International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 05 | May 2022 www.irjet.net p-ISSN: 2395-0072

A Survey on Emotion Recognition Using CNN

Prof, S. S. Kale, Prachi Naidu, Priyanka Gaikwad, Akanksha Kaundinya, Sakshi Gajbhiye

Computer Science, NBN Sinhgad School of Engineering, Ambegaon, Maharashtra, India

---

Abstract - *Many automated system applications, such as robotics, artificial intelligence, and security rely heavily on face expression recognition. Accurately recognising facial expressions can be quite difficult for machines. It has a wide range of applications in the areas of training, online business, health, and security. This study examines a variety of CNN-based face expression recognition systems. It comprises methods proposed by various researchers. The study also demonstrates how to use CNN for FER. This paper also examines CNN-based techniques and problems to consider when deciding whether or not to use CNN to solve FER.)*

Key Words: *CNN, FER, Emotion Recognition, Survey, Deep learning*

1. INTRODUCTION

One of the important ways humans interact is through facial expressions. In humans this ability is highly developed. Today, computers are used to automate everything. Developing the ability for computers to recognize emotions is a very popular subject. Emotion recognition can be very useful in marketing research, academics, robotics and security.

This paper survey focuses on models that are aimed at real-time facial emotion recognition as static emotion recognition is not very useful. A study on various feelings of humans put forward by Darwin stated that a human's face is a significant factor on how humans communicate **[1]**. The six basic emotions which Ekman et al proposed are Fear, Disgust, Anger, Happiness, Sorrow and Surprised **[2]**. CNN, which is part of the Deep Learning system, may be used to learn facial expressions. The features like ears, eyes, mouth and hair play an important part in achieving the outcome. The extraction of face traits reduces the amount of time and resources required for the procedure without sacrificing important information.

For this, task deep learning methods are used which has hidden layers and classification layers. Hidden layers are also called feature extraction layer. Hidden layer consists of convolution layer followed by pooling layer used for feature extraction. This is followed by classification part.

2. TRADITIONAL APPROACH

Most papers have proposed their research using MLP (Multi-layer Perceptron Model), SVM (Support vector machine) and KNN (K-Nearest Neighbors). The difference between these traditional methods and CNN is that in old methods features need to be extracted manually whereas CNN learns these features to detect an emotion.

3. DATASET

There are various datasets available online for different purposes. The most commonly used dataset was FER-2013 with almost 30,000 images. FER-2013 has 7 categories. (Sad, angry, happy, surprise, disgust, neutral, fear)

4. PRE- PROCESSING

Images come in a range of sizes and colors. They also come from different sources. All adjustments on the raw data before it is input to the machine learning or deep learning algorithm are referred to as preprocessing. Sometimes, the data quantity we have is not sufficient to perform classification. We use data augmentation to solve that problem. Data augmentation can be done by flipping, rotating, adding noise, cropping, etc.

5. CNN ARCHITECHTURE

After pre-processing the data, it is fed into CNN model. CNN is a deep learning algorithm; it is different from machine learning languages as it has hidden layers for image processing. It involves convolution operations. CNN has two main layers 1, Hidden layers 2. Fully connected layers. Hidden layers are used for feature extraction. Hidden layers also have two layers, convolution and pooling layers. The first layer is convolution layer. It is used to extract features from the input images. A filter of size NXN glides over the input image and a dot product is taken of that area. The output that is stored is called feature map feature. Pooling layer is followed by fully connected layers. The output from all the previous layers is flattened and fed into fully connected layer. This layer usually comes before output. In this layer classification process begins. To avoid overfitting, i.e., when a model works so well on training data that it has negative impact on testing data, we use dropout layers. Dropout layers drop nodes from the neural network randomly. Activation is the last component used to increase non-linearity of the output. Mainly, ReLu function is used in CNN. This function returns 0 if your value is negative or returns the same value you gave if it is between 0 to infinity.by pooling layer. The aim of this layer is to reduce size of the feature. Convolution layer is followed map which reduces computational cost and time. Max pooling takes the max value from the filter as the output whereas average pooling calculates the average. It reduces the size of the feature map without losing any important

