A Project Report on

FACIAL EMOTION DETECTION USING CONVOLUTIONAL NEURAL NETWORKS BY COMPUTING IMAGE EDGES

submitted to

**JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY ANANTAPUR,**

**ANANTHAPURAMU**

in partial fulfillment of requirements for the award of the degree of

# BACHELOR OF TECHNOLOGY

**in**

# COMPUTER SCIENCE AND ENGINEERING

**Submitted by**

**S HEMA 18121A05K8**

**S HEMANTH 18121A05K9**

**S MADHURI 18121A05L3**

**S RAKESH 18121A05L5**

**S BHARATH REDDY 18121A05L6**

**Under the guidance of**

#### Mr. V. SIVA PRASAD, M.Tech

Assistant Professor of CSE



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



**)**

(Affiliated to JNTUA, Ananthapuramu and Approved by AICTE, Accredited by NBA & NAAC-‘A’ grade)

Sree Sainath Nagar, Tirupati – 517102

**(2018-2022)**

# ABSTRACT

Human emotion recognition plays an important role in the interpersonal relationship. Emotion detection from image is one of the most powerful and challenging research task in social communication. The automatic recognition of emotions has been an active research topic from early eras. Therefore, there are several advances made in this field. Emotions are reflected from speech, hand and gestures of the body and through facial expressions. Hence extracting and understanding of emotion has a high importance of the interaction between human and machine communication.

When it comes to image processing, deep learning (DL)-based emotion recognition outperforms classical methods. This paper proposes a face expression identification approach based on a convolutional neural network (CNN) and image edge detection to avoid the complex process of explicit feature extraction in traditional facial expression recognition. The edge of each layer of the image is retrieved in the convolution process after the facial expression image is normalized. To maintain the texture picture's edge structure information, the retrieved edge information is placed on each feature image. The maximum pooling method is then used to reduce the dimensionality of the extracted implicit features. Finally, a Softmax classifier is used to classify and recognize the expression of the test sample image.

**Keywords:** Convolutional Neural Networks, Edge Computing, Max Pooling, Histogram Equalization, SoftMax Classifier, HAAR - Like Features, KIRSCH Edge Operator.

# INTRODUCTION

Emotion is a mental state associated with the nervous system associated with feelings, perceptions, behavioral reactions, and a degree of gratification or displeasure. One of the current applications of artificial intelligence (AI) using neural networks is the recognition of faces in images and videos for various applications. Most techniques process visual data and search for general pattern present in human faces in images or videos. Face detection can be used for surveillance purposes by law enforcers as well as in crowd management. This paper presents a method for identifying seven emotions such as anger, disgust, neutral, fear, happy, sad, and surprise using facial images.

Human-computer interaction technology refers to a kind of technology which takes computer equipment as the medium, so as to realize the interaction between human and computer. The facial expression recognition, as an important means of intelligent human- computer interaction, has a broad application background. It has been applied in the fields of assistant medicine, distance education, interactive games and public security. Under the trend of artificial intelligence, the communication between human and computer becomes easier and easier. The facial expression recognition is also applied to the medical field. To know the effect of new antidepressants, more accurate drug evaluation can be made according to the daily record of patients’ facial expressions. It helps to interpret the emotions of autistic children and help doctors understand the psychological changes in them.

It can be used in traffic field to judge the fatigue state of pilots or drivers. It is also used in life management robots.

The research of expression recognition in computer field mainly focuses on the feature extraction and feature classification. There are many methods of feature extraction. According to the type of data input, the existing methods of feature extraction can be divided into two categories: one is based on static images and the other is based on a dynamic sequence. Feature extraction methods based on static images include Gabor wavelet transform, Haar wavelet transform, Local Binary Pattern (LBP), and Active Appearance Models (AAM). Generally speaking, the dimension of feature is large before and after the completion of feature, and thus the dimension reduction is usually carried out. Commonly used methods of facial expression classification are Hidden Markov Model (HMM), Support Vector Machine (SVM), AdaBoost, and Artificial Neural Networks (ANN).

To avoid the complex process of explicit feature extraction in traditional facial expression recognition, a facial expression recognition method based on CNN and image edge detection is proposed in this paper. The main innovations of this method are as follows:

1. The edge of each layer of the input image is extracted, and then the extracted edge information is superimposed on each feature image to preserve the edge structure information of the texture image.
2. In this paper, the maximum pooling method is used to reduce the dimension of the extracted implicit features, which shortens the training time of the convolutional neural network model.

