Recognition of Facial Expressions using Deep Learning Model

Preprint · July 2020

DOI: 10.13140/RG.2.2.27834.41926

1 author:

Arunkumar Senthamaraikannan
Adhiparasakthi Engineering College
4 PUBLICATIONS **6** CITATIONS

SEE PROFILE

International Journal of Advanced Science and Technology
Vol. 29, No. 04, (2020), pp. 6813 – 6820

Recognition of Facial Expressions using Deep Learning Model

Arunkumar S[1], Geetha.A[2]

[1]*Research Scholar, Dept. of Computer and Information Science, Annamalai University*
[2]*Professor, Dept. of Computer Science and Engineering, Annamalai University*
arunkumar@adhiparasakthi.in, aucsegeetha@yahoo.com

Abstract

People are using social media to express their Expressions. Automatic Facial expression recognition is an emerging topic for researchers. Emotions of a human being are expressed through Facial Expression. This work talks about the research works done in the field of facial expression identification and recognition method using Convolutional Neural Network (CNN). The proposed method produces a better result than the existing method found in the literature.

Keywords: *Facial Expression recognition, feature selection, Convolutional Neural Network.*

I.Introduction

1.1 Facial Expression

Human expressions and emotions play a vital role in Emotional Recognition. One's expression helps to know the mood of the human. There are many ways to inspect the recognition of human expressions by their final expressions, voice tone and body language, etc. In this proposed work, we have focused on facial expression recognition. "A facial expression recognition system is an automated system that can analyze the features of a face from a static image" [1].

1.2 Image Annotation

Picture comment and recovery have been a solid region of exploration these days. In picture explanation and recovery, text comment goes about as a significant job. The productive marking of photographs has been dynamic exploration at present. The photographs are clarified dependent on assortments or gatherings as opposed to singular people.

1.3 Deep learning

In the previous 10 years, profound learning made an advancement and delivered cutting edge brings about numerous application spaces, beginning from PC vision, at that point discourse acknowledgment, and all the more as of late, NLP [2]. The renaissance of neural systems are regularly credited to a few variables.

2. Existing System

"The most important ones include the supply of computing power thanks to the advances in hardware (e.g., GPUs), the supply of giant amounts of coaching data, and therefore the power and flexibility of learning intermediate representations [3]. In a nutshell, deep learning uses a cascade of multiple layers of nonlinear processing units for feature extraction and transformation. The lower layers close to the data input learn

ISSN: 2005-4238 IJAST

simple features, while higher layers learn more complex features derived from lower-layer features. The architecture forms a hierarchical and powerful feature representation".

Facial expression recognition has received attention during the last 2 decades. The rationale for facial expression recognition based on motion can be derived from studies in psychology.

Research of Psychologist Mehrabian [4] shows that only 7% of the actual information is transmitted orally and 38% passes by the auxiliary of language such as the rhythm and speed of speech, tone, etc. The information ratio which is transmitted by the expression of the face has reached 55%. "In 1872 Darwin [5] announced the expression consistency. The expression of the face cannot be estimated by the gender and race of the person. Ekman and Friesen [6, 7] made a research work for recognizing the facial expression and they defined to basic categories of facial expressions such as Happiness, Sadness, surprise, fear, disgust and anger" [1].

Such investigations have shown that at any rate 5 feelings are generally connected with unmistakable outward appearances [8]: anger, disgust, happiness, sadness and surprise.

Most mental investigations of outward appearances have utilized mug-shot pictures that catch the statement of the subject at its pinnacle [9]. Just a couple of studies have explored the impact of movement and twisting of facial highlights on the understanding of outward appearances. Bassili [10] recommended that movement in the picture of a face could permit feelings to be distinguished even with the negligible course of action of the highlights.

In the engineering literature, early efforts [11, 12] were based on analysis of the optical flow field of the image sequence, which provides clues to the spatial changes in the facial features [13]. This demonstrated successful facial expression recognition in extensive laboratory experiments involving 40 subjects as well as in television and movie sequences.

In principle, facial expression recognition can be integrated into a face recognition system so the system is robust to expression variations. In practice, however, it seems that moderate, non-dramatic expressions can be handled by many existing face recognition systems.

3. Proposed System

Social networks like Facebook and Whatsapp are used to share their emotions through images. By viewing their Facebook page images or the profile photos status they kept in Whatsapp we can know the emotional status of the person. By analyzing the mood employee's work allotment can be done. Work efficiency, as well as the quality of the work, will be great if the employee is in a good mood. Productivity is synchronized with human emotions. Public opinion is important to predict the results of the elections. TV channels are broadcasting election result predictions through social media.

