AN ADVANCED TECHNIQUE FOR FACIAL EMOTION DETECTION USING ARTIFICIAL INTELLIGENCE

[1]M. Sai Krishna Murthy, [2]R.Radhika, [3]M. A. Nidhi, [4]K. SaiSandeep, [5]K. Chaithanya Kumar
[1]Assistant Professor,[2,3,4,5]B. Tech,
[1,2,3,4,5]Department of Computer Science and Engineering,
[1,2,3,4,5]St.Martin's Engineering College, Hyderabad, India

Abstract: Facial expression of a person tells us about their emotion. In the same way a machine can also detect the face of a human and recognize the emotion based on the facial expression. If we expect a machine to exhibit human like performance Artificial Intelligence is needed. In this prototype system, machine automatically recognizes the emotion represented on a face when our face is detected in the webcam. Thus a Convolutional Neural Network based solution combined with image processing is used in classifying the universal emotions.

Keywords **- Artificial Intelligence, Dataset, Support Vector Machines(SVM),Active Shape Model(ASM), Active Appearance Model (AAM), Relevance Vector Machine(RVM), gabor filter, KNN Regression algorithm, eight faces, Electroencephalography (EEG) , openCV, numpy, keras, dlib and tensorflow, Convolutional neural networks, AlexNet, ResNet, Inception, Visual Geometry Group (VGG)**

I. INTRODUCTION

Ekman and Friesen proposed six emotional expressions which are said to be universal. They are happy, sad, disgust, surprise, neutral and fear.To make a system recognize these emotions,it needs to be trained with a dataset containing images of people with different expressions. Convolutional Neural Networks (CNNs) are used for this training process.

Facial emotion detection can be done with some of the popular networks like AlexNet, VGG (Visual Geometry Group), Inception and ResNet (Residual Networks). In our work we have used Lighter Version of VGG for training the system for facial emotion recognition.

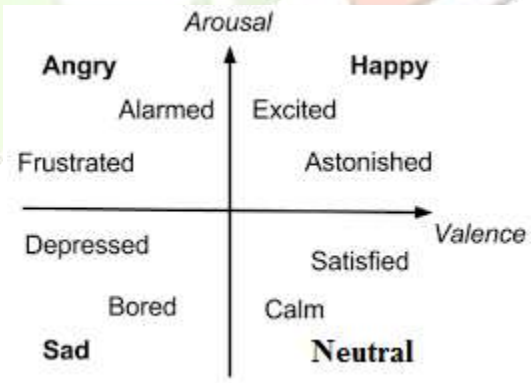

FIG 1: A Two - Dimensional Representation of Emotion

II. LITERATURE SURVEY

Myunghoon Suk; Prabhakaran, B. et al. [7] developed a system which uses a set of Support Vector Machines(SVMs). Active Shape Model(ASM) is used for extracting facial expression features and generation of dynamic features can be done by displacement between neural and expression features for emotion recognition.

Kai-Tai Song; Chao-Yu Lin et al. [3], presented a reinforced approach for enhancement of emotion recognition from facial images. Active Appearance Model(AAM) is used for computation of shape and texture models of facial images. From this facial

feature points and geometrical feature values were extracted. Relevance Vector Machine(RVM) uses the extracted features for recognizing the emotional states.

Singh, M.; Majumder, A.; Behera, L.et al. [6] presented a facial expression recognition system which uses Bayesian network. Probabilistic modelling which is used for drawing relationship between facial features, action units is used for training the network and finally recognizes six basic emotions. To get Geometric feature vector they proposed feature extraction methods. Geometric feature vector contains angular information. Over certain facial regions gabor filter is applied. The moments extracted from here are present in appearance feature vector. With the help of these feature vectors relationships are drawn among Action Units.

NavdeepKaur, Er. VarinderjitKaur et al. [8] have proposedan algorithm for facial expression detection called KNN Regression algorithm with SURF feature. Firstly, the eigenspacewith eigenvalues and eigenvectors were created. Using Principal Component Analysis (PCA) the most relevant eight faces were selected. But before selection from this space, the eightfaces will be constructed.