# STATEMENT OF THE PROBLEM

The scientific study have been made by many organizations on human sociology identifies the real importance of the human emotion in each and every sector. So if we understand the human emotions automatically it will help us to process more accurately and efficiently. Improving the skills of reading expressions is an important step towards successful relations. Recognizing the facial expressions which are considered universal among all walks of cultures like anger, disgust, neutral, fear, happy, sad, and lastly surprise, by image edge computing unlike traditional methods that work on feature extraction and consume more time. And building a model based on Convolutional Neural Networks (CNN) that helps in detecting and classifying the emotions through weighted assumption of the facial emotions of a person.

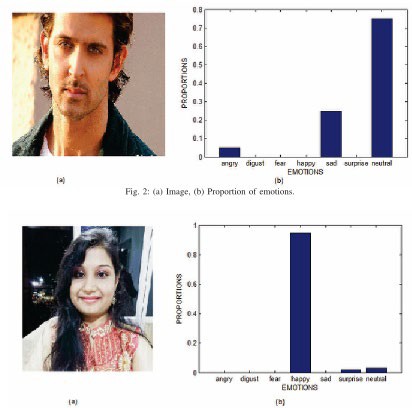


Fig. 3: (a) Image, (b) Proportion of emotions.

# OBJECTIVES

Building a model that uses the deep learning (DL) open library “Keras” for recognizing facial emotions of a person. The model should use Haar Cascade classifier in order to detect the faces and to compute the facial edges and AdaBoost algorithm is to be applied to enhance the classifier. The Eigen values for the Haar- like features are to be calculated using integral graphs. The input images are to be pre-processed first and executed by the model. As part of pre-processing, firstly the images are scaled to a specific size and they are normalized to enhance the contrast or its gray level equalization using Histogram Equalization method. Image edges are to be computing using Kirsch Edge operator.

The main objective is that the model should be able to detect seven different types of emotions like anger, disgust, neutral, fear, happy, sad, and lastly surprise. So the model is to be trained on the basis of FER-2013 dataset that is split into training data and test data. The image data is to be sent through various layers of convolutional neural networks. Besides working with convolutional layers, the data should be sent through maximum pooling layer for dimensionality reduction simultaneously applying ReLU function to avoid the problem of overfitting. The resulting data need to be sent through full connection layer and Softmax layer for final classification of emotions.

### SOFTWARE AND HARDWARE REQUIREMENTS

**HARDWARE SPECIFICATIONS:**

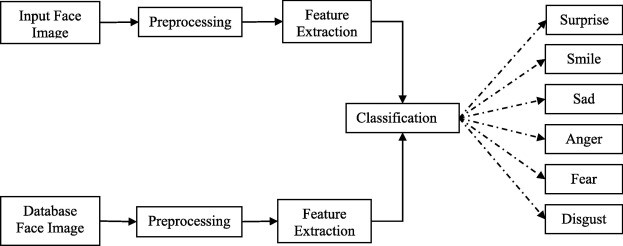
* Processor : i5/Intel Processor
* RAM : 8GB (Min)
* Hard Disk : 128 GB

### SOFTWARE SPECIFICATIONS:

* + Operating System : Windows 10
  + Server-side Script : Python 3.6
  + IDE : Jupyter notebook, Google Colab
  + Libraries Used : Numpy, Flask, keras, pandas, FER(Face Emotion Recognizer)

# EXISTING SYSTEM

The traditional method for emotion detection includes three steps. They are preprocessing, feature extraction and classification.



#### Preprocessing

Image preprocessing includes different types of processes such as image clarity and scaling, contrast adjustment, and additional enhancement process to improve the expression frames.

The cropping and scaling processes were performed on the face image in which the nose of the face is taken as midpoint and the other important facial components are included physically. The [Gaussian filter](https://www.sciencedirect.com/topics/computer-science/gaussian-filter) is used for resizing the input images which provides the smoothness to the image.

Normalization is the preprocessing method which can be designed for reduction of illumination and variations of the face images with the [median filter](https://www.sciencedirect.com/topics/computer-science/median-filter) and to achieve an improved face image. It is also used for the extraction of eye positions which make more robust to personality differences for the FER(Facial Emotion Recognition) system and it provides more clarity to the input images. Localization is a preprocessing method and it uses the Viola-Jones algorithm to detect the facial images from the input image. Detection of size and location of the face images using

Adaboost learning algorithm and haar like features. It is mainly used for spotting the size and locations of the face from the image.