3.1 Architecture

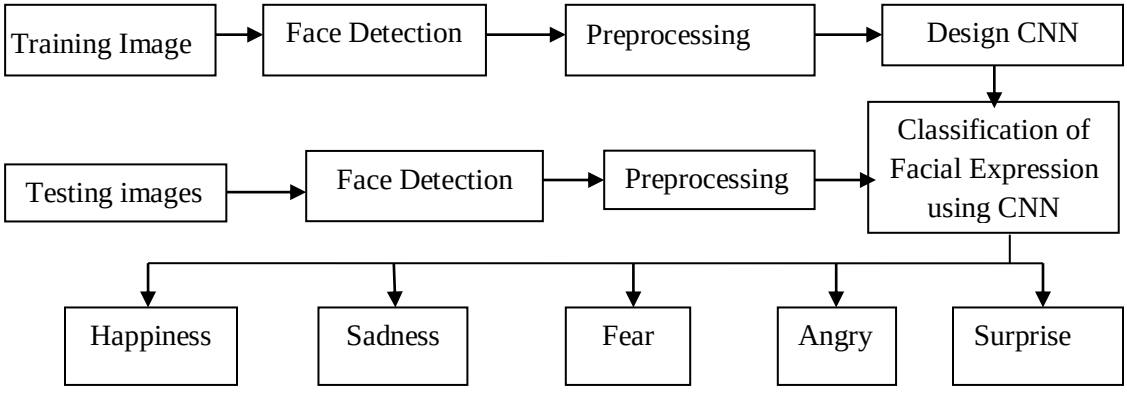

Figure 1. The architecture of the Proposed System

ISSN: 2005-4238 IJAST

The architecture of the proposed system is shown in the figure.
The face regions are detected from input images and preprocessed, Convolution Neural Network is designed for the classification of emotions and test images are classified.

3.2 Haar-cascade Classifier

OpenCV comes with a trainer also as a detector. Here we will deal with detection. OpenCV already contains many pre-trained classifiers for face, smile etc.
Haar Cascade Classifier is loaded as follows:
Import numPy as np
Import cv2
Face_cascade=cv2.CascadeClassifier("haarcascade_frontalface_default.xml")
Img=cv2.imread("arun.jpg")
Gray=cv2.cvtcolor(img,cv2.COLOR_BGR2GRAY
the appearances are presently recognized by the Haar Cascade classifier which restores the situation of identified faces as Rect(x,y,w,h)

3.3 Convolutional Neural Network

"Convolutional neural system (CNN) as appeared in Fig.2 might be an extraordinary kind of feedforward neural system initially utilized inside the field of PC vision. Its structure is propelled by the human visual region, a noticeable instrument inside the creature mind. The visual area contains tons of cells that are liable for detecting light in small and overlapping subregions of the visual fields, which are called receptive fields. These cells act as local filters over the input space. CNN comprises of various convolutional layers, every one of which plays out the capacity that is handled by the cells inside the visual zone" [14].

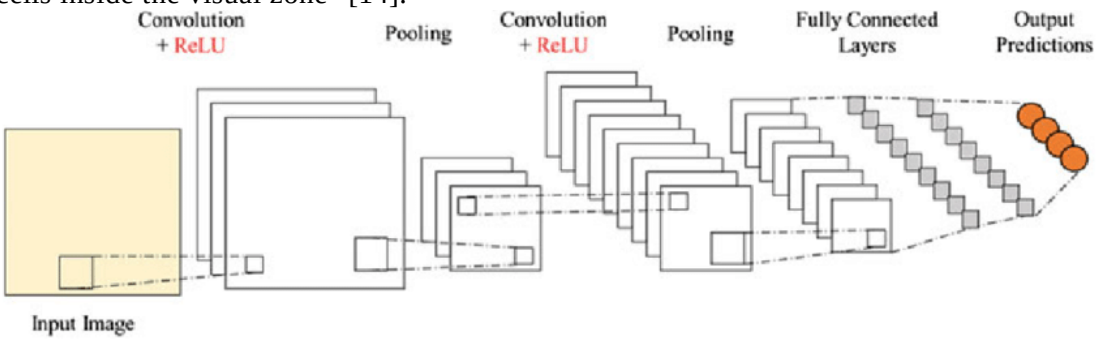

Figure. 2 CNN Architecture

3.4 Classification

Image processing involves some basic operations namely image restoration/rectification, image enhancement, image classification, image fusion, etc. Image classification forms a crucial a part of image processing. The target of image classification is that the automatic allocation of the image to thematic classes [15]. Two sorts of classification are supervised and unsupervised. The strategy for picture characterization includes two stages, preparing of the framework followed by testing. The preparation procedure implies, to require the trademark properties of the photos (structure a class) and structure a particular depiction for a particular class. The technique is finished for all classes relying on the kind of arrangement issue; double order or multi-class grouping. The testing step intends to arrange the test pictures under different classes that the framework