III. OVERVIEW

In present times, the solutions are accessible but to a limitation which will not help common people. Two major scientific fields are being used to attain Emotion Recognition so far. They are:

- Interactive Voice Response (IVR) system that uses audio signal analysis: The solution of Emotion Recognition depends on the emotion that has to be recognized and the purpose of doing it. The wide range of applications of Emotion Recognition are the talking toys, video and computer games, and call centers. Of which the call centers are specifically interested in the application of this technology for IVR systems. An apparent example of this is the automatic call routing of angry customers to the customer representatives and monitoring the quality of agents' performance. These systems are deliberate and thus the utterances are short. This technique though effective has a flaw that some customers are very cool and calm while the discussion however inside they may be furious. Such calls may go undetected while applying Emotion Recognition using IVR. Additionally, only 20 per cent of what is said is handled by the electronic telephony while noticing the frequency range of human voice. Thus IVR systems are not reliable.
- Emotion Recognition using Brain Activity: Electroencephalography (EEG) is a monitoring method to record electrical activity of the brain. To analyze EEG signals these signals propose a system. This system is able to categorize them into 5 classes based on two emotional dimensions, valence and arousal. This system is meant to test the quality of recognizing emotions using EEG signals. So, a dataset with EEG signals is used to perform this assessment by measuring these signals from the people that were stimulated by the pictures. This way enabled to teach the system the relationship between the brain activity and emotions. These EEG signals have enough data to differentiate five different classes on valence and arousal dimensions. Eventually, using a 3-fold cross validation method for training and testing, the classification rates of 32% for valence dimension and 37% for arousal dimension were reached. And by using only the extreme values, the rates were 71% and 81%.

On comparison, the latter is considered to be better but is highly expensive and out of reach of common people. EEG cannot be used in common practice and needs a lot of computational expertise. Hence, a reliable solution to detect emotions is done in this project that uses digital images which is easy to acquire and analyse.

Our approach will be beneficial for the common people as with the advancement of the technology users can communicate with each other through their cell phones and will be able to recognize the emotion prior the interaction. This technology will help the call centres to check the employee-customer relationship. Even the lie detectors can be updated by using all these methods were better results can be achieved. And the methods of detecting smiling faces can be extended further in the industry of Digital Camera.

3.1Architecture

The aim of our work is to use computer vision techniques to detect the face and recognize the emotion from the frames through webcam given as input.

In our work we used Python as a programing language and the libraries used areopenCV, numpy, keras, dlib andtensorflow.

The Basic architecture in our work is as follows

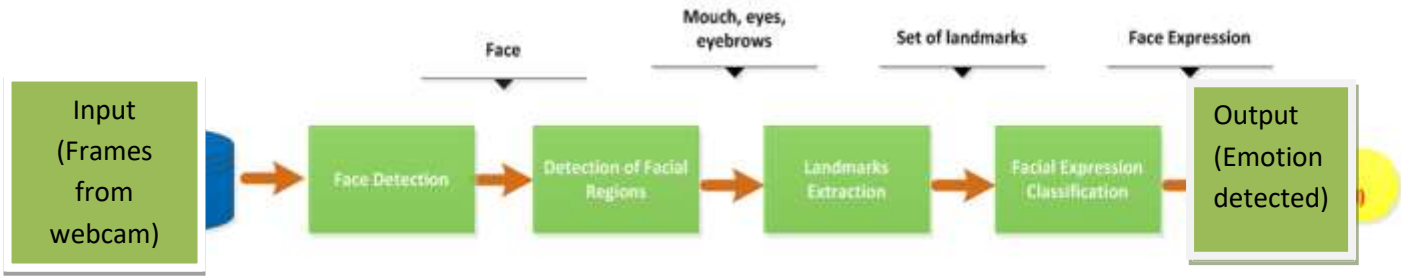

FIG 2: Architecture

3.2 Phases

In our work, training the system should be done first. And is done by using lighter versions of VGG to which 48X48 input images are given. Here is the brief description of phases included in our work.

