**MINOR-1 PROJECT**

**SYNOPSIS on**

**MELANCHOLY RADAR TOWARDS CLIENT SEGMENTATION**

Submitted By :

|  |  |  |
| --- | --- | --- |
| **Name** | **Roll No** | **Branch** |
| Sagar Jain | R970215059 | CSE OGI |
| Nikunj Patel | R970215045 | CSE OGI |
| Harsh Mittal | R970215029 | CSE OGI |
| Harshita Dwivedi | R970215031 | CSE OGI |

**Under the guidance of**

**Dr. Durgansh Sharma**

**Associate Professor**

**HOD Department of Symantics**



School of Computer Science

**UNIVERSITY OF PETROLEUM AND ENERGY STUDIES**

**Dehradun-248007 2017-18**

**Approved By**

(Dr. Durgansh Sharma) (Dr.Nilima. S. Fulmare)

**Project Guide Department Head**



**School of Computer Science**

University of Petroleum & Energy Studies, Dehradun

**Synopsis Report (2017-18)**

**1 Project Title**

MEANCHOLY RADAR TOWARDS CLIENT SEGMENTATION

# Abstract

Facial expressions depicting various moods play an essential role in communications in social interactions with other human beings which deliver rich information about their emotions. The most crucial feature of human interaction that grants naturalism to the process is our ability to infer the emotional states of others. Our goal is to categorize the various human moods from their eye and lips expressions. The proposed system presents mood radar that analyses the human eye and lips region from different frames. As we are heading towards a world working with bots, handling artificial intelligence, working to bring more digitisation to the different sectors, understanding human behaviour is a growing emergence. As it would be a harder task to convince a person a work if he/she is in bad mood also positive mood may lead to more positive expectations concerning source trustworthiness or likability than a negative mood, therefore, our project can be used to give a predictive analyses. The Convolutional Neural Networks and the ReLU function are used in the system which extract the features from the image and classify it and it will rectify the feature image.The system uses the image data for the training of the system. We are taking the review of the client by the picture and scale it on the basis of the moods.

Keywords: CNN, ReLU, Moods, review, picture.

# Introduction

Human emotion recognition is an important component for efficient human computer interaction. It plays a critical role in communication, allowing people to express oneself beyond the verbal domain. Analysis of moods from human eye and lips expression involves the detection and categorization of various human moods or state of mind. For example, in security and surveillance, they can predict the offender or criminals behaviour by analysing the images of their faces from the frames of the video sequence. State-of- the-art face recognition is dominated by industry- and government-scale datasets. Example applications in this space include person of interest identification from mounted cameras and tagging a users friends in pictures. Training is often an offline, batch operation and produces a model that can predict in hundreds of milliseconds. The time to train new classification models in these scenarios isnt a major focus because the set of people to classify doesnt change often. Many of the most successful marketing campaigns and initiatives are focused on moods[1].

TakeDoves Real Beauty campaign, which was an attempt to change the conversation around womens beauty. OrNikes Just Do It campaign, firing up the inner-athlete in everyone OrMasterCards Priceless campaign, which has been powering brand success for a whopping 19+ years.

All of this marketing aims to move your heart, and then relies on the experience to build a lasting bond with the brand. What it doesnt do is rely on feature sets, discounts, or new technology to move you to a sale. None of it speaks to the number of offices, employees, or years in business of the brand.Google Senior VP of Global Marketing,Lorraine Two hill,states, If we dont make you cry, we fail. Its about mood[2].

In this project we are using different technologies like CNN, ReLU. Convolutional Neural Networks (CNN)[3] are a category ofNeural Networksthat have proven very effectivein areas such asimage recognition and classification. CNN have been successful in identifying faces, objects and trafficsigns apart from powering vision in robots and self driving cars. CNN being used forrecognizing everyday objects, humans and animals. Lately, CNNhave been effective in several Natural Language Processing tasks (such as sentenceclassification) as well.

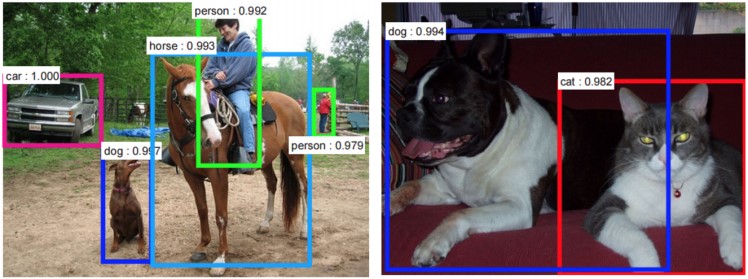


Figure 1: *Feature Extraction[4]*

The basic building blocks ofeveryConvolutional Neural Network.

