions identified. Lastly, in section 6 the paper will be concluded.

II. BACKGROUND INFORMATION

A- Emotion Recognition

The process of recognizing variety of emotions and telling exact emotion is called as Emotion recognition, there are various technology that identifies your emotions from the facial expression you give to the machine. Evolution in the field of technology specifically Machine learning and Artificial Intelligence, made the emotion recognition much easier that before. It seems to be that emotions have to be the most used medium of communication with computers in near future. And this will increase the requirement of automatic Facial expression precognition rapidly. Research work in the area focuses on finding the expressions form the videos or any auditory information. Majority of research work acknowledges and matches but only some of them have used the convolutional neural networks that pervade emotions from images.

Emotional recognition is the field which deals with the evaluation of emotions and identifying it, there are variety of techniques and methods used for this identification. Diversity of the methods are there to deduce the emotions such as, artificial intelligence, emotional intelligence, neural networks, machine learning. Many problems are being solved as the importance of emotion recognition is increasing rapidly. The images are given as input to the system and then facial expression evaluates the emotion in the image.

B- Facial Expression Recognition

Facial Expression Recognition is rapidly growing research area, it is the task of differentiating the facial expressions on input image into number of categories such as happiness, fear, sadness, anger and so on. FER is very diverse in nature. Detecting emotions is not an easy job but it requires to focus on each and every parameter to get the accurate and precise emotions. Thus, survey is to be done to explore facial emotion recognition.

Artificial Intelligence is used for detecting various emotions machine is tending to learn all the emotions and what of the different facial expressions means. Further it applies its previous experience to recognize present emotion. This AI based system can read, interpret, imitate and also respond to the facial expressions of the human being.

C- Applications of Facial Emotion Recognition

FER is used for payments, user doesn't even want a card for withdrawal but FaceTech uses the facial technology for seamless payment. FER is used for criminal identification, if FER can be used for keeping unauthorized people away from the services, then it can be also used for criminal identification. FER can be used in Access and security, instead of using passcodes and patterns biometrics can be used to log in to the device or physical locations.

D- Deep learning

Various deep learning algorithms and methods are used to grasp large amount of datasets of the different faces with every emotion, and are able to identify various variety of faces and emotions. There was a method that was designed in 1991 called as Face recognition using Eigen faces. This is an emanate technique has evolved the research field of FER that has been launched by the development of Deepface and DeepID techniques.

Deep learning has a widespread application in image recognition with the help of neural networks. There are much more techniques like multimodal deep learning that is used in the field of image recognition, selection etc.

III. LITERATURE STUDY

This paper is basically being a research to run face emotion recognition on browser by implementing several CNNs (Convolutional Neural Networks) to solve face detection, face recognition, face landmark detection and tiny face detection optimized for the web browsers and for mobile devices. We implement CNNs using some image datasets from different sources. Following paper will investigate about the reactions found by the images of the emotions, by the means of variety of classifiers. There are different types of algorithms that are used to classify these emotions, like K-Nearest Neighbor, Random Forest. Facial Emotion Recognition has great importance in the field of research, FER has been analyzed and studied in all the areas by the researchers. Highly accurate technique to recognize emotion by using the filter banks and Deep CNN and another reason to use this technique that it has high accuracy and we will also know that Deep learning is the way we can use the emotion detection. In this paper we have used image spectrogram in combination with deep convolution networks. This paper explores various types of database that helps in face emotion recognition, the selected images form the collection of datasets, different class of emotions that are classified. There is large array of data is tested and the technique used is bottleneck, Long Short-term memory is taken to test the facial emotion recognition. As for speech recognition real-time results are being processed, similarly image recognition is done by the RGB technique of image classification for real-time results. Markov model has been used along with deep belief networks along with unweighted average recall of approximately 56.36%. Diverse images and datasets were inspected for recognizing the emotion by facial expressions using variety of classifier techniques like SVM, KNN, HMM. The following paper will explain significant features for example, Local invariant feature training, support vector machine training. Variety of different emotional datasets are being investigated, databases like Emo-DB, MES. The paper will describe the technique of detecting or identifying facial expression using Deep neural networks, whereas the traditional method of

identifying the expression is standard, that has been used to capture the change in emotion.