The [histogram equalization](https://www.sciencedirect.com/topics/computer-science/histogram-equalization) method is used to conquer the [illumination variations](https://www.sciencedirect.com/topics/computer-science/illumination-variation). This method is mainly used for enhancing the contrast of the face images and for exact lighting also used to improve the distinction between the intensities.

#### Feature extraction

Feature extraction process is the next stage of FER system. Feature extraction is finding and depicting of positive features of concern within an image for further processing. In [image processing](https://www.sciencedirect.com/topics/computer-science/image-processing) [computer vision](https://www.sciencedirect.com/topics/computer-science/computer-vision) feature extraction is a significant stage, whereas it spots the move from graphic to implicit data depiction. Then these data depiction can be used as an input to the classification. The feature extraction methods are categorized into five types such as texture feature-based method, edge based method, global and local feature-based method, geometric feature-based method and patch-based method. The traditional methods are mostly based on texture feature and global and local features.

An **image texture** is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image. Some of the filters used for texture synthesis are as follows. Gabor filter is a [texture descriptor](https://www.sciencedirect.com/topics/computer-science/texture-descriptor) for feature extraction and it includes the magnitude and phase information. The Gabor filter with the magnitude feature confines the information about the organization of the face image. The phase feature precincts the information about the complete description of the magnitude features. Local Binary Pattern (LBP) is also a texture descriptor and it can be used for feature extraction. Generally LBP features are produced with the binary code and it can be obtained by using thresholding between the center pixel and its locality pixels. The facial texture features are extracted using the Gaussian Laguerre (GL) function which grants a steering pyramidal structure which extracts the texture features and the facial related occurrence information. Comparing to [Gabor function](https://www.sciencedirect.com/topics/computer-science/gabor-function) GL uses the single filter

instead of multiple filters. Moreover, another descriptor which is used namely Vertical Time Backward (VTB) which also extracts the texture features of face images. Moments descriptor extracts the shape related features of significant facial components. Both VTB and moments descriptors are effective on spatiotemporal planes. Weber [Local Descriptor](https://www.sciencedirect.com/topics/computer-science/local-descriptor) (WLD) is a feature extraction technique that extracts the high discriminant texture features from the segmented face images. Feature extraction is performed with three stages using Supervised Descent Method (SDM). At first, the facial main positions are extracted. Next the related positions are selected. Finally it estimates the distance between the various components of the face.

The descriptors which extract the features based on the global and local feature-based methods are described as follows. Principal Component Analysis (PCA) method is used for feature extraction. It extracts the global and low dimensional features. Independent Component Analysis (ICA) is also a feature extraction method which extracts the local features using the multichannel observations. Stepwise [Linear](https://www.sciencedirect.com/topics/computer-science/linear-discriminant-analysis) [Discriminant Analysis](https://www.sciencedirect.com/topics/computer-science/linear-discriminant-analysis) (SWLDA) is the feature extraction technique which extracts the localized features with backward and forward regression models. Depends on the class labels the F-test values are estimated for both regression models.

#### Classification

Classification is the final stage of FER system in which the [classifier](https://www.sciencedirect.com/topics/computer-science/classification-machine-learning) categorizes the expression such as smile, sad, surprise, anger, fear, disgust and neutral.

[Support Vector Machine](https://www.sciencedirect.com/topics/computer-science/support-vector-machine) (SVM) is one of the [classification techniques](https://www.sciencedirect.com/topics/computer-science/classification-technique) in which two types of approaches are involved. They are one against one and one against all approaches. One against all classification means it constructs one sample for each class. One against one classification means it constructs one class for each pair of classes and SVM is one of the strongest classification methods for advanced dimensionality troubles. SVM is the supervised [machine learning technique](https://www.sciencedirect.com/topics/computer-science/machine-learning-technique) and it uses four types of kernels for its better performance. They are linear, polynomial, [Radial Basis Function](https://www.sciencedirect.com/topics/computer-science/radial-basis-function) (RBF) and sigmoid. The linear kernel maps the [high dimensional data](https://www.sciencedirect.com/topics/computer-science/high-dimensional-data) and it is linearly separable. The RBF kernel uses the

function that maps the single feature into the high dimensional data. The polynomial kernel learns the [nonlinear models](https://www.sciencedirect.com/topics/computer-science/nonlinear-model) and also resolves their similarity.

The Hidden Markov Model (HMM) classifier is the statistical model which categorizes the expressions into different types. Hidden [Conditional Random](https://www.sciencedirect.com/topics/computer-science/conditional-random-field) [Fields](https://www.sciencedirect.com/topics/computer-science/conditional-random-field) (HCRF) representation is used for classification. It uses the full covariance Gaussian distribution for superior classification performance.