Channelis a conventional term used to refer to a certain component of an image. An image from a standard digital camera will have three channels red, green and blue you can imagine those as three2dmatrices stacked over each other (one for each color), each having pixel values in the range 0 to255. Agrayscaleimage, on the other hand, has just one channel.For the purpose of this post, we will only consider grayscale images, so we will have a single 2d matrix representing an image. The value of each pixelin the matrix will range from 0 to 255 zeroindicating blackand 255 indicating white[5].

In CNN terminology, the 33 matrix is called a filter or kernel or feature detector and the matrix formed by sliding the filter over the image and computing the dot product is called the Convolved Feature or Activation Map or the Feature Map. It is important to note that filters acts as feature detectors from the original input image.

A filter (with red outline) slidesover the inputimage (convolution operation) to produce a feature map. The convolution of another filter (with the green outline), over the same image gives a different feature map as shown in Figure 2. It is important to note that the Convolution operation capturesthe local dependencies in the original image. The two different filters generatedifferent featuremaps from the same original image[6].

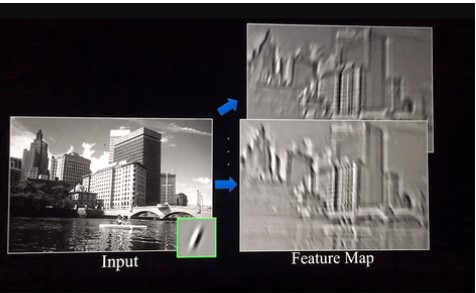


Figure 2: *The Convolution Operation[7]*

The size of the Feature Map (Convolved Feature) is controlled by three parameters that we need to decide before the convolution step is performed:

* Depth:Depthcorresponds to the number of filters weuse for the convolution operation.
* Stride:Stride isthe number of pixelsbywhich we slide ourfilter matrix over the input matrix. When the stride is 1 then we move the filters one pixel at a time. When the stride is 2, then the filters jump 2 pixels at a time as we slide them around. Having a larger stride will produce smaller feature maps.
* Zero-padding:Sometimes, it is convenient to pad the input matrixwith zeros around the border, so that we can apply the filter to bordering elements of our input image matrix. Anice feature of zero padding is that it allows us to control the size of the featuremaps.Adding zero-padding is also calledwide convolution,andnot using zero-padding would be anarrow convolution.
* ReLu[8]: ReLUstands for Rectified Linear Unit and is a non-linear operation.ReLUis an element wise operation (applied per pixel) and replaces all negative pixel values in the feature map by zero. The purpose of ReLU is to introduce non-linearity in our ConvNet, since most of the real-world data we would want our ConvNet to learn would be non-linear.

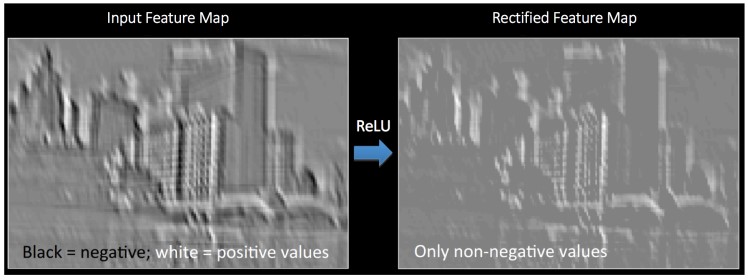


Figure 3: *ReLU operation[8]*

* Fully Connected Layer: The Fully Connected layeris a traditional Multi Layer Perceptron that uses a softmax activation functionin the output layer.The term Fully Connected implies that every neuron inthe previous layer is connected to every neuron on the next layer. The output from the convolutional and pooling layers represent high-level features ofthe input image.The purpose of the Fully Connected layer is to use these features for classifyingthe input image into various classes based on the training dataset. Apart from classification, adding a fully-connected layer is also a cheap way of learning non-linear combinations of these features. Most ofthe features from convolutional and pooling layers may be good for the classification task, but combinations of those features might be even better. The sum of output probabilities from the Fully Connected Layer is 1. This is ensured by using theSoftmaxas the activation function in the output layer of the Fully Connected Layer.The Softmax functiontakes a vector of arbitrary real-valued scoresand squashes it to a vector of values between zero and one that sum to one.

