Facial Expression Using Deep Convolutional Neural Networks for Higher Secondary Students in Class

M.Lakshana, R. Devi Prabha, M.Deeptihananthini
B.E., Computer Science and Engineering, K.Ramakrishnan College of Technology, Samayapuram, Tiruchirappalli-621112, Tamil Nadu, India.

Abstract

As we are stepping forward from one generation to another, innumerous technologies are abiding us according to our necessities. Facial expression recognition is still a challenging problem that waits for more and more accurate algorithms. This project presents a new model that is capable of recognizing facial expression by using deep Convolutional Neural Network (CNN).

Keywords --- *Deep Convolutional neural networks, Facial Expression*

I. INTRODUCTION

A Convolutional neural network (CNN) is a class of deep neural networks, most commonly applied to analyzing visual imagery. Facial recognition could be a class of biometric code that maps somebody's facial expression mathematically and stores the information as a face print.

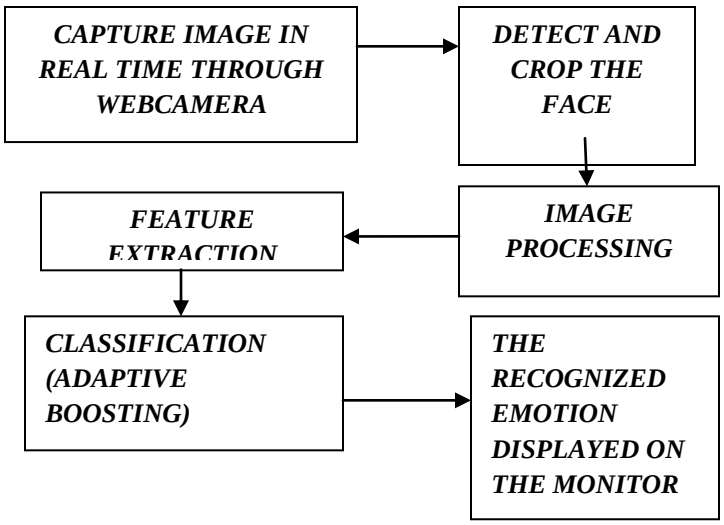
Facial expression recognition methods are separated into two categories: geometric-based methods and appearance-based methods. Geometric-based methods concern about the feature vectors encoding some facial geometric properties such as distance, angle, and position to determine the shapes and locations of the invariance points of face. For instance, invariance points belong to a face image was extracted for facial expression recognition.

Facial expression recognition problem constructs a classification algorithm to separate emotions into the classes such as sadness, surprise, anger, happiness, fear, disgust, and neutrality by using the extracted features. In proposed system, it recognizes facial expression of the students in the classroom by capturing image of students during class hours. Facial expressions are classified into interest, boredom, amazement, distraction, happiness, anger, sad, disgust, fear and neutrality. Appearance-based methods use the features extracted directly from the images but does not include information relating to the facial points. There are a lot of Appearance based methods. The most important ones are Local Binary Pattern (LBP) , Gabor Wavelets , Local Gabor Binary Patterns (LGBP) , Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradient (HOG) , and Curvelet Transform.

Artificial Neural Networks (ANNs), Support Vector Machines, (SVMs), Spherical Classifiers, Hidden Markov Models (HMMs), K-Nearest Neighbors (KNNs), and compressive sensing based sparse classifiers for recognizing facial expressions. Facial expressions make the certain regions of face change, which causes interest in just the special regions. The salient features were extracted from local patches.

Classifier receives descriptive features obtained from above methods as its inputs. Hence the classifier performances depend on the quality of feature vectors. Moreover the Methods extract low-level features belong to facial expression but don't extract high-level feature. Moreover the methods extract low-level features belong to facial expression but don't extract high-level feature. However recently the methods as referred to as deep learning have appeared as a promising one to perform facial expression recognition.

II. STRUCTURE



III. METHODOLOGY

Facial expression recognition methods are separated into two categories:

- geometric-based methods
- appearance-based methods

There are different approaches to create face recognition systems: graph matching techniques.

Graph matching is the problem of finding a similarity between graphs.