The Multilayer [Feed Forward Neural Network](https://www.sciencedirect.com/topics/computer-science/feedforward-neural-network) (MFFNN) classifier uses three layers such as input, hidden and output layers and [back propagation algorithm](https://www.sciencedirect.com/topics/computer-science/backpropagation-algorithm) for classification. In the training stage the weights are initialized and the activation units are estimates. Bayesian [neural network](https://www.sciencedirect.com/topics/computer-science/neural-networks) classifier is the classification method which also includes three layers such as input, hidden and output layers. The classical back propagation algorithm is used with [Bayesian classifier](https://www.sciencedirect.com/topics/computer-science/bayesian-classifier) for its better accuracy. Convolution Neural Network (CNN) consists of two layers such as [convolutional](https://www.sciencedirect.com/topics/computer-science/convolutional-layer) [layer](https://www.sciencedirect.com/topics/computer-science/convolutional-layer) and subsampling layer in which the two dimensional images are taken as input. In convolutional layer the feature maps are produced by intricate the [convolution](https://www.sciencedirect.com/topics/computer-science/convolution-kernel) [kernels](https://www.sciencedirect.com/topics/computer-science/convolution-kernel) with the two dimensional images where as in the subsampling layer, pooling and redeployment are performance. The CNN also contains two important perceptions likely shared weight and sparse connectivity. In FER, the CNN classifier used as multiple classifiers for the different face regions. If CNN is framed for entire face image then first frame the CNN for mouth area and next for eye area likely for each other area CNNs are framed.

According to several classifiers SVM classifier gives better recognition accuracy and it provides better classification. The neural network-based classifier CNN gives better accuracy than the other neural network based classifiers. In FER, SVM classifier is more exploitable comparing with other classifiers for recognition of expressions.

## System Requirements Specification

#### Contents

* 1. **Introduction**
  2. Purpose
  3. Project scope

#### 2.0 General Description

* 1. **Functional Requirements**
  2. Jupyter Notebook / Google Colab
  3. Python Programming Language
  4. Webpage
  5. Database

#### Non-Functional Requirements

* 1. Usability
  2. Availability
  3. Reliability
  4. Robustness

#### System Specifications

* 1. Hardware specifications
  2. Software specifications

## Introduction

#### Purpose

The purpose of this document is to present a detailed description of the Facial emotion recognition using Convolutional Neural Networks. In this project a new approach is described which classifies the expressions with the help of image edge computing. Besides classifying facial emotions, this model requires lesser training time with more accuracy and robustness unlike traditional methods that uses feature extraction.

#### Scope of the Project

The model designed will be used for recognizing seven types of facial expressions that are considered as universal among all walks of cultures. They are anger, disgust, neutral, fear, happy, sad, and lastly surprise. Facial expression recognition systems run with an objective of understanding and analyzing the true perceptions of people with the help of their expressions.

Another important aspect to be noted is that this model is more efficient and accurate when compared to the traditional methods because it is well trained and tested on FER-2013 dataset and uses maximum pooling to reduce the training time. This model can be used by any person who is interested in FER (Facial Expression Recognition) systems.

## 2.0 General description

The main theme of the project is to detect and classify the emotions of a person. Facial expression emotion recognition is an intuitive reflection of a person’s mental state, which contains rich emotional information, and is one of the most important forms of interpersonal communication. In this project a model is build to which facial images are given as input. The file name and its path of these images are taken by the source code to access the images in that address. This is done using a Webpage in which user upload the image. Once the execution is completed the bar charts are displayed as output that describes the classification of emotions.

## Functional Requirements

#### Jupyter Notebook / Google Colab

The source code for this model needs to be executed in either Jupyter Notebook or Google Colab. Jupyter Notebook provides a simple platform for executing python programs. It needs some of the disk memory to save python files. It also provides various tools that can help to import the python files from local disk and execute them whenever necessary.

Colaboratory, or “Colab” for short, is a product from Google Research. Colab allows anybody to write and execute arbitrary python code through the browser, and is especially well suited to machine learning, data analysis and education. The main advantage of Google Colab is that it provides free GPU (Graphics Processing Unit) and some other additional tools.

#### Python Programming language

We use python to design the model that works on Convolutional Neural Networks. This program consists of various python libraries like “Keras”, “matplotlib”, “tensorflow” etc,.

#### Webpage

This is where the images are uploaded and taken as input. The webpage consists of an “Upload” button through which the images are stored into the database. The uploaded image is analyzed and facial expressions present in it are classified and displayed as a bar chart.