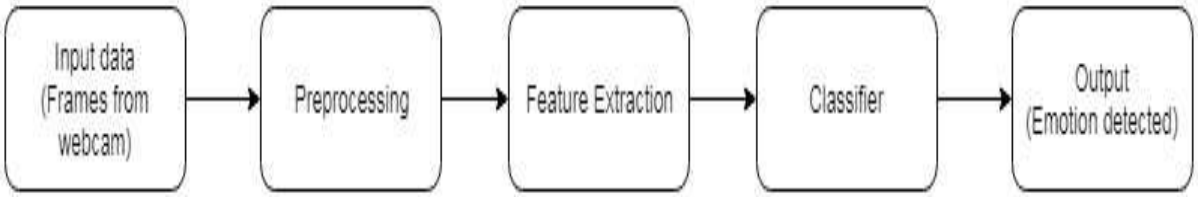

FIG 3: Phases for Facial Emotion Detection

Frames from the webcam are given as input to the system. In our project we have used dlib library for face detection. Using this library the face will be detected and converted to grayscale image in the pre-processing phase and mouth, eyes, eyebrows which are the main facial regions for emotion detection are detected in the feature extraction phase. In the classifier phase, the system will predict the emotion on our face. Based on the training we give to the system, the system recognizes the emotion of our face expressions. If the training is perfect with large data set we can achieve correct results. This process is done when at least one face is detected through webcam. In our project the system detects the emotion for every ten frames from through the webcam. System is able to detect and recognize the emotion for any number of faces at a time.

3.3 Libraries Used

Below are the libraries used in these phases

OpenCV which is basically called as Open Source Computer Vision is a machine learning software Library. This supports windows, Linux, iOS, Android and Mac OS and has C++, JAVA, Python interfaces.

Numpy is a fundamental package for scientific computing exclusively in python. This library is used for providing support for large multi-dimensional arrays or matrices. This library consists of large number of mathematical functions which are useful to operate on the multi-dimensional arrays.

Tensorflow is a framework which comprises of tensors. A tensor is basically generalization of matrices or vectors to high dimensions. Tensorflow defines and runs computations involving these tensors. Internally tensors are represented as n-dimensional arrays.

Keras is written in python programming language which is a high level Neural Networks API. Keras is developed for fast experimentation and this is capable of running on top of tensorflow.

Dlib is a toolkit that contains tools to solve real world problems and create complex software. This also contains machine learning algorithms. Some of the domains in which dlib can be used are robotics, embedded devices, mobile phones and high performance computing environments.

IV. METHODOLOGY

Our work describes a convolutional neural networks based solution for facial emotion detection.

Convolutional Neural Networks is a kind of neural networks which is very effective in areas like image recognition and imageclassification. Convolutional Neural Networks is very successful in identifying faces or objects.CNN consists of one or more convolution layers which are followed by fully connected layer/s. Convolution layer is the core building block of Convolution network.

Facial emotion detection can be done with following network architectures

4.1 AlexNet

AlexNet is proposed by Alex Krizhevsky which consists of five convolutional layers and three fully connected layers. Tanh and sigmoid activation functions were earlier standards for traditional neural networks. In this network instead of using earlier standards, ReLU (Rectified Linear Unit) is used for non-linear part.

By using dropout layer after fully connected layer AlexNet reduces outfitting problem.

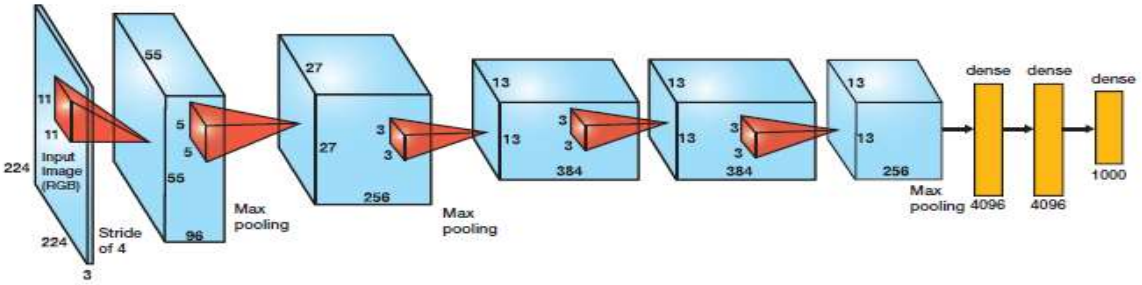

FIG 4: AlexNet

4.2 Inception

Inception is also called as GoogLeNet. This network is built on an idea that in deeper networks most of the activations are either redundant because of correlation between them or unnecessary. Every output channel is connected to every input channel in every convolutional operation at one location in Inception. Hence it is called as Dense Connected architecture. This network devised a module called Inception. The major point of this module is that it has a bottleneck layer which is helpful in massive reduction of computational requirement. Before feeding into the convolutions, reduction of dimension of the input channels is done. So, Inception module uses 1X1 convolutions. Later, larger sized kernels are applied.