Graphs are commonly used to encode structural information in many fields, including computer vision and pattern recognition, and graph matching is an important tool in these areas. In these areas it is commonly assumed that the comparison is between the data graph and the model graph.

Exact graph matching is known as the graph isomorphism problem. The problem of exact matching of a graph to a part of another graph is called sub graph isomorphism problem.

Inexact graph matching refers to matching problems when exact matching is impossible, e.g., when the number of vertices in the two graphs are different. In this case it is required to find the best possible match.

There are a lot of Appearance based methods. The most important ones are Scale Invariant Feature Transform (SIFT), Histogram of Oriented Gradient (HOG), K-Nearest Neighbors (KNNs),

Scale-invariant feature transform (**SIFT**) is a feature detection algorithm in computer vision to detect and describe local features in images. Applications include object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, individual identification of wildlife and match moving.

Histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, scale-invariant feature transform descriptors, and shape contexts, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy.

K-Nearest Neighbors (KNNs)

In pattern recognition, the k-nearest neighbors algorithm (k-NN) is a non-parametric method used for classification and regression .In both cases, the input consists of the *k* closest training examples in the feature space. The output depends on whether k-NN is used for classification or regression:

- In k-NN classification, the output is a class membership. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its *k* nearest neighbors (k is a positive integer, typically small). If $k = 1$, then the object is simply assigned to the class of that single nearest neighbor.
- In k-NN regression, the output is the property value for the object. This value is the average of the values of its k nearest neighbors.

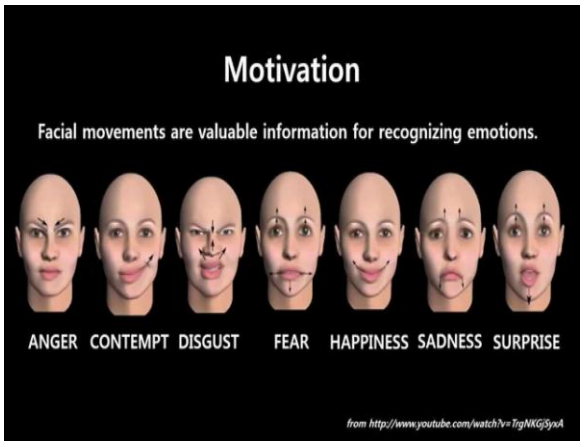
Hidden Markov Model (**HMM**) is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved (i.e. hidden) states. The hidden Markov model can be represented as the simplest dynamic Bayesian network.

IV. FACIAL EXPRESSION RECOGNITION USING DEEP CONVOLUTION NEURAL NETWORK

Human facial expressions are extremely essential in social communication. Normally communication involves both verbal and nonverbal. Non-verbal communications are expressed through facial expressions. Face expressions are the delicate signals of the larger communication. Non-verbal communication means communication between human and animals through eye contact, gesture, facial expressions, body language, and paralanguage.

Eye contact is the important phase of communication which provides the mixture of ideas. Eye contact controls the contribution, discussions and creates a link with others. Face expressions include the smile, sad, anger, disgust, surprise, and fear. A smile on human face shows their happiness and it expresses eye with a curved shape. The sad expression is the feeling of looseness which is normally expressed as rising skewed eyebrows and frown. The anger on human face is related to unpleasant and irritating conditions. The expression of anger is expressed with squeezed eyebrows, slender and stretched eyelids. The disgust expressions are expressed with pull down eyebrows and creased nose. The surprise or shock expression is expressed when some unpredicted happens. This is expressed with eye-widening and mouth gaping and this expression is an easily identified one. The expression of fear is related with surprise expression which is expressed as growing skewed eyebrows.