We have a data sets, using our own data base of images and the images from the Cohn-Kanade (CK) database. We will show the results of five facial expression i.e. cheerful, happy, bad, sad, angry. The review is based on these moods. As it help in taking the review of the user as most of the users are not willingly in giving the reviews.

# Literature Review

CNN and ReLU operation are the best to use in the machine learning as these are the more error less classifier and operation then the other. Here is the conclusion of some of the reference paper that we review to make oour project more better and to know more technologir=es that we can use in our system.

* In the paper[9] by Javier G. Razuri, David Sundgren, Rahim Rahmani, Antonio Moran Cardenas, and quot;Automatic emotion recognition through facial expression analysis in merged images based on an Artificial Neural Network,and quot; in 12th Mexican International Conference on Artificial Intelligence, in 2013, they discussed about their system of recognizing humans emotion from a detected humans face. They used eye and mouth region to form a merged new image to identify various facial expressions pertaining to six universal basic facial emotions. They fed this image to the feed forward neural network which included the use of backpropogation to train the model and then classify them accordingly. Paper Full Text
* In the paper [3] by Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman, and quot; Deep Face Recognition,and quot; they explained the use of Convulational Neural Network (CNN) deep learning neural network to get the optimal detection of the persons face. Their algorithm works on the still images or the frames from a video. They used Labeled Faces in the Wild dataset (LFW), which contains 13,332 images of 5,749 different identities, and YouTube Faces (YTF). It contains 3,425 videos of 1,595 people collected from YouTube, with an average of 2 videos per identity, which is a standard benchmark for face verification in video. Paper Deep face recognition
* In the paper [10] by Brandon Amos, Bartosz Ludwiczuk,y Mahadev Satyanarayanan, and quot; OpenFace: A general-purpose face recognition library with mobile applications,and quot; June 2016, They built an OpenFace library that provides near- human accuracy for face recognition, on the LFW benchmark and presented a new classification benchmark for mobile scenarios. This paper provides a light introduction to the deep neural network techniques used like CNN. OpenFace is an python based library that used the Google Facenet architecture. Paper OpenFace: A generalpurpose face recognition library with mobile applications
* In the paper[11] by S.R.Vinotha, R.Arun and T.Arun, and quot;Emotion Recognition from Human Eye Expression,and quot; in International Journal of Research in Computer and Communication Technology, April 2013, they explained the human sentiment recognition based solely on the persons eyes. They their complemented their project in 3 different phases .i.e., edge detection, classification by SVMs(support vector machines) and HMM(hidden markov model). This paper helped us to shape our ideas towards the sentiment recognition based on the human eyes.

# Problem Statement

In todays era, we do the work by seeing its review and on that basis we decide what we should do like, buy a product, watch a movie, etc. So the problem at hand is to get an on hand and effective review on the companys sales products. Peoples are generally lazy to give their review on their purchase as it involves their active participation. Even if we receive the reviews we cant be very much sure to believe them as they are not from trustworthy source. Current review system can also be easily mingled and thus we receive a corrupted review. Also, with the tremendous growth in successful AI implementations in industry, such feedback systems that use AI can be11 considered a better option to increase the companys sales as well as to ease the process.

# Objectives

Here we are providing the reviews on the basis of the human sentiment that he/she is having at the time of purchase of that product. Our system will show the product review in terms of the number of person showing a particular mood on the particular product. The use of melancholy radar will help the company to know about their client more effectively and thus improvise their thinking towards customer satisfaction which will aid their financial growth.

Sub-Objective:

* We use **Convolutional Neural Networks** for the image recognition and classification.
* Using Alexnet architecture having the Convolution, Pooling layers act as Feature Extractors from the input image while Fully Connected layer acts as a classifier[3].
* Cohn-Kanade (CK) database of image for the training set[9].

# Methodology

Convolutional Neural Networks are a category of Neural Networks that have been very effective in areas such as image recognition and classification. The Convolutional Neural Network architecture classifies and used mainly for character recognition tasks. There are four main operations are[12]:

* Convolution
* Non Linearity (ReLU)
* Pooling or Sub Sampling
* Classification (Fully Connected Layer)

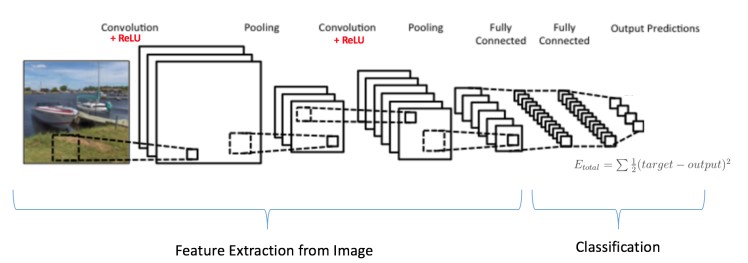


Figure 4: *Training the CNN[4]*

These operations are the basic building blocks of every Convolutional Neural Network. The Convolution + Pooling layers act as Feature Extractors from the input image while Fully Connected layer acts as a classifier.