ISSN: 2005-4238 IJAST

was prepared. This relegating of the classification is finished bolstered the parceling between classes upheld the preparation highlights.

Since 2006, profound organized learning, or all the more usually called profound learning or various leveled learning, has risen as a substitution region of AI research [16]. Several definitions are available for Deep Learning; coating one among the various definitions from [16] Deep Learning is defined as: a category of machine learning techniques that exploit many layers of nonlinear information science for supervised or unsupervised feature extraction and transformation and for pattern analysis and classification. This work aims at the appliance of CNN for image classification.

In this proposed work countenance emotions are detected and recognized as Happiness, Sad, Anger and Surprise by CNN. The detected Emotions work are often extended to more applications just like the industry can use this face recognition for his or her employees. Education institutions can adapt this face recognition expression analysis system for his or her students. At home, parents can use the face recognition system for his or her wards. The Media can use this technique to understand the general public opinion before polling to predict the election results.

4.Implementation

4.1 Software

"The deep learning environment has been set by installing Jupyter notebook v0.27.0, python 3.5 versions along side anaconda library. OpenCV (version 3.3.0) library is successfully linked with python 3.5 interpreter" [14]. Using anaconda install command required libraries like numpy, Sequential, Activation, Dropout, Flatten, Dense, ImageDataGenerator, Convolution 2D, Maxpooling 2D, Zeropadding 2D, Optimizers, Applications and Model are imported.

4.2 Dataset Used

Dataset is taken from Cohn Kanade and Whatsapp Profile pictures.

4.3 Training Phase

"The top gestures are captured from an internet camera and labeled as five types like happiness, sadness, anger, disgust and surprise. About 2500 training samples like happiness 500, sadness 500, angry 500, disgust 500 and surprise 500 are used for training. Every frame is resized as 150 x 150 and every one images are converted as grayscale for better performance" [14].

4.4 Testing Phase

About 1000 samples like happiness 200, sadness 200, angry 200, disgust 200 and fear 200 are validated and happiness 50, sadness 50, angry 50, disgust 50 and fear 50 are tested. The pictures are resized as 150 x 150 and converted into gray image. The Haar cascade classifier model classifies the happiness, sadness, angry, disgust and fear.

5. Experimental Results

"In this work, we've used static images of the Cohn-Kanade database for the training process and testing process. The CK database consists of images of university students that they ranged in age from 20 to 35 years. Image sequences for the frontal views are digitized into 640 x 490 pixel array with 8 bits grayscale" [17].

ISSN: 2005-4238 IJAST

Figure 3. Examples of Original Images from Cohn-Kanade Database

Figure 3 shows the pictures of a private that are taken from the Cohn-Kanade database. Image sequences for frontal views and 30-degree views were digitized into either 640x490 or 640x480 pixel array with 8- bit grayscale or 24-bit color values. 95 subjects which have five expressions for every subject and for every expression six static frames are assigned which are used for training the system.

5.1 Confusion Matrix

A confusion matrix may be a table that's wont to project the performance of a classification model on a group of test data that truth values are known. It's visualizing the performance of an algorithm.

Table 1 Confusion matrix for CNN

Actual Class		Predicted Class			
		Happy	Sad	Angry	Surprise
	Happy	50	0	0	0
	Sad	5	35	5	5
	Angry	5	5	30	10
	Surprise	0	0	0	50

5.2 Classification Report

The performance of the annotation system is calculated by using precision and recall. Precision and Recall values are evaluated and thought of because the performance of the system. Accordingly,

$$\text{Precision} = \frac{\text{number of correct annotated label}}{\text{total annotated label}} \quad (1)$$

$$\text{Recall} = \frac{\text{number of correct annotated label}}{\text{total label in the test set}} \quad (2)$$

$$\text{F-score} = \frac{2 \cdot \text{precision} \cdot \text{recall}}{(\text{precision} + \text{recall})} \quad (3)$$

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+FN+TN)} \quad (4)$$

ISSN: 2005-4238 IJAST

Table 2 shows the Classification report

	Precision (%)	Recall (%)	F-Score (%)	Accuracy (%)
Happy	100	83	90	95
Sad	70	88	78	90
Angry	60	86	70	88
Surprise	100	77	87	93
Avg/Total	83	84	81	92

CNN classifies Happy emotion with **the very best** accuracy of 95.00%. CNN produces **a mean** accuracy of 92.00% in classifying emotions.

