

Emotion Recognition from Facial Expression using Deep Learning

Nithya Roopa. S

ABSTRACT--- Facial expression recognition is the part of Facial recognition which is gaining more importance and need for it increases tremendously. Though there are methods to identify expressions using machine learning and Artificial Intelligence techniques, this work attempts to use deep learning and image classification method to recognize expressions and classify the expressions according to the images. Various datasets are investigated and explored for training expression recognition model are explained in this paper. Inception Net is used for expression recognition with Kaggle (Facial Expression Recognition Challenge) and Karolinska Directed Emotional Faces datasets. Final accuracy of this expression recognition model using Inception Net v3 Model is 35%(~).

Index Terms— Facial recognition; expression recognition; deep learning; image recognition; Facial technology; signal processing; image classification

I. INTRODUCTION

Recognizing human expressions and emotions has drawn the attention of researchers, as the capability of recognizing one's expressions helps in human-computer interaction, to right advertising campaigns, and crowning with an augmented and enhanced human communication ,by amending the emotional intelligence ("EQ") of humans. There are many ways to inspect the recognition of human expressions, ranging from facial expressions, body posture, voice tone etc. In this paper we have focused on facial expression recognition. Facial Emotion Recognition(FER) is a thriving research area in which lots of advancements like automatic translation systems, machine to human interaction are happening in industries. In contrast the paper focus to survey and review various facial extraction features, emotional databases, classifier algorithms and so on. This paper is organized as follows. Section 2 describes background information about expression recognition, emotion recognition system and applications of emotion recognition. Section 3 explains the Feature selection methods and Image optimization. Section 4 compares various Facial emotional database. Section 5 addresses various classifier algorithms for classifying images according to the expression identified. The paper is concluded in Section 6.

II. BACKGROUND INFORMATION

A. Emotion Recognition

Facial Recognition is the technology that deals with methods and techniques to identify the emotions from the facial expression. Various technological developments in the area of Machine Learning and artificial Intelligence, made the emotion recognition easier. It is expected that

expressions can be next communication medium with computers. A Need for automatic emotion recognition from facial expression increases tremendously. Research work in this area mainly concentrates on identifying human emotions from videos or from acoustic information. Most of the research work recognizes and matches faces but they have not used convolutional neural networks to infuse emotions from images.

Emotion Recognition deals with the investigation of identifying emotions, techniques and methods used for identifying. Emotions can be identified from facial expressions, speech signals etc. Enormous methods have been adapted to infer the emotions such as machine learning, neural networks, artificial intelligence, emotional intelligence. Emotion Recognition is drawing its importance in research which is primary to solve many problems. The primary requirement of Emotion Recognition from facial expressions is a difficult task in emotional Intelligence where images are given as an input for the systems.

B. Facial Emotion Recognition

Facial Emotion Recognition is research area which tries to identify the emotion from the human facial expression. The surveys states that developments in emotion recognition makes the complex systems simpler. FER has many applications which is discussed later. Emotion Recognition is the challenging task because emotions may vary depending on the environment, appearance, culture, face reaction which leads to ambiguous data. Survey on Facial emotion recognition [2] helps a lot in exploring facial emotion recognition

C. Deep Learning

Deep Learning [3] is machine learning technique which models the data that are designed to do a particular task. Deep learning in neural networks has wide applications in the area of image recognition, classification, decision making, pattern recognition etc. [4]. Other deep Learning techniques like multimodal deep learning is used for feature selection, image recognition etc.

D. Applications of Facial Emotion Recognition

Emotion Recognition is used in BPO's for identifying calls based on their emotions. Emotion Recognition serves as the identifier for conversational analysis[7] for identifying the unsatisfied customer, customer satisfaction so on. FER is used in car board system depending on information of the mentality of the driver can be provided to the system to initiate his/her and the customer safety.

Revised Manuscript Received on August 14, 2019.

Nithya Roopa. S, Assistant Professor, Kumaraguru College of Technology, Coimbatore. T.N, India. (Nithyaroopa.s.cse@kct.ac.in)