The overall training process of the Convolution Network[13]:

* Step1: We initialize all filters and parameters / weights with random values.
* Step2: The network takes a training image as input, goes through the forward propagation step (convolution, ReLU and pooling operations along with forward propagation in the Fully Connected layer).
* Step3: Error is calculated at the output layer.
  + Total Error = (target probability output probability)
  + We take the partial derivative of the total error to find the gradient.
* Step4: Use Backpropagation to calculate the gradients of the error with respect to all weights in the network and use gradient descent to update all filter values / weights and parameter values to minimize the output error.
  + The weights are adjusted in proportion to their contribution to the total error.
  + The network has learnt to classify the image correctly by adjusting its weights / filters such that the output error is reduced.[14]
  + Parameters like number of filters, filter sizes, architecture of the network etc. have all been fixed before Step 1 and do not change during training process only the values of the filter matrix and connection weights get updated[14].

The above steps train the CNN, this essentially means that all the weights and parameters of the CNN have now been optimized to correctly classify images from the training set. When a new (unseen) image is input into the system, the network would go through the forward propagation step and output a probability for each class (for a new image, the output probabilities are calculated using the weights which have been optimized). If our training set is large, the network will generalize well to new images and classify them into correct categories.

We have a data sets, using our own data base of images and the images from the Cohn-Kanade (CK) database. We will show the results of five facial expression i.e. cheerful, happy, bad, sad, angry. The review is based on these moods. As it help in taking the review of the user as most of the users are not willingly in giving the reviews. As the image are divided in: the facial emotion detection using merged images in our own data base, the isolated analysis in the mouth and eye zones using our own data base, and finally, the facial emotion detection using merged images from the Cohn-Kanade (CK) database. The use of the five moods for the review works as the scale of five or five stars, but more efficient then any other reviews. As it helps in getting the correct review and helps the company to know more about its client on the basis of the particular location, product, category, etc. The data sets are segrigated on the basis of the moods and their landmarks which classifies the features in am image.

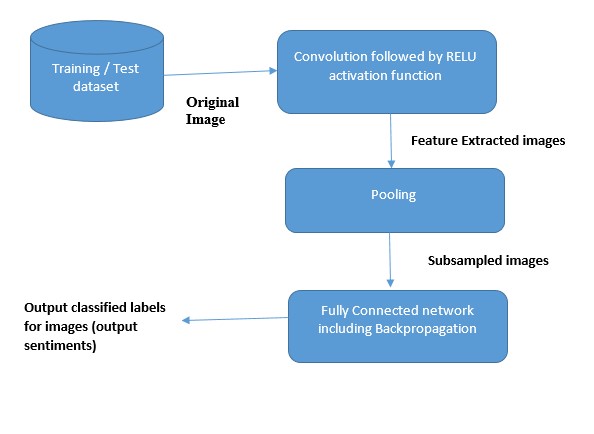


Figure 5: *Data Flow Diagram*

# System Requirements (Software/Hardware)

* Hardware Interface:
  + 64 bits processor architecture supported by windows.
  + Minimum RAM requirement for proper functioning is 4 GB.
  + Required input as well as output devices.
  + Required webcam and sufficient Graphic card for image processing.
* Software Interface:
  + This system is develop in Java programming language.
  + Eclipse compiler