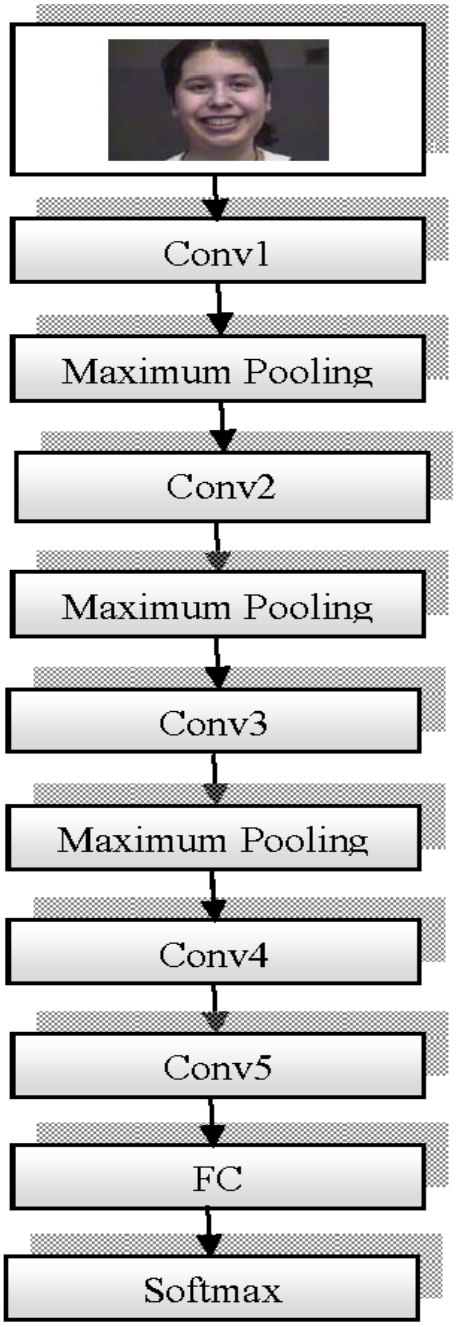
FER has the important stage is feature extraction and classification. Feature extraction includes two types and they are geometric based and appearance based. The classification is also one of the important processes in which the above-mentioned expressions such as smile, sad, anger, disgust, surprise, and fear are categorized. The geometrically based feature extraction comprises eye, mouth, nose, eyebrow, other facial components and the appearance based feature extraction comprises the exact section of the face.



The slow signals are permanent wrinkles which include the changes in facial appearance such as muscle tone and skin texture changes that happen slowly with time. The rapid signals are raising the eyebrows which include the face muscles movement, impermanent face appearance changes, impermanent wrinkles and changes in the location and shape of facial features. These flashes on the face remain for a few seconds. These three signals are altered with individual option while it is very hard to alter static and slow signals.

CNN is an updated version of multi-layer neural networks. CNNs receive as the input the images unlike conventional neural networks. The basic layers of CNNs are called as convolutional, pooling, rectified linear units, fully Connected and loss layer. The input size is reduced to the end layers. In the sequential layers, the properties from low-level and high level Features are extracted. In convolutional layer of CNNs, each input image is convoluted with kernels. The obtained output images are imported to the second layer.

In the pooling layer of CNNs, it is selected the salient features in the previously determined regions. In other way, each region is down–sampled by a non-linear down-sampling operation such as maximum, minimum or average. The fully connected layer of CNNs is located after the above defined layers. All neurons at the layer are fully connected to all activations in the previous layer. In the loss layer of CNN, different loss function is applied .For instance, softmax loss function is used for multi class classification problems



e graphical pipeline of the proposed CNN m

V. CONCLUSION AND FUTURE WORK

In this project, a new facial expression recognition algorithm is proposed based on CNN. An appropriate CNN architecture was selected. Convolutional layer numbers, kernel and stride sizes, and pooling structures were determined. The proposed system can be applied at real life.

Expression recognition is a task that humans perform daily and effortlessly, but it is not yet easily performed by computers, despite recent methods have presented accuracies larger than 95% in some conditions (frontal face, controlled environments, high-resolution images).

Some limitations of image processing are image quality, image size, face angle, processing and storage.

Image processing, image size, face angle affects how well facial-recognition algorithms work. The image quality of scanning video is quite low compared with that of a digital camera.

The proposed system can be applied at real life applications. However, recently the Method called deep learning has appeared as a promising one to perform facial expression recognition.

REFERENCES

[1] D. C. Ali Mollahosseini and M. H. Mahoor. Going deeper in facial expression recognition using deep neural networks. IEEE Winter Conference on Applications of Computer Vision, 2016.