IV. PROPOSED METHODOLOGY

In this following section we will discuss the methodology, the expression datasets used in the further research, and Inception model.

A-Emotional Datasets

In this research we have explored mainly two datasets that are Karolinska Directed Emotional Faces (KDEF) and Kaggle’s Facial Expression Recognition Challenge. As the data we have is rarely used. Corpus Data mainly consist of about fifteen characters or types of actors in session with markers on the face, head and hands, that helps to examine crucial information of the facial emotions and hand movements. The character in the image has designed scenarios designed to deliver particular type of emotions. The database that we have downloaded has tones of audio-visual data. Based on different types of actors, there are clips of 10seconds those are further known as emotions classes. The whole database is divided into five sessions of data in .mp4 format. While taking the data form the actors, their facial expression is taken and analyzed with the seven ranges of the expressions.

B- Transfer learning

Transfer learning is the kind of machine learning algorithm in which the knowledge is obtained from the solution of one problem can be applied in finding solution of another problem. Transfer training is best technique as it increases the accuracy and reduces the cost too, it solves the problem in very short duration.

Transfer learning has been significantly applied and has been used in taking the weights of model and have also been implemented by repairing the layers and also helped in resting other layers too. In this paper transfer learning has to be implemented by the weights of the Kaggle database, it is larger dataset and also holding some particular layers of the KDEF database, it is smaller dataset. By this method as both datasets Kaggle and KDEF has similar dataset with same seven emotions.

C- Inception Net Model

Inception Net model helps in training and building expression recognition model that can be used universally in any application. GoogleLeNet architecture significantly uses inception.