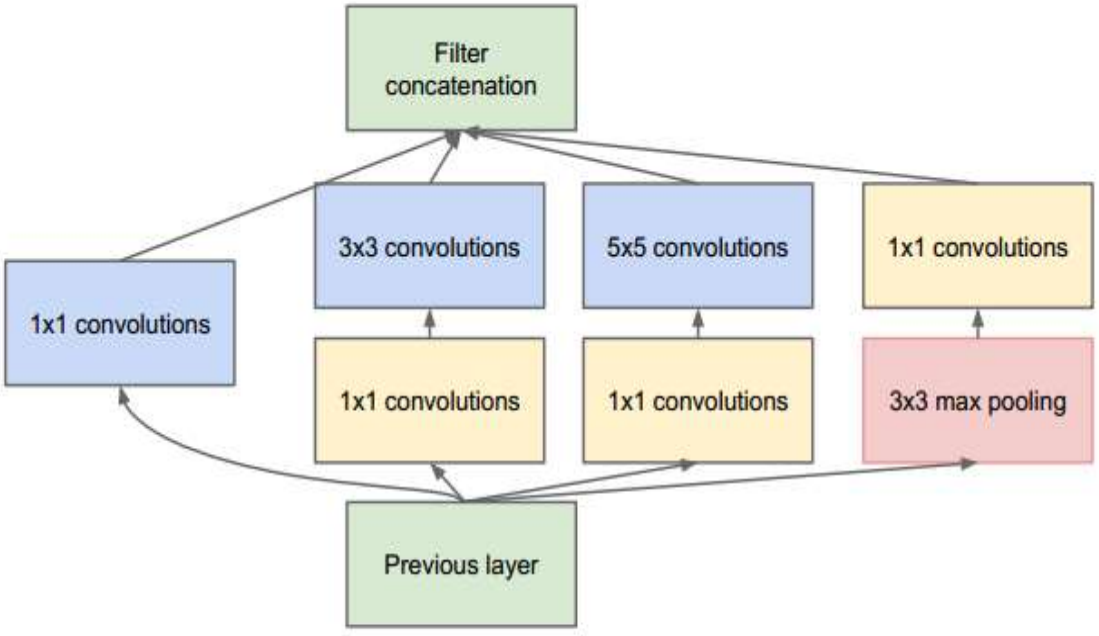

FIG 5: Inception

4.3 ResNet

Basically, if we increase the depth then accuracy of the network increases when over fitting problem is taken care. If we perform optimization on huge parameter space and there by naively add the layers then this may lead to high training error. But residual

network (ResNet) allows training of deep networks. Construction of networks through modules called residual modules is done in ResNet.

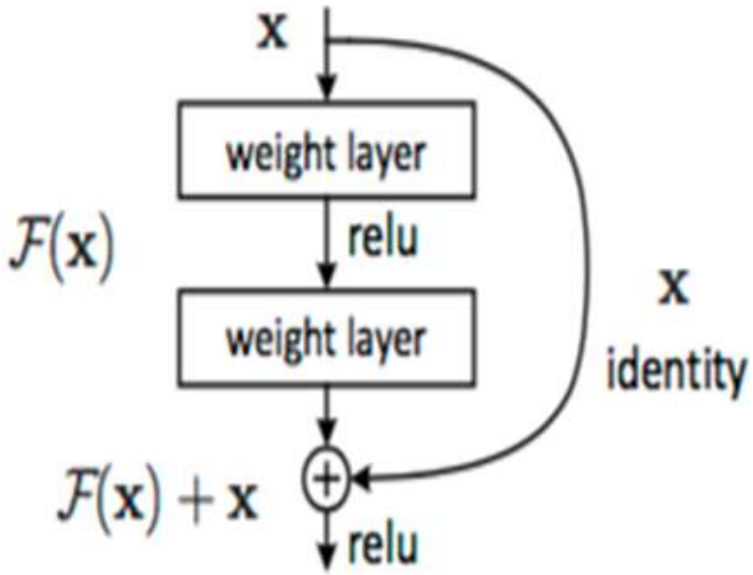

FIG 6: ResNet

4.4 VGG Network

In our work we have used VGG network for training the system.

