TEST Engineering & Management

January-February 2020
ISSN: 0193-4120 Page No. 4057- 4061

A Deep Learning based Facial Emotion Recognition System

Bevish Jinila .Y, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India
Shyamal Shankar .S, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India
Ajitha .P, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India.
Sivashangari .A, School of Computing, Sathyabama Institute of Science and Technology, Chennai, India

Article Info
Volume 82
Page Number: 4057- 4061
Publication Issue:
January-February 2020

Article History
Article Received: *18 May 2019*
Revised: *14 July 2019*
Accepted: *22 December 2019*
Publication: *21 January 2020*

Abstract:
In the recent times, it is well-known that the elderly people are left alone in their home during working hours, as most of the present generation work to earn. It is difficult to analyze their behavior and emotion at all times. This paper presents a deep learning based facial emotion recognition system to track and analyze the expressions of the elderly from the video data. This system generates a report and an alert of the elderly being monitored.

Keywords: *deep learning, elderly, facial emotion recognition, image processing*

I. INTRODUCTION

In this modern era, technological advancements has acquired more of the lifetime of the younger generations. People spend most of their lifetime earning for their living and has less time to spend with the family. Meanwhile, spending time with elderly and caring them has been a bigger outset for the past decade. Several innovations and advancements have been done in the field of facial expression detection. There are a number of approaches introduced for detection and recognition of facial emotions. Though many research works has been introduced to track the behavior and movements of elderly, there are little work done to apply emotion recognition in the elderly.

Facial emotions are generated as an outcome of distortions of facial features and this reshapes the facial structure with respect to emotions. It is more common that, humans are subject to six different kinds of emotions like happy, sad, anger, fear, surprise and disgust. The facial expression varies with respect to humans on varied situations. The process of emotion recognition is categorized into three different phases namely; face recognition, feature extraction, and emotion classification. In this paper, a deep learning based approach is used for facial emotion recognition.

II. RELATED WORKS

A. Principal component analysis (PCA)

Donoto et al. [1] has proposed a scheme on classifying facial actions. This scheme mainly focuses on recognizing the emotions and cognitive behaviour of an individual. A comparison of various approaches on classification of facial actions is analyzed. This paper shows the importance of the usage of local filters and spatial frequencies for classifying facial actions. But, the limitation with this approach is the lack of coordination between motion pictures and textual information.

Yang et al. [2] has introduced a 2D principal component analysis for image representation. This is a classical approach .This 2D approach was compared with the classic Principal Component Analysis and it shows better results in accuracy. Though it shows good accuracy across databases, it seems that in some experiments, there was degradation in performance. It can also be noted that 2D-PCA image representation is also not as efficient as PCA with respect to requirements in storage. Further, the mean square error is high in 2D-PCA compared to PCA.

B. Machine learning approach

A deep CNN based transfer learning [6] is suggested for emotion recognition in the wild. This paper mainly focuses on illuminations, occlusions, backgrounds and noise. The main focus is based on least squares regression classifier and weighted score level. The inference of this approach seems to show that there is imbalance in the classification process.

Li et al. [7] proposed a deep CNN based approach with emotion joint learning. In this paper, both facial and identity features are extracted using two separate CNN networks. But, there are issues related to alignment mismatch.

Chang et al. [8] has proposed a deep leaning based framework for facial expression recognition for restaurants. Since the world is automated, the upcoming restaurants are unmanned and atomized. The deep learning approach proposed in this paper identifies the expression of the customers by trained CNN models. A scoring system was introduced to assess the satisfaction of the customers based on the expressions retrieved. A rough estimation is made based only on two inputs and also it does not integrate with the existing platforms.

A deep learning approach on emotional big data in audio and video [9] uses CNN for emotion recognition. In this scheme, two CNN's are joined together to form a higher learning machine. Finally, support vector machine is used to classify the emotions. In [10,11] rough set based schemes are also used in medical images to classify the outcomes.

