Facial Expression Detection using Deep Neural Networks

M. Sandeep Reddy, Ch. Chinmai, B. Sai Teja, P. M. Ashok Kumar

Abstract: ***Facial Expression conveys nonverbal communication, which plays an important role in acquaintance among people. The facial expression detection system is an activity to identify the emotional state of the person. In this system, a captured frame is compared with trained data set that is available in the database and then state of the captured frame is defined. This system is based on Image Processing and Machine Learning. For designing a robust facial feature descriptor, we apply the Xception Modelling algorithm. The detection performance of the proposed method will be evaluated by loading the dataset and pre-processing the images for feeding it to CNN model. Experimental results with prototypic expressions show the superiority of the Xception-Model descriptor against some well-known appearance-based feature representation methods. Experimental results demonstrate the competitive classification accuracy for our proposed method.***

Keywords: CNN, Facial Expression, ImageNet, Xception Modelling.

I. INTRODUCTION

Facial expression is the representation of the affective state, intention, personality and psychopathology of a person and plays a communicative role in interpersonal relations. Though the implementation has been made, recognizing facial expression with a more accuracy and reducing the noise remains to be burdensome due to the complexity and having many types of various expressions [1].

Generally, non-verbal ways of communication are used by humans when they want to express their feelings and are things that humans accomplish in day to day life. These include gestures, facial expressions, and unconditioned languages. This system can be significantly used by people for effective communication.

On frequent basis, humans repeatedly identify emotions by characteristic features, viewed as a piece of facial expression. For instance, sad is irrefutably connected with a tragic or a downward development of the edges of the lips. Investigation into programmed discovery of the outward appearances tends to the issues encompassing the portrayal and classification of static or dynamic attributes of these misshapenness of face pigmentation [2].

The principle motivation behind this framework is proficient cooperation between people and machines utilizing eye stare, outward appearances, psychological displaying and so forth.

The framework's power fluctuates from individual to individual and furthermore changes alongside age, sexual orientation, estimate and state of face, and further, even the appearances of similar individuals don't stay consistent with time.

Be that as it may, the natural inconstancy of facial pictures brought about by various variables like varieties in enlightenment, present, arrangement, impediments make demeanor discovery a difficult errand. A few reviews on facial highlight portrayals for face recognition and demeanor investigations tended to these difficulties and potential arrangements in detail [3].

In today's world, the need to maintain security has become a challenging issue. In countries like Venezuela, the rate of crimes is increasing day by day. No automatic systems are there that can track person's activity. If it's possible to track facial expressions of person automatically, then finding criminals is done easily.

II. METHODOLOGY

A. Xception Modelling

Xception is a convolutional neural network that is trained on more than million images from ImageNet database. This modelling consists of 71 layers in the format of deep structure and classification of images into more than 1000 different categories.

B. Architecture

Xception is deep convolutional neural network architecture that involves depth-wise separable convolution. Google researcher's developed Xception Modelling. Initially, Inception modules are proposed by google and later, a depth wise separable convolution is brought into existence and is perceived as an Inception module with great number of towers. Precedingly, depth wise separable convolutions replaced inception modules.

The reason behind using Xception modelling is:

- As compared to CNN, they've a smaller number of parameters.
- In mobile applications, due to fewer computations, they're cheaper in in vision.

Revised Manuscript Received on February 06, 2020.

Dr. P. M. Ashok Kumar, Assoc. Prof., Department of Computer Science, K L Educational Foundation, Guntur, India.
Email: pmashokk@gmail.com

M. Sandeep Reddy, Department of Computer Science, K L Educational Foundation, Guntur, India. Email: sandeep.reddym@icloud.com

Ch. Chinmai, Department of Computer Science, K L Educational Foundation, Guntur, India. Email: chaluvadi.chinmai99@gmail.com

B. Sai Teja, Department of Computer Science, K L Educational Foundation, Guntur, India. Email: saitejabalusu3@gmail.com

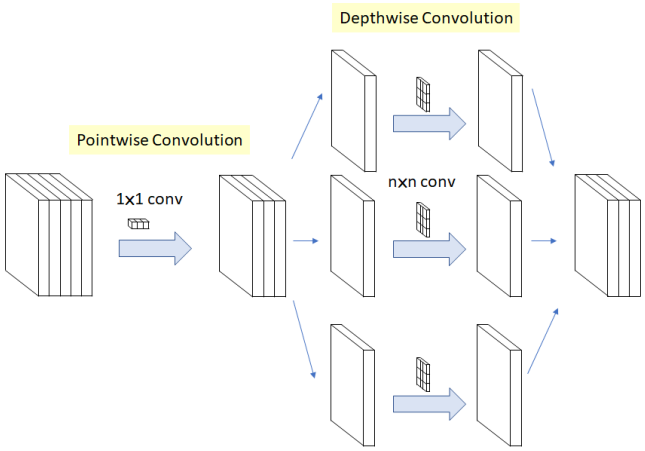
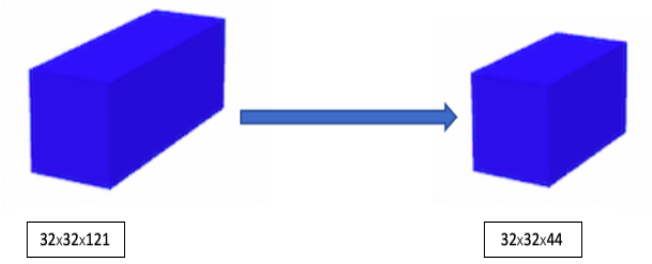


Fig. 1. Architecture of Xception Modelling

C. Point-wise Convolution

A 1x1 convolution is called as pointwise convolution. Its work is uncomplicated task which connects the input pixel to output similar pixels. It reduces depth channel size and the drawback is that it is simpler and slower for the large data inputs to calculate multiplication process.