VGG stands for Visual Geometry Group which consists of five convolutional layers and three fully connected layers. It replaces larger kernel sized filters used in AlexNet with multiple 3X3 kernel sized filters. Multiple non-linear layers increase the depth of the network. This makes the system learn more complex features at low cost. So using multiple stacked smallersized kernel is better than the larger sized kernel. Thus, we used the lighter version VGG. The activation functions we have used are ReLU and softmax.

ReLU stands for Rectified Linear Unit. This is the most used activation function. We can find this activation function in most of the convolutional neural networks or deep learning. This is half rectified. The function and its derivative are monotonic. The other activation function used is softmax. Softmax activation function is a more generalized logistic activation function which is used for multiclass classification.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Network	A,A-LRN	B	C	D	E
Number of parameters	133	133	134	138	144

FIG 7: VGG

V. RESULTS

The overall result of the project is positive.The key phase in our work is the Pre-processing step. The face is properly detected by the system using dlib library in openCV. The network used enhances the effectiveness of a machine in performing tasks. For every ten frames, the system detects the emotion correctly. We have trained the system using a lighter version of VGG network which has been completed successfully.A test was conducted on several individual and found that the results are positive. This is a low cost solution which gives 90% correct results.

VI. CONCLUSION

A VGG based solution gives us good results and our work has high success rate where the solution is independent of many factors like skin colour, gender, face shape and so on. Thus, this solution will work reliably for any person's face. This solution is very easy to use and reduces the learning work. This work can recognize emotions of many faces at a time and the algorithm is more flexible and can be used in real time applications. This can be used as a base to take decisions by a machine. Further improvements can be done by how a machine can recognize the emotion and mood of a person based on both facial expressions and gestures.

BIBLIOGRAPHY

[1] "Face Recognition: A Convolutional Neural-Network Approach," Steve Lawrence, Member, IEEE, C. Lee Giles, Senior Member, IEEE, Ah Chung Tsoi, Senior Member, IEEE, and Andrew D.Back, Member, IEEE.

[2] J. Schmidhuber, "Deep learning in neural networks: An overview,"Neural Networks, vol. 61, pp. 85–117, 2015. 76

[3] Kai-Tai Song, Chao-Yu Lin, "Robust facial emotion recognition using a temporal-reinforced approach,"14th International Conference on Control, Automation and Systems (ICCAS), pp. 804-807, 22-25 Oct. 2014.

[4] M. Abdulrahman, T.R. Gwadabe, F.J. Abdu, A. Eleyan, "Gabor wavelet transform based facial expression recognition using PCA and LBP," 2014 IEEE 22nd Signal Processing and Communications Applications Conference (SIU), pp. 2265-2268, 23-25 April 2014.

[5] M.C. Sobia, V. Brindha, A. Abudhahir, "Facial expression recognition using PCA based interface for wheelchair," 2014 International Conference on Electronics and Communication Systems (ICECS), pp. 1-6, 13-14 Feb. 2014.
[6] M. Singh, A. Majumder, L. Behera, "Facial expressions recognition system using Bayesian inference," 2014 International Joint Conference on Neural Networks (IJCNN), pp. 1502-1509, 6-11 July 2014.
[7] M. Suk, B. Prabhakaran, "Real-Time Mobile Facial Expression Recognition System -- A Case Study," 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 132-137, 23-28 June 2014.
[8] NavdeepKaur, Er. VarinderjitKaur, "Improved Emotion Detection by Regression Algorithm with SURF Feature and SVM", International Journal of Engineering and Computer Science, vol. 3, no. 10, pp. 8639-8642, October, 2014.
[9] Real Time Facial Emotion Recognition based on Image Processing and Machine Learning(International Journal of Computer Applications (0975 – 8887) Volume 139 – No.11, April 2016).
[10]StefanosZafeiriou, Cha Zhang and Zhengyou Zhang, "A Survey on Face Detection in the wild: past, present and future," 2015.
[11] Swati Mishra, Avinash Dhole" A Survey on Facial Expression Recognition Techniques," 2013.
[12] WathsalaNayomiWidanagamaachchi"Facial Emotion Recognition with a Neural Network Approach" August 2009.