III. METHODOLOGY

This section briefly explains about the proposed approach on facial emotion recognition. The various steps involved in this process of emotion recognition include pre processing, facial detection, feature extraction, feature matching and outcome generation.

A. System Architecture

The Convolution Neural Network (CNN) architecture consists of the mainly convolution and pooling. The main component in CNN is convolution. It is the major building block. It is a mathematical representation that helps to merge two sets of data. The convolution is applied to a set of data using the convolution filter and this produces a feature map. The convolution filter is also known as the kernel. In this proposed method, a 3x3 convolution filter is used. As a next step, activation function is applied. Stride is applied to the convolution filter. This shows how much the convolution filter is moved. If less overlap is required between receptive fields, then a bigger stride is required. Padding is applied to maintain the same functionality. Then, pooling is applied to reduce the dimensionality. This paves way for the reduction of parameters and reduces over fitting. The layers in pooling samples each feature independently by reducing the height and width. In this proposed method, the most commonly used method, max pooling is used. The pooling does not have any parameters. The pooling and convolution layers together perform feature extraction. Then, feature mapping and classification is performed.

A. Process Flow

Figure 1 shows the process flow of the proposed deep learning based method on emotion recognition.

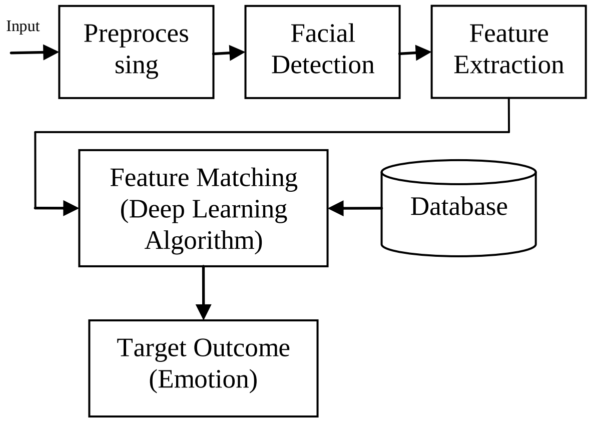

Fig. 1. Process flow of the proposed system

Pre-processing

To overcome the inaccuracies in the detection of emotion that discriminates the facial features, the inter-class feature mismatch should be reduced. To make this possible, image normalization should be applied. The Gaussian normalization is done using Gaussian mean and standard deviation. The input image is represented as p(x,y) and q(x,y) is the normalized output image, where x and y represents the row and column of the processed image.

Facial Detection

The first step in facial emotion recognition is to detect the face. This requires analysis of the entire image. Then, the next step is using the detected face to extract every feature for emotion recognition.

Feature Extraction

Feature extraction is done to determine the type of emotion namely happy, sad, anger, fear, surprise and disgust. In this proposed approach Action Units (AU) are used to recognize the facial expressions using facial features like eyes, mouth and lip.

Feature Mapping

The feature mapping is done using CNN. It performs both convolution and pooling. Using this, it performs feature mapping and classification.

I. EXPERIMENTAL RESULTS AND ANALYSIS

The dataset used for emotion recognition is the one from Cohn [12] expression database. This data set contains six emotions. All the tests of the algorithm are performed for 50 different samples. The analysis is performed bas ed on the six outcomes. It is considered that a person many or may not go to neutral state, while changing emotions.

Fig.2. Sample images on facial expression

The analysis is performed based on three major metrics namely accuracy, precision and recall. Accuracy shows the amount of trueness in terms of percentage. Precision is the accurateness in terms of percentage. Recall refers to the percentage of total results classified correctly by the program.

The proposed method uses CNN for emotion recognition. This algorithm is compared with SVM (Support Vector Machine). The outcome is measured based on the three metrics namely accuracy, precision and recall.

Precision is computed by the following equation

$$\text{Precision} = T_p / \quad T_p + F_p \qquad (1)$$

In equation (1), T_p represents true positive, F_p represents false positive.