[2] C. Ruoxuan, L. Minyi, and L. Manhua "Facial Expression Recognition Based on Ensemble of Multiple CNNs," CCBR 2016, LNCS 9967, pp. 511-578, Springer International Publishing AG 2016.

[3] D. C. Ali Mollahosseini and M. H. Mahoor. Going deeper in facial expression recognition using deep neural networks. IEEE Winter Conference on Applications of Computer Vision, 2016.

[4] Networks outperform state-of-the-art methods on the well-known FERC-2013 dataset provided on the Kaggle facial expression reorganization.

[5] Bettadapura, Vinay (2012). Face expression recognition and analysis: the state of the art. Ar Xiv preprint ar Xiv: 1203.6722.

[6] A. Krizhevsky, I. Sutskever, G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 2012: 1097 – 1105.

[7] Akputu, K. O., Seng, K. P., & Lee, Y. L. (2013). Facial Emotion Recognition for Intelligent Tutoring Environment. In *2nd International Conference on Machine Learning and Computer Science (IMLCS'2013)* (pp. 9-13).

[8] Challenges in representation learning: Facial expression recognition challenge. https://www.kaggle.com/c/challenges-in representation-learning-facial expression-recognition-challenge, 2013.

[9] S. Kherchaoui and A.Houacine, 2010, "Face Detection Based On A Model Of The Skin Color With Constraints And Template Matching", Proc. 2010 International Conference on Machine and Web Intelligence, pp. 469 - 472, Algiers, Algeria.

[10] Y. Miyakoshi, S. Kato, "Facial emotion detection considering partial occlusion of face using Bayesian network," in: Proceedings of IEEE Symposium on Computers & Informatics (ISCI), pp. 96–101, 2011.

[11] L. Zhang, D. Tjondronegoro, V. Chadran, "Gabor based templates for facial expression recognition in images with facial occlusion," Neurocomputing, Vol. 145, no.5, pp. 451–464, 2014.

[12] Z. Zhang , L. Wang , Q. Zhi , S. Chen , Y. Chen, "Pose invariant face recognition using facial landmarks and Weber local descriptor," Knowl.Based Syst., Vol. 84, pp. 78–88, 2015.

[13] .13. N. Perveen, S. Gupta, K. Verma, "Facial expression recognition using facial characteristic points and Gini index, in: IEEE StudentsConference on Engineering and Systems, pp.1–6, 2012.

[14] S. He, S. Wang, Y., Lv. "Spontaneous facial expression recognition based on feature point tracking," in: IEEE2011 Sixth InternationalConference on Image and Graphics (ICIG), pp.760–765, 2011.

[15] A.Lorincz, L.A.Jeni, Emotional expression classification using timeseries kernels, in: IEEE Conference on Computer Vision and PatternRecognition, pp.889–895, 2013.

[16] S. Bashyal, and G. Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization,"Eng. Appl. Artif. Intell., Vol. 21, no. 7, pp.1056_64, Oct. 2008.

[17] T. H. Zavaschi, A. S. Britto Jr, L. E. Oliveira, A. L. Koerich, "Fusion of feature sets and classifiers for facial expression recognition," ExpertSyst. Appl., Vol. 40, no. 2, pp. 646–655, Feb. 2013.

[18] G. Littlewort, M.S. Bartlett, I. Fasel, J. Susskind, J. Movellan,"Dynamics of facial expression extracted automatically from video",Image Vis. Comput., Vol.24, no.6, 615–625, 2006.

[19] C. Shan, S.Gong, P.W. McOwan, "Facial expression recognition based on local binary patterns: a comprehensive study," Image Vis. Comput.Vol. 27, no.6, pp. 803–816, 2009.

[20] G. W, Xiang C, Venkatesh YV, D. Huang, H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis," Pattern Recognit., Vol. 45, no. 1, pp. 80–91, 2012.

[21] X. Sun, H. Xu, C. Zhao, J.Yang, "Facial expression recognition based on histogram sequence of local Gabor binary patterns," in: IEEE Conference on Cybernetics and Intelligent Systems, pp.158–163, 2008.