# Schedule (PERT Chart)

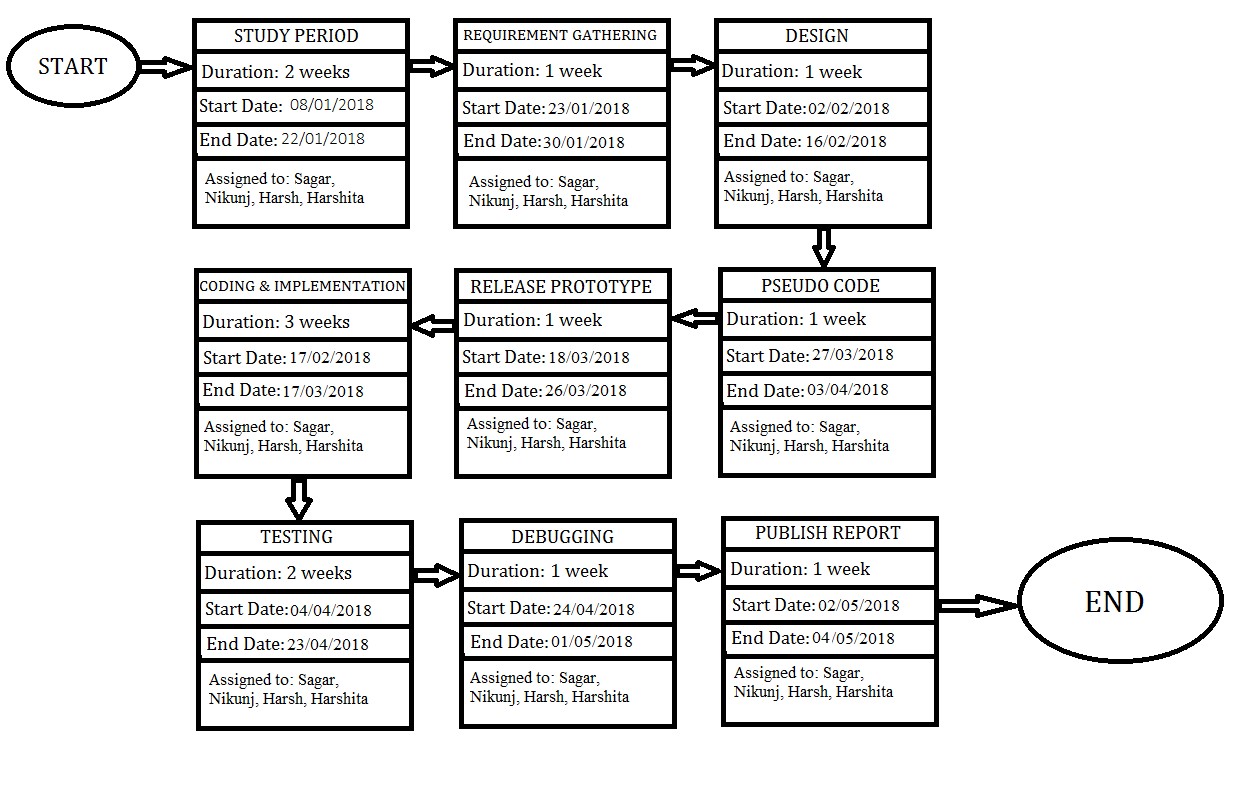


Figure 6: *Pert Chart*

# References

1. https://www.A.W.Harley,œAnInteractiveNode-LinkVisualizationofConvolutionalNeuralNetworks, inISVC,pages867-877,2015(link).Demo//.
2. https://www.cse.unsw.edu.au/~cs9417ml/MLP2/BackPropagation.java//.
3. O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep Face Recognition,” *Procedings of the British Machine Vision Conference 2015*, no. Section 3, pp. 41.1–41.12, 2015.
4. https://www.analyticsvidhya.com/blog/2014/10/ann-work-simplified//.
5. https://www.WhatdothefullyconnectedlayersdoinCNNs?//.
6. https://chatbotsmagazine.com/machine-learning-neural-networks-and-algorithms-5c0711eb8f9a/

/.

1. https://www.NeuralNetworksbyRobFergus,MachineLearningSummerSchool2015//.
2. https://www.CS231nConvolutionalNeuralNetworksforVisualRecognition,Stanford//.
3. J. G. Razuri, D. Sundgren, R. Rahmani, and A. M. Cardenas, “Automatic Emotion Recognition through Facial Expression Analysis in Merged Images Based on an Artificial Neural Network,” *2013 12th Mexican International Conference on Artificial Intelligence*, pp. 85–96, 2013.
4. B. Amos, B. Ludwiczuk, and M. Satyanarayanan, “OpenFace: A general-purpose face recognition library with mobile applications,” *CMU School of Computer Science, Technical Report CMU-CS*, vol. 16, no. 118, 2016.
5. R. Work, “Emotion Recognition from Human Eye Expression,” vol. 2, no. 4, pp. 158–164, 2013.
6. https://wwwConvolutionalNeuralNetworks,AndrewGibiansky//.
7. https://www.ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets///.
8. https://wwwDeepLearningMethodsforVision,CVPR2012Tutorial//.