Recall refers to the total relevant results that are correctly classified by the algorithm. Recall is computed as shown in equation (2).

$$\text{Recall} = T_p / \quad T_p + F_n \qquad (2)$$

In equation (2), T_p represents true positive, T_n represents true negative, F_p represents false negative. Figure 3 shows the data set representation as a scalar graph.

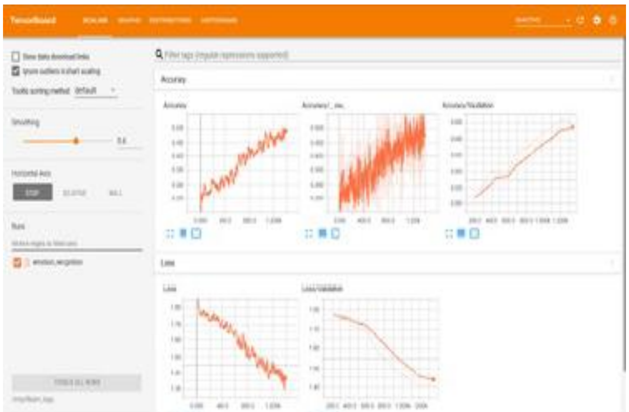

Fig. 3. Data set representation

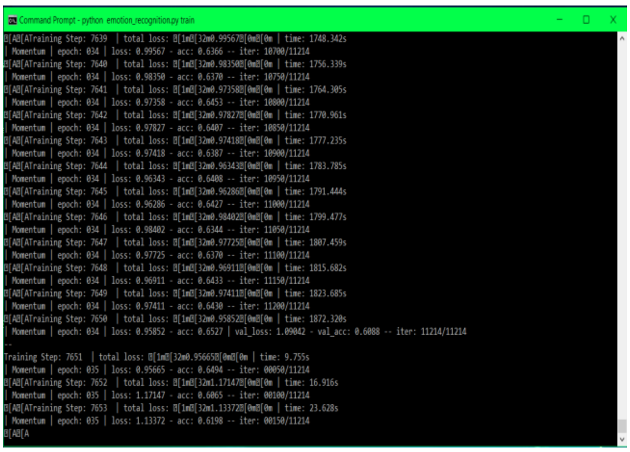

Fig.4 Data training

January-February 2020
ISSN: 0193-4120 Page No. 2732 - 2742

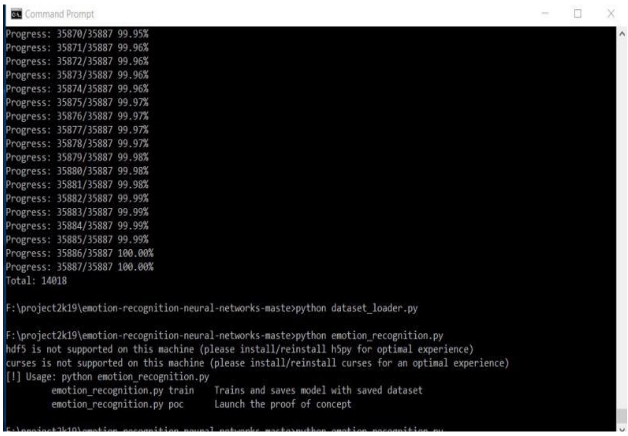

Fig.5 Metric outcomes

Figure 4 and 5 shows the training and testing of the data. Table 1 shows the comparative analysis of the methods SVM asnd CNN.

Table 1. Comparison of Results

	Accuracy	**Precision**	**Recall**
SVM	67.50%	63.20%	68.90%
CNN	95.40%	96.20%	93.30%