6. ScreenShots

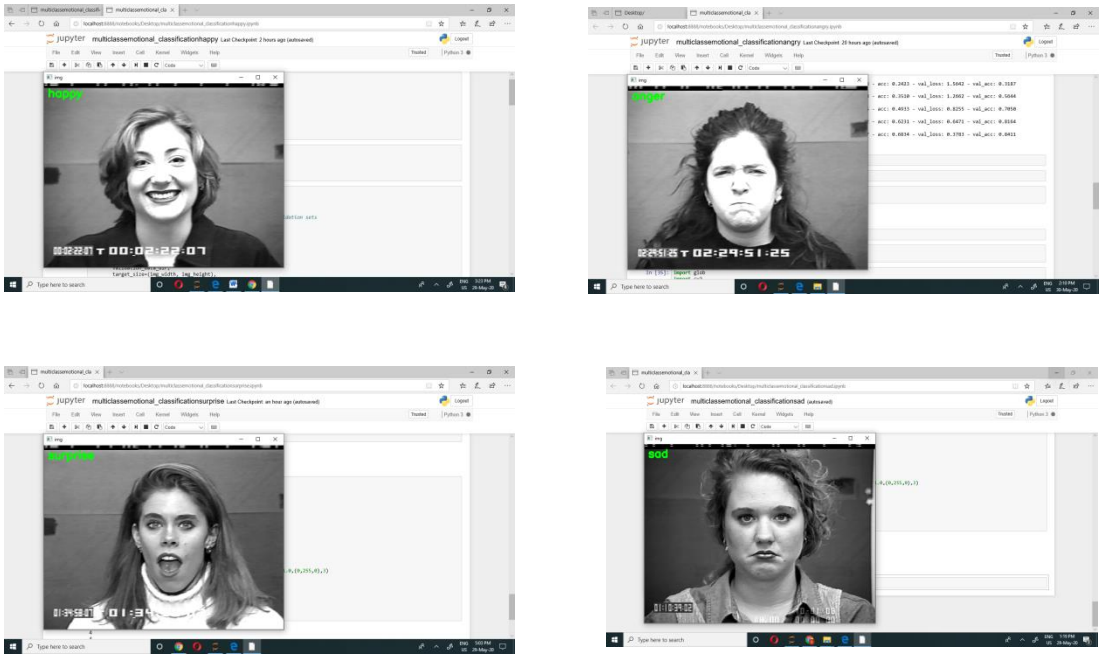

Figure 4. Result of the testing image classification predictions

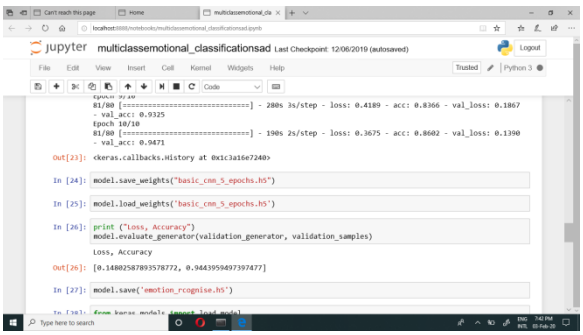

Figure 5. Accuracy of the testing image classifications

ISSN: 2005-4238 IJAST

7. Conclusion

In this paper, the face is detected from the input image which image is assessed supported their emotions. Experimental results validate that the proposed algorithm are able to do competitive performance in terms of accuracy and efficiency. For datasets with quite 200 images, CNN offers a faster prediction speed with a mean accuracy of 92% in comparison to traditional methods. Supported these results, we conclude that CNN may be a promising method for image predictions with large image collections. The work are often extended that images are often annotated automatically.