III. LITERATURE SURVEY

A detailed study on the facial emotion recognition is discussed in [9] which exposes the properties of dataset, facial emotion recognition study classifier. Visual features of image is examined and some of the classifier techniques are discussed in [10] which is helpful in the further inspection of the methods of emotion recognition. This paper [11] examined the prediction of the future reactions from images based on the recognition of emotions, using different classes of classifiers. Some of the classification algorithms like K-Nearest Neighbour, Random Forest are applied in [11] to classify emotions. Neural network arises tremendously which attempts to solve problems in data science. Deep RNN like LSTM, Bi-directional LSTM modelled for audio visual features are used in [12]. Various range of CNN, modelled and trained for facial emotion recognition are evaluated in [13]. Facial emotion Recognition is drawing its own importance in the research field. Facial emotion recognition is inspected and analysed on all research areas [14]. Emotion is identified from facial images using filter banks and Deep CNN [15] which gives high accuracy rate with which we had an inference that deep learning can also be used for emotion detection. Facial emotion recognition can be also performed using image spectrograms with deep convolutional networks which is implemented in [16]. All the above methods mentioned used some of the conventional methods of feature selection from MFCC's, wave parameters such as pitch are used in the paper [6]. This paper [17] studies different database used for facial emotion recognition, features selected from facial expression images, classifiers used to classify different classes of emotions. As the amount of data array is taken and the method of bottleneck is used, Long Short-term Memory (LSTM) is used for Facial emotion recognition [18]. Though speech emotion recognition is done and desired results are shown, research real time facial emotion is still ongoing. Real-time facial emotion recognition is done through RGB image classification using transfer learning methodologies in which knowledge gained from solving one problem and that is implemented for the another problem [19]. Emotion has been recognized from facial expressions using hidden markov models and deep belief networks with unweighted average recall (UAR) of about 56.36% (~) [20]. Different image types and emotions were examined for detecting expressions from the facial expressions using different classifiers such as KNN, HMM, GMM, SVM [21]. This paper [22] explains about learning significant features such as Support vector machine training, local invariant feature learning, salient discriminative feature analysis for facial emotion recognition. Various significant features are examined and trained to detect emotions using Convolution neural networks in which dataset is obtained from various emotional databases such as SAVEE, Emo-DB, DES, MES. This paper describes Facial emotion recognition using Deep neural Networks. Although Probabilistic method of identifying emotion is conventional, that is used to recognize emotion changes [24].

IV. PROPOSED METHODOLOGY

This section explains the proposed methodology, emotion database used for research, Inception model.

A. Emotion Database

The two datasets we have used in our work are the Kaggle's Facial Expression Recognition Challenge and Karolinska Directed Emotional Faces (KDEF) datasets.. Since this data is rarely used, this work explores significantly on this dataset. Corpus Data consists of fifteen actors in sessions with markers on the face, head, and hands, which provide detailed information about their facial expression and hand movements. The actors shows selected emotions and also improvised scenarios designed to express specific types of emotions (happiness, anger, sadness, frustration and neutral state). Database consist of twelve hours of audio-visual data. We chose video clips of various sessions. Based on certain actors, these video clips of 10 seconds (approx.) are classified into one of the emotions classes. All the audio-visual data are divided into five sessions video data in .mp4 format. When capturing the data, actor's emotions are analysed by various actors into seven range of emotions. All the data are explored along with the database.

B. Transfer Learning

Transfer learning is one of the machine learning methods which uses the knowledge obtained from solving one problem to solve another problem. It is true that Transfer learning solves problems within short duration of time. Transfer Learning is implemented whenever the computation cost has to be reduced and to achieve accuracy with less training.

Transfer learning is a frequently applied technique which works by taking the learned weights of a model (e.g. ImageNet) and implement those by fixing other layers and holding the rest of the layers in the network. In this project, we implement transfer learning by taking the learned weights from the Kaggle dataset, a larger dataset, and holding a few layers on the KDEF dataset, the smaller dataset. This approach is chosen as both the KDEF and Kaggle has similar data, images having one of the seven emotions.

C. Inception Net v3 Model

Inception Net v3 [26] Model is used to train and build an emotion recognition model which can be used in any applications. Inception is evolved from GoogLeNet Architecture with some enhancements and advancements. Inception model is used for automatic image classification and image labelling. Inception-v3 is used for image classification in Google Image Search. Inception-v3 achieved top 5.6% error rate in ILSVRC 2012 classification challenge validation. Figure 1 illustrates the complete architecture of Inception Net v3 Model. Inception Net model consists of Inception module which combines all the output of 1x1, 3x3, 5x5 filters. Inception net consists of network in a network which consists of three inception modules that are

embedded inside the inception architecture which helps in modelling the emotional recognition database.

The Inception model which is used for image classification model building, is used for analysing and recognizing the expressions from the spectrogram of the image. The refined structure from each layer reveals the exact value of the emotion to be predicted by the trained model. The image classification and analysis of each segment by matching it with the existing segments is done parallelly by the system to produce the desired results.