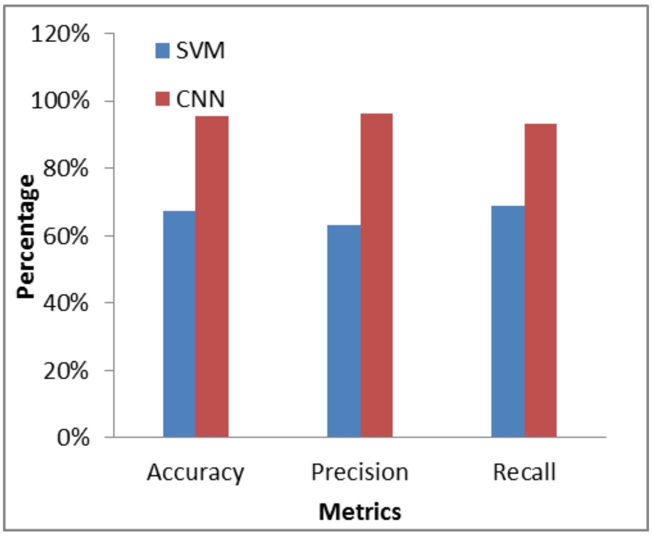

Fig.6. Comparison of metrics

Table 1 and figure 6 show the comparison of the metrics namely accuracy, precision and recall for SVM and CNN methods. Results show that CNN has better accuracy, precision and recall compared to SVM.

I. CONCLUSION

This paper presents a CNN based deep learning approach for facial expression recognition. Experimental results shows that CNN based approach provide better accuracy, precision and recall compared to the SVM approach. The outcomes generated are based on six emotions namely happy, sad, anger, fear, surprise and disgust. In future, the analysis can be extended to support further activities of the elderly people.

REFERENCES

1. Donato, Gianluca, Marian Stewart Bartlett, Joseph C. Hager, Paul Ekman, and Terrence J. Sejnowski. "Classifying facial actions." IEEE Transactions on pattern analysis and machine intelligence 21, no. 10 (1999): 974-989.
2. Yang, Jian, David D. Zhang, Alejandro F. Frangi, and Jing-yu Yang. "Two-dimensional PCA: a new approach to appearance-based face representation and recognition." IEEE transactions on pattern analysis and machine intelligence (2004).

TEST Engineering & Management
January-February 2020
ISSN: 0193-4120 Page No. 4057- 4061

3. Gosavi, Ajit P., and S. R. Khot. "Emotion recognition using principal component analysis with singular value decomposition." In 2014 International Conference on Electronics and Communication Systems (ICECS), pp. 1-5. IEEE, 2014.
4. Zhao.G, M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions", *PAMI*, vol. 29, no. 6, pp. 915-928, 2007.
5. Gosavi, Ajit P., and S. R. Khot. "Facial expression recognition using principal component analysis." International Journal of Soft Computing and Engineering (IJSCE) 3, no. 4 (2013): 2231-2307.
6. Kaya, Heysem, Furkan Gürpınar, and Albert Ali Salah. "Video-based emotion recognition in the wild using deep transfer learning and score fusion." Image and Vision Computing 65 (2017): 66-75.
7. Li, M., Xu, H., Huang, X., Song, Z., Liu, X., & Li, X."Facial Expression Recognition with Identity and Emotion Joint Learning" IEEE Transactions on Affective Computing, 2018.
8. Chang, W. J., Schmelzer, M., Kopp, F., Hsu, C. H., Su, J. P., Chen, L. B., & Chen, M. C., " A Deep Learning Facial Expression Recognition based Scoring System for Restaurants." In 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC) (pp. 251-254), 2019, IEEE.
9. Hossain, M. Shamim, and Ghulam Muhammad. "Emotion recognition using deep learning approach from audio–visual emotional big data." Information Fusion 49 (2019),pp. 69-78.
10. Madhukeerthana, Y. Bevish Jinila, "A review on rough set theory in medical images", Research Journal of Pharmaceutical, Biological and Chemical Sciences, Vol.7, Issue 1, pp.815-822, ISSN : 0975-8585,2016.
11. Madhukeerthana, Y. Bevish Jinila, Deepika, "Enhanced rough set theory for denoising brain MR images using bilateral filter design ",Research Journal of Pharmaceutical, Biological and Chemical Sciences, Vol.7, Issue 3, ISSN : 0975-8585,2016.
12. T. Kanade, J. Cohn, Y. Tian, Comprehensive database for facial expression analysis, 2000