Here, in Xception modelling, to attain formulation, the process for dividing each frame into single task is explained above the frames are sent to dept- wise convolution step.



Here, in Fig. 2, through point-wise convolution, the depth of initial frame is filtered and reduced to point wise kernels.

D. Depth-wise Convolution

Depth-wise convolutions are a variation in regular 2D convolution equated over multiple number if input channels. The filter undergoes in deep format as it can combine the channels and generate each element in the output. Depth-wise convolution do not have combined channels. Here, each channel is kept separate and that is the reason for its name Depth-Wise.

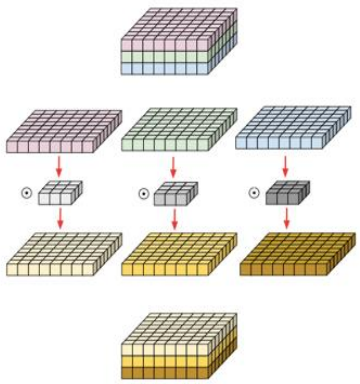


Fig. 3. Architecture of Depth-wise Convolution

There are three conceptual stages:

- Division of input and filter into separate channels.
- Combining input with preceding filter to obtain the final layer.
- Combination of the outputs of previous channels is the final channel.

III. RESULT AND DISCUSSION

Confusion Matrix: A confusion matrix is the measurement of the effectiveness and performance of an algorithm. It is a table with a combination of actual observations and predicted observations and it is useful for measuring Accuracy, Precision, Recall, and F1score. We used four performance measures like True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

- True Positive implies if anticipation is positive and it is true.
- True Negative implies if anticipation is negative and it is true.
- False Positive implies if anticipation is positive and it is false.
- False Negative implies if anticipation is negative and it is false.

Accuracy, in general, tells about how regularly our algorithm or model is correct and it is the proportion of correctly estimated observations to the total observations.

$$Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (1)$$

In general, there are 7 basic expressions (Happy, Sad, Disgust, Angry, Surprised, Scared + Neutral).

Total Images Taken: 70

Table-I: Confusion Matrix

	Neutral	Happy	Sad	Disgust	Angry	Surprised	Scared
Neutral	9	1	0	0	0	0	0
Happy	1	9	0	0	0	0	0
Sad	1	0	8	1	0	0	0
Disgust	1	0	1	6	1	1	0
Angry	0	0	0	1	8	1	0
Surprised	1	0	0	0	1	7	1
Scared	0	0	0	1	1	1	7

For each category of expression, we've considered 10 different images and classified them based on majority of percentage obtained.

As per Equation (1),

	(TP+TN)/(TP+TN+FP+FN)	Accuracy
Neutral	9/10	0.9
Happy	9/10	0.9
Sad	8/10	0.8
Disgust	6/10	0.6
Angry	8/10	0.8
Surprised	7/10	0.7
Scared	7/10	0.7

Total Accuracy = 0.7142857

IV. CONCLUSION

This research introduces an approach of categorization of facial expressions. Face recognition and extraction of expressions play a crucial role in many real-time applications like digital cameras, mobile cameras, robotics, security and human-computer interaction. This research objective is to develop a facial expression detection system that enhances feature extraction and classification of expressions. In our system, 7 basic expressions of different person images have been analyzed from different datasets. Our research deals with preprocessing of facial expression of captured images followed by modelling through Xception methodology. This project validates more facial expressions that is based on FER database. Our system can be used to detect and track a person's state of mind for example, in shopping center, to view the feedback of clients to enhance the business.

REFERENCES

1. Shan, C., Gong, S., & McOwan, P. W. (2018, September). Robust facial expression recognition using local binary patterns. In Image Processing, 2018. ICIP 2018. IEEE International Conference on (Vol. 4, pp. II-370). IEEE.
2. Chibelushi, C. C., & Bourel, F. (2017). Facial expression recognition: A brief tutorial overview. CVonline: On-Line Compendium of Computer Vision, 11.
3. Ahmed, F., Bari, H., & Hossain, E. (2019). Person-independent facial expression recognition based on compound local binary pattern (CLBP). Int. Arab J. Inf. Technol., 11(2), 195-203.

AUTHORS PROFILE

Dr. P M Ashok Kumar obtained his Ph.D. from Anna University, B.Tech in ECE from JNTU Hyderabad and M.E Computer Science Engg. from Anna University, Chennai, India. Currently he is working as Associate Professor in the Department of Computer Science Engineering, KL University, India. He authored more than 20 articles in both SCIE & Scopus indexed journals, conferences. His main research interests include Image processing, Data mining, Machine learning. He is currently working on Face Detection & Recognition, Video Traffic surveillance, Texture feature extraction, Object tracking features.

Mr. M Sandeep Reddy is pursuing his B.Tech in K L University. He is passionate about research and his rea of interests are Machine Learning, Artificial Intelligence, Network and Security, Cloud Computing and Application Development. In addition to it, he is good at coding and is also interested to work in health care organizations.

Mr. Ch Chinmai is pursuing his B.Tech in K L University. He is passionate about research and his rea of interests are Machine Learning, Artificial Intelligence, Network and Security and Cloud Computing and Application Development. In addition to it, he is interested to work in health care organizations.

Mr. B Sai Teja is pursuing his B.Tech in K L University. He is passionate about research and his rea of interests are Machine Learning, Artificial Intelligence, Network and Security and Cloud Computing and Application Development. In addition to it, he is interested to work in health care organizations.