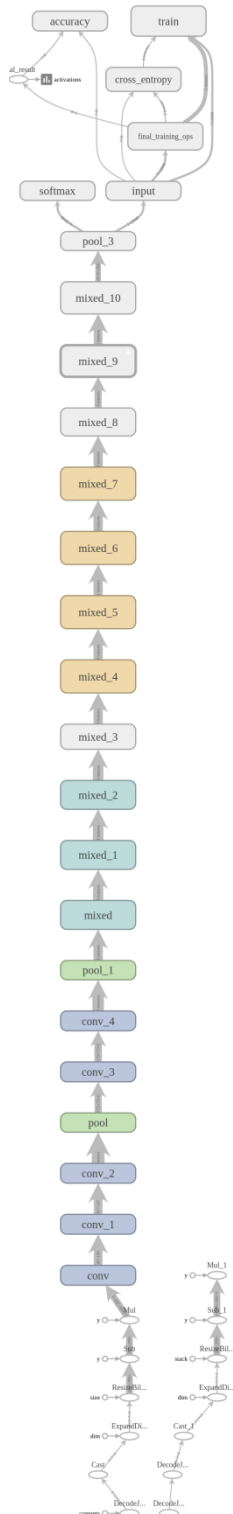


Fig. 1 Inception Architecture

D. Spectrogram

The Spectrogram represents the image from which the expression has to be recognized. Representation of spectrogram like 2D plots, 3D plots varies based on the type of the applications. Spectrogram is a plot in which Y axis represents frequency and x axis represents time. Spectrogram can consist of various color maps such as RGB, CMKYK etc. Weak signals are represented by dark blue colors and stronger signals are represented by brighter colors. Sample spectrograms are shown in Fig. 2.

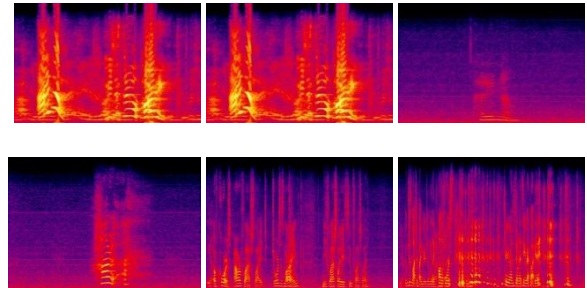


Fig 2. Spectrograms representing various classes of emotions

It is computed from the image by applying Fast Fourier transform (FFT) to spectrogram, which form time-frequency representation. In order to find the frequencies, it is divided into small segments and Fast Fourier Transform is applied to each segment of the image spectrogram.

E. Preparation of Training dataset

All the Video clips from the Kaggle’s and Karolinska Directed Emotional Faces (KDEF) databases are drawn from various sessions. Using the emotion identification report given in the database, various wav files are labeled and classified into seven range of emotions.

V. EXPERIMENTAL SETUP

This section which explains about experimental setup, libraries used for Deep learning which helps in emotion recognition.

A. Training method

All emotions labeled images are trained for the model. The designed model was implemented using TensorFlow. The spectrogram of the images with expressions were generated from the Kaggle’s and Karolinska Directed Emotional Faces (KDEF) are resized to 500 x 200. More than 500 spectrograms were generated from all the desired images. For each emotion, the part of the spectrogram above certain threshold values or class of emotion is collected from the image database. The training process was applied for 20 epochs with a batch size set to 50. Initial learning rate was set to 0.02 with a decay of 0.2 after each 20 epochs. Training data model was applied on a single NVidia GeForce GT 730 with 1 GB onboard memory. The training took around 25 minutes and the best accuracy was achieved

after 30 epochs. On the training set, a loss of 0.81 was achieved, whereas 0.85 loss was recorded on the test set. An accuracy of 35.91 % was achieved per spectrogram of the image. It is evident that the overall accuracy is very low. These may be due to transfer learning and less dataset for each class of emotion. It is proposed to use the transfer learning with large labelled data set to give better accuracy.

VI. RESULT & ANALYSIS

An accuracy rate of about 35.6% is achieved from the image data model for predicting the emotions from the expressions. The classification of the expressions from the image was challenging. It is evident from the below figure that 0.8 is the best accuracy rate achieved during validation of data and it can be more if it is classified and evaluated before conversion.

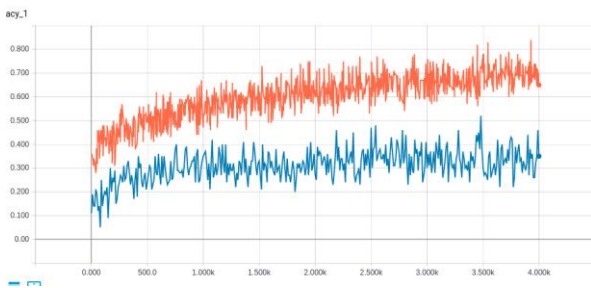


Fig.3 Accuracy rate of the Data Model

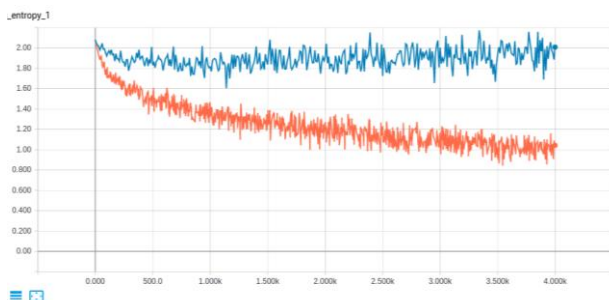


Fig.4 Cross Entropy

The reasons stated for less accuracy rate is that, Transfer Learning, which is used to train the model and it could've generated less number of spectrograms for training, which leads to the less accuracy. There are also less data set used for the training process which also leads to the case.