Name	Researcher	Dataset	Samples	Model	Result
Facial Emotion Recognition Using Deep Convolutional Neural Network	Pranav E. et al.	Manually collected using 48 MP camera	2550	CNN	78.04%
Facial Expression Recognition with Convolutional Neural Networks	Shekhar Singh et al.	FER-2013	35,887	CNN	75.25%
Emotion recognition from facial expression using deep convolutional neural network	D. Y. Liliana	CK+	10,708	CNN	92.81%
Learners Mood Detection Using Convolutional Neural Network (CNN)	Rosa Ariani et al.	Manually collected	220	CNN	76.66%
Facial expression recognition using convolutional neural network with weighted loss function	Jiawei Luan et al.	FER-2013	35,887	R-CNN	69.69%
Research on Face Expression Detection Based on Improved Faster R-CNN	Weiran Hua et al.	SFEW 2.0	1765	CNN	81.33%
Facial Expression Classification Using Convolution Neural Networking and Its Applications	Ruhi Jaiswal	FER-2013	35,887	CNN	66%
Facial Emotion Recognition using Convolutional Neural Networks	Akash Sarvanan et al.	FER-2013	35,887	CNN	60%
Facial Expression Recognition with CNN Ensemble	K Liu ae al.	FER-2013	35,887	CNN	65.03%
Facial Expression Recognition using Visual Saliency and Deep Learning	M. Viraj et al.	CFEE	1610	CNN	74.79%
		RaFD	1407		95.71%
Going deeper in facial expression recognition using deep neural networks	A. Mollahosseini et al.	MMI	1280 videos	CNN	77.90%
		DISFA	4845 video frames		55.00%
		FERA	289		76.7%
		SFEW	663		47.7%

6. Analysis

The above table shows analysis of the previous papers. After analysis, we found that we need only one or two CNN layers for full frontal face images as they are not very complicated. Then, it is followed by pooling layers. Mostly, max pooling layer is used. All these layers are followed by fully connected layers which contains neurons.

More complicated images require 3 to 6 layers followed by pooling layers. But these pooling layers are not placed after each convolution layer. Pooling layers are mostly followed after two convolution layers. This is done so that the model can learn the features before its size is reduced. More layers lead to more parameters. This can be fixed by introducing dropout layers. This layer randomly drops neurons. This also reduces the risk of overfitting.

7. CONCLUSION

In this paper we have surveyed the effectiveness of CNN on facial expression recognition. We have also highlighted the difference between traditional architecture and CNN. Traditional architecture is more complicated, time consuming and less accurate as compared to CNN.

We have also compared previous research paper to understand the most useful CNN architecture and dataset.

From our survey we have concluded that a smaller number of layers and parameters are required for full frontal images. Whereas, a greater number of layers is required for more complicated images.

REFERENCES

[1] Darwin C. The expression of emotions in animals and man. London: Murray 1872

[2] Ekman P, Friesen WV. Constants across cultures in the face and emotion. Journal of personality and social psychology 1971

[3] Pranav, E.; Kamal, Suraj; Satheesh Chandran, C.; Supriya, "Facial Emotion Recognition Using Deep Convolutional Neural Network", IEEE 2020 6th International Conference on Advanced Computing & Communication Systems (ICACCS), doi:10.1109/ICACCS48705.2020.9074302

[4] Singh, Shekhar; Nasoz, Fatma, "Facial Expression Recognition with Convolutional Neural Networks" IEEE 2020 10th Annual Computing and Communication Workshop and Conference (CCWC)

[5] D. Y. Liliana, "Emotion recognition from facial expression using deep convolutional neural network", 2018 International Conference of Computer and Informatics Engineering.

[6] Rosa Ariani, Munir, Siswo Handoko, "Learners Mood Detection Using Convolutional Neural Network (CNN)", 2017 3rd International Conference on Science in Information Technology (ICSITech)

[7] Jiawei Luan, "Facial expression recognition using convolutional neural network with weighted loss function", 2020 IEEE 3rd International Conference on Information Communication and Signal Processing (ICICSP)

[8] Weiran Hua, Qiang Tong, "Research on Face Expression Detection Based on Improved Faster R-CNN", 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)

[9] Ruhi Jaiswal, "Facial Expression Classification Using Convolution Neural Networking and Its Applications", 2019 4th International Conference on Computer Science and Engineering (UBMK)

[10] Akash Sarvanan et al., "Facial Emotion Recognition using Convolutional Neural Networks"

[11] Kuang Liu, Mingmin Zhang, Zhigeng Pan," Facial Expression Recognition with CNN Ensemble", 2016 International Conference on Cyberworlds (CW)