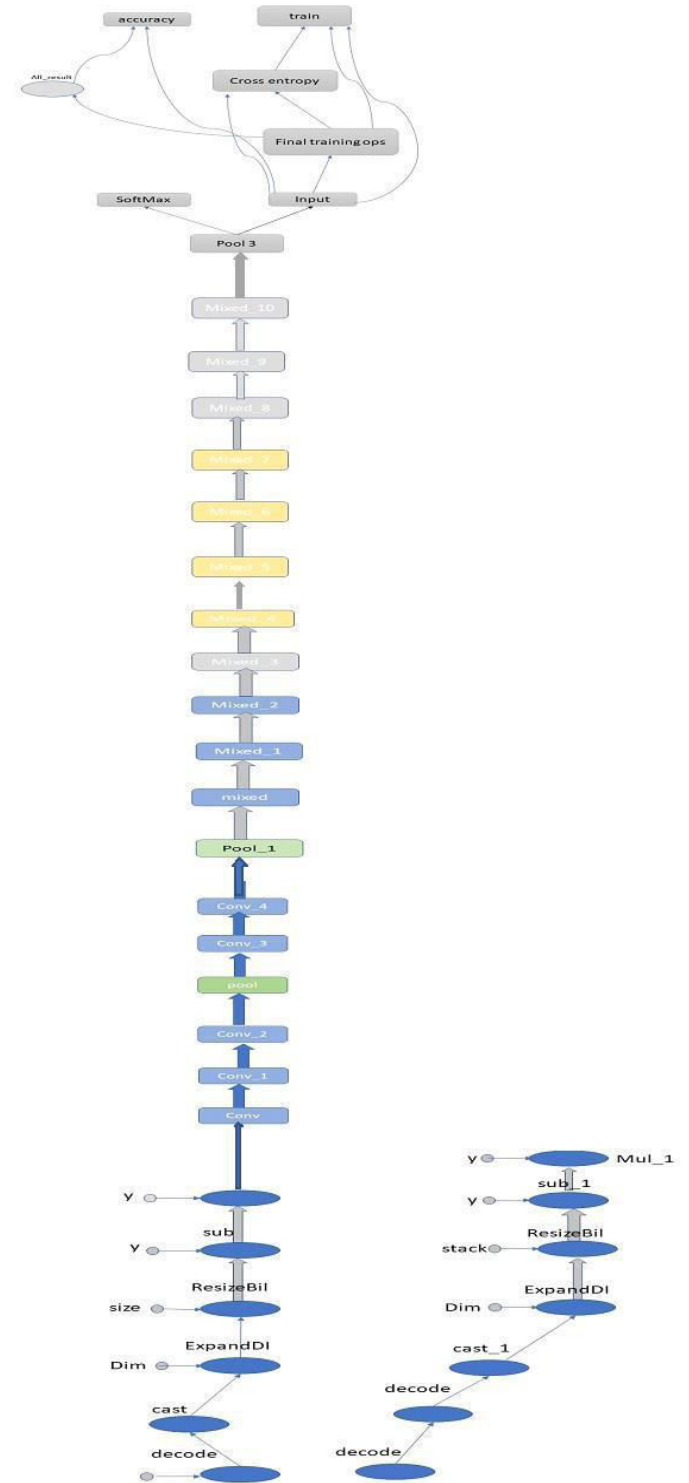


Fig. 1 Inception Model architecture

This model also used for Automatic Image categorization and image labeling. Inception Net v3 model is first runner up in image classification in google image search. With **fewer parameters like** AlexNet(60 million parameters), VGGNet(3X more parameters than AlexNet), GoogLeNet/ Inception-v1(7 million parameters) and with 42 layers deep learning network error rate in image classification obtained is very lower and this become Inception model the 1st runner up for image classification in ILSVRC(ImageNet Large Scale Visual Recognition Competition) 2015. Inception Net model is a combination of all the output of 1X1,3X3,5X5(Factorizing Convolutions) filters. This factorizing convolution comes into existence to decrease the number of parameters without reducing the deep learning network efficiency. According to inception Net model architecture there are some typos for the architecture. This architecture is consisting of 5X Inception, 4X Inception, 2X Inception modules connected with Grid size reduction (with some advancements) which helps to framework the facial emotion recognition databases. This model uses the image spectrograms to read and analyze the facial expressions and emotions that's why Inception Net model is used for image classification. After every layer of this refined architecture of Inception Net Model we get a accurate value of expressions and facial emotions to be affirmed by the trained model. The system to produce the desired output compare and analyze the new segment with the existing one.

D- Spectrogram

Spectrograms are basically 2-D graphs with 3-D color representation used to visualize the signal strength or loudness of a signal where time varies with respect to different frequencies in a particular waveform. Time varies from left to right (or oldest to youngest) on the x-axis. In this paper we have use spectrograms to imitates the image in a form though which we can recognize the expressions and emotions. As we know spectrograms are graphical depiction of frequency with respect to time where frequency lies on y-axis and time lies on horizontal axis. Spectrograms are also consisting different colors in third dimension

such as RGB, CMKYK etc. dark colors such as blue, black etc. imitate the week signal and brighter colors shows the strong signal. Sample of spectrogram in Fig. 1.

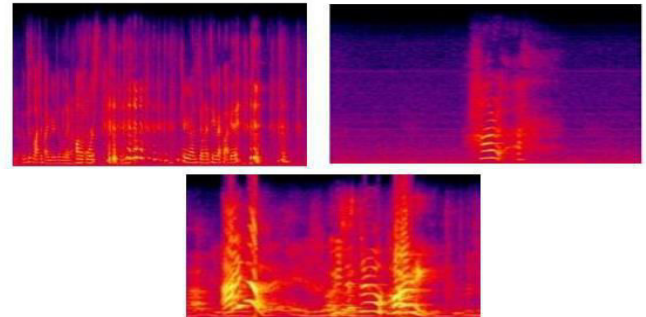


Fig 1. Different classes of emotions represented by spectrogram

These time-frequency graphs are computed by implementing Fast Fourier transform (FFT) to spectrograms. In order to get different frequencies first of all it is broke into different segments and then each segment of image spectrograms assigned fast Fourier transform.

E- Preparation of training dataset

All the image samples and video clips from Kaggle's and Karolinska Directed Emotional Faces (KDEF) datasets are pinched from different sessions. By using various emotion recognition output and result reports provided by image datasets, different .wav audio files are tagged and assigned into seven different range of emotions.