References:

[1] S. NithyaRoopa, "Research on Face Expression Recognition", International Journal of Innovative Technology and Exploring Engineering (IJITEE), vol. 8, pp. 88-91, 2019.
[2] Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., & Kuksa, "Natural language processing from scratch", Journal of Machine Learning Research, pp. 2461-2505, 2011.
[3] Bengio, Y., Courville, A., & Vincent, P., "Representation learning: A review and new perspectives", IEEE Transactions on Pattern Analysis and Machine Intelligence, 35, pp. 1798–1828, 2013.
[4] Tuark MA, Peintland AP, "Feeling recognition mistreatment eigen faces In: laptop Vision and Pattern Recognition", Proceedings, IEEE laptop Society Conference on. IEEE, pp. 586-91, 1991.
[5] M. El Ayiadi, M. S. Kaamel, and F. Kaerray, "Survey on feeling recognition: options, classification schemes, and information", Pattern Recognition, vol. 44, no. 3, pp. 572– 587, 2011.
[6] Y. Leicun, Y. Beingeo, and G. Henton, "Deep learning", Nature, vol. 521, no. 7553, pp. 436–444, 2015.
[7] J. Schmedhuber, "Deep Learning in convolutional neural networks: an summary", Neural Networks, vol. 61, pp. 85–117, 2015.
[8] P.Ekman, "Facial Expression of Emotion: An old Controversy and New Findings", Philosophical Transactions of the Royal Society of London, vol.335, pp 63-69, 1992.
[9] A.W. Young and H.D.Ellis, Eds., "Handbook of Research on Face processing New York", Elsevier, 1989.
[10] J.N.Bassili, "Emotion Recognition: The role of Facial Movement and the Relative Importance of Upper and Lower Areas of the Face", Journal of Personality and Social Psychology, vol. 37, pp 2049-2059, 1979.
[11] Y.Yacoob and L.S. Davis, "Computing Spatio-Temporal Representation of Human Faces", in Proceedings, IEEE Conference on Computer vision and Pattern Recognition, 1994.
[12] M.Rosenblum, Y.Yacoob and L.Davis, "Human Emotion Recognition from Motion using a Radial Basis Function Network Architecture", in Proceedings, IEEE Workshop on Motion of Non-rigid and Articulated Objects, 1994.
[13] M.Black and Y.Yacoob, "Tracking and Recognising Facial Expressions in Image Sequences, Using Local Parameterised Models of Image motion", Technical Report CS-TR-3401, Centre for Automation Research, University of Maryland, 1995.
[14] R.Indhumathi, A.Geetha, "Emotional Interfaces for Effective E-Reading using Machine Learning Techniques", International Journal of Recent Technology and Engineering, vol.8, pp 4443-4448, 2019.
[15] Lillesand, T.M. and Kiefer, R.W. and Chipman, J.W., "Remote Sensing and Image Interpretation", 5th ed. Wiley, 2004.
[16] Li Deng and Dong Yu "Deep Learning: methods and applications", vol.7, Microsoft, 2014.
[17] S. Fazli, R. Afrouzian and H. Seyedarabi, "High- performance facial expression recognition using Gabor filter and Probabilistic Neural Network," 2009 IEEE International Conference on Intelligent Computing and Intelligent Systems, Shanghai, pp. 93-96, 2009.

Authors

S.Arunkumar, M.C.A.,M.Phil.,
Research scholar in Department of Computer and Information Science at Annamalai University. He has been completed his Master of Computer Applications degree in the year 1998 at Annamalai University, Chidambaram, Tamilnadu, India and completed his Master of Philosophy in Computer Science in the year of 2008 at Annamalai University, Chidambaram, Tamilnadu, India. He also published one International journal. His area of research is image processing.

Dr. A. Geetha, M.E., Ph.D.,
She obtained her B.E degree in Computer Science and Engineering from Periyar Maniammai University, Thanjavur in the year 1994. She received her M.E degree in Computer Science from

Regional Engineering College, Trichy in the year 1997. She has been working in Annamalai University, since 1999. She was awarded Ph. D in Computer Science and Engineering at Annamalai University in the year 2014. Presently she works as a Professor in the same department. She has published around 40 papers in national & international journals and conferences. She has coordinated a workshop in the field of speech and image processing. An international conference was also organized by her to her credit. Also she was one of the editors for the International Journal of Applied Engineering Research which was published in the international conference. Two book chapters were published by her. She has won the best paper award in National Conference on Multimedia Signal Processing. She has delivered special lectures on several topics in her area, in workshops and staff development programme. Presently she is the coordinator of the Value Added Course, FEAT, and Annamalai University. She has visited and presented papers in international conferences conducted in Malaysia and Japan. Currently, she is a member of the Doctoral Committee in the Annamalai University and Research Advisory Committee at Anna University. Her areas of research interest include pattern classification, image & video processing and human-computer interaction.

ISSN: 2005-4238 IJAST