VII. CONCLUSION

Numerous researches and studies about Emotion Recognition, Deep learning techniques used for recognizing the emotions are conducted. It is required in future to have a model like this with much more reliable, which has limitless possibilities in all fields. This project tried to use inception net for solving emotion recognition problem. various databases have been explored, Kaggle's and Karolinska Directed Emotional Faces (KDEF) is used as dataset for carrying out the research. Tensor Flow is used to train the model. Accuracy rate of about 39% is achieved. In future, real time emotion recognition can be developed using the same architecture.

REFERENCES

1. M. El Ayaadi, F. Karraeand M. S. Kamal "Survey on emotion recognitions: Features, classification scheme, and database," *Pattern Recognit.*, vol. 44, no. 3, pp. 572–587, 2011.
2. Y. Lecan, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
3. J. Schmidhubar, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.
4. J. Ngiem, A. Khoslaa, M. Kam, J. Nim, H. Lei, and A. Y. Nig, "Multimodal Deep Learning," *Proc. 28th Int. Conf. Mach. Learn.*, pp. 689–696, 2011.
5. S. Lugivic M. Harva and I. Dander "Techniques and applications of emotion recognition," *2016 39th Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2016 - Proc.*, no. November 2017, pp. 1278–1283, 2016.
6. B. Schaller, G. Rigull, and M. Ling, "Emotion recognition combining acoustic features and information in a neural network - belief network architecture," *Acoust. Speech, Signal Process.*, vol. 1, pp. 577–580, 2004.
7. J. Rieng, 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.
8. 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.
9. G. Hinton ,Greves and A. Mohamed, "Emotion Recognition with Deep Recurrent Neural Networks," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 6645–6649.
10. 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.
11. A. M. Yusuf, M. B. Mustafaafa, M. and M. Malekzedeh, "emotion recognition research: an analysis of research focus," *Int. J. Speech Technol.*, vol. 0, no. 0, pp. 1–20, 2018.
12. L. Cavidon ,H. M. Fayak, M. Lich, "Evaluating deep learning architectures for Emotion Recognition," *Neural Networks*, vol. 92, pp. 60–68, 2017.
13. A. M. Badshaah, J. Ahmed and S. W. Baek, "Emotion Recognition from Spectrograms with Deep Convolutional Neural Network," *2017 Int. Conf. Platf. Technol. Serv.*, pp. 1–5, 2017.
14. A. Routrey, M. Swaen and P Kabisetpathy, "Database, features and classifiers for emotion recognition: a review," *Int. J. Speech Technol.*, vol. 0, no. 0, p. 0, 2018.
15. K.-Y. Hueng, C.-H. Wiu, T.-H. Yieng, M.-H. Sha and J.-H. Chiu, "Emotion recognition using autoencoder bottleneck features and LSTM," in *2016 International Conference on Orange Technologies (ICOT)*, 2016, pp. 1–4.
16. M. N. Sttilar, M. Leich, R. S. Bolie, and M. Skinter, "Real time emotion recognition using RGB image classification and transfer learning," *2017 11th Int. Conf. Signal Process. Commun. Syst.*, pp. 1–8, 2017.
17. D. Lie and E. M. Provost, "EMOTION RECOGNITION USING HIDDEN MARKOV MODELS WITH DEEP BELIEF NETWORKS."
18. H. Jiaang *et al.*, "Investigation of emotions for identifying depression using various classifiers," *Expression Commun.*, vol. 90, pp. 39–46, 2017.
19. Q. Maio, M. Diong, Z. Huaeng, and Y. Zhaen, "Learning Salient Features for Emotion Recognition Using Neural Networks," *IEEE Trans. Multimed.*, vol. 16, no. 8, 2014.
20. 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.

21. R. Ashrafidust, S. Setaeyeshi, and A. Shaarifi, "Recognizing Emotional State Changes Using Facial Expression," *2016 Eur. Model. Symp.*, pp. 41–46, 2016.
22. C. Busiso *et al.*, "IEMOCAP: Interactive emotional dynamic motion capture database," *Lang. Resour. Eval.*, vol. 42, no. 4, pp. 335–359, 2008.
23. C. Szeigedy, V. Vanhouicke, J. Shleins, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2014.