V. EXPERIMENTAL SETUP

All experimental setup, libraries and modules we used for deep learning explained in this section which helps a lot in facial emotion recognition.

A. Training method

All moving faces and images labelled with emotions are trained for the model. We have used TensorFlow to implement the designed model. We have used Kaggle's and Karolinska Directed Emotional Faces for generation of spectrograms of images with emotions labelled. These spectrograms of images are resized to 500 x 300. For the accurate results more

than 300 spectrograms were created from almost all images. For the different emotions, spectrograms were generated on various classes of emotions or above certain threshold values from our image database. The whole training method was enforced for 20 epochs with a batch size set to 45. Initial learning rate was set to 0.02 with a decay to 0.2 after each 20 epochs. We have applied this training model on a single Nvidia GeForce GT 730 with 2 GB RAM. This experimental training took around 30 minutes and best accuracy case was achieved after 35 epochs. On the training set, a loss of 0.80 was also achieved, this loss registered on the test set. The accuracy of 35% per spectrogram of the image which shows a very low overall accuracy. limitation in dataset for each class of emotion and transfer learning, these may be the cause of lower accuracy. So it is intended that use of transfer learning along with large labelled data set for each class to accord better accuracy.

VI. RESULT & ANALYSIS

For predicting the exact emotions from face expression, an accuracy rate of about 35.3% is achieved from the image data set for each class. The most challenging part of this paper was classification of different expressions from the image. But the resultant accuracy achieved is very low. The reasons stated for less accuracy rate is that, transfer learning, which is used to train the inception model and it could've use of a smaller number of data sets for each class for the training process, which are the major cause for low accuracy rate. It could've generated less number of spectrograms for images for training, which leads to the less accuracy.

VII. CONCLUSION

We have used a lot of researches and studies related to Emotion Recognition, Deep learning techniques used for facial expression recognition and emotion recognition are attended. In this project we have used Inception model to solve emotion recognition problem through generating many spectrograms for images. To carrying out the research various databases have been tested, Kaggle's and Karolinska Directed Emotional Faces (KDEF) is used to train the model. Tensor FLOW is also used to train the model.

An accuracy rate of about 36% is achieved. In future, the same architecture we can use for real time facial emotions recognition.

REFERENCES

- [1] "Survey on emotion recognitions: classification scheme and database" by M.El Ayaadi, F.Karraeand M.S. Kamal
- [2] "Application of face emotion recognition" by disruptionhub.com
- [3] "Inception model v3" in towardsdatascience.com
- [4] J. Ngiem, A. Khoslaa, M. Kam, J. Nim, H. Lei, and A. Y. Nig, "Multimodal Deep Learning," Proc. 28th Int. Conf. Mach. Learn.
- [5] F. Noruzi, G. Anbarjafaari and N. Akraami "Expression-based emotion recognition and next reaction prediction," 2017 25th Signal Process. Commun. Appl. Conf. SIU 2017, no. 1, 2017.
- [6] <https://machinelearningmastery.com/>
- [7] K.Weint and C.-W. Huaang, "Characterizing Types of Convolution in Deep Convolutional Recurrent Neural Networks for Robust Speech Emotion Recognition," pp. 1–19, 2017.
- [8] P. Hairár, R. Burgeet, and M. K. Duitta, "Facial Emotion Recognition with Deep Learning," in Signal Processing and Integrated Networks (SPIN), 2017, pp. 4–7.
- [9] R. Ashrafidust, S. Setaeyeshi, and A. Shaarifi, "Recognizing Emotional State Changes Using Facial Expression," 2016 Eur. Model. Symp., pp. 41–46, 2016.
- [10] C. Busiso et al., "IEMOCAP: Interactive emotional dynamic motion capture database," Lang. Resour. Eval., vol. 42, no. 4, pp. 335–359, 2008.
- [11] J. Riang, G. Leie, and Y. P. P. Chen, "Acoustic feature selection for automatic emotion recognition from speech," Inf. Process. Manag., vol. 45, no. 3, pp. 315–328, 2009.
- [12] J. Schmidhubar, "Deep Learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85–117, 2015.