#### Database

To store the image data some of the local disk memory is provided. The webpage is linked to this database. So the images uploaded in the webpage are directly sent to this database for storing.

## Nonfunctional requirements

#### Usability

This model can be used by anyone through the designed webpage. The page is very simple to use for anyone who knows a little English vocabulary. The user does not need to understand any of the concepts in the source code. They just need to upload their facial images to know the weighted comparison of different expressions present in their faces.

#### Availability

This project can be used by anyone anytime without any constraints or restrictions. The webpage can be accessed by anybody who is willing to understand and classify various expressions present in the facial images.

#### Reliability

This model is quite efficient when compared to the traditional methods with less accuracy rates around 65%. The proposed model runs with an average accuracy rate of around 80%. Even training time is less when compared to other methods. The main advantage of this model lies in its performance and accuracy.

#### Robustness

This model is efficient even when the facial images are fuzzy. It provides accurate results and detects faces clearly in various types of backgrounds. To increase the robustness the model is trained with the scientific mixture of two datasets ie, Fer- 2013 and LFW. And to know the standards of this model its reliability factor is compared to algorithms like FRR-CNN and R-CNN.

## System specifications

#### Hardware specifications

* + - Network connectivity.
    - Minimum 4GB RAM is required.
    - i5/ Intel Processor
    - Graphic Card is required.

#### Software specifications

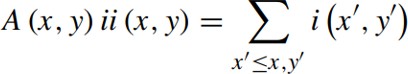
* Operating System : Windows 10
* Server-side Script : Python 3.6
* IDE : Jupyter notebook, Google Colab
* Libraries Used : Numpy, Flask, keras, pandas, FER(Face Emotion Recognizer)

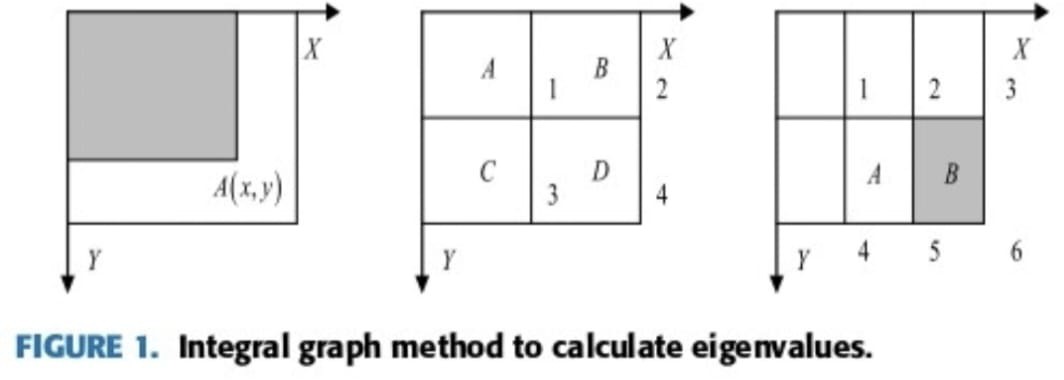
# PROPOSED SYSTEM

The difference between the proposed system and the existing system is that the concept of image edge computing is included in this model. The facial emotions are detected and classified using these edge information. Firstly, we locate the face in the image and cut out the face image. Then, we normalize the face image to a specific size. Next, we equalize the histogram of the image to reduce the influence of illumination and other factors. Finally, we extract the edge of each layer of the image in the convolution process.

### FACE DETECTION AND LOCATION

This paper uses a Haar classifier for human detection. The Haar classifier is trained by Haar-like small features (linear, edge, center and diagonal) and an integral graph method combined with the AdaBoost algorithm. This method uses the Haar-like to extract facial features, and uses an integral graph to realize fast calculation of Haar- like features. However, the calculation of eigenvalues is very time-consuming. In order to improve the calculation speed, this paper uses the integral graph method to calculate the Haar like eigenvalues. The integral graph of the coordinate A (x, y) in a graph is defined as sum of all the pixels in its upper left corner.

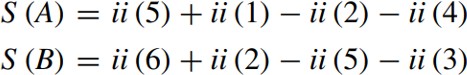




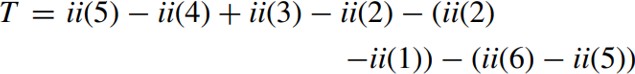
Here, *ii (x,y)* represents the integral image. *i (x`,y`)* represents the original image; for gray image, here represents the gray value and for color image, here represents the color value. The pixel value of an area can be calculated by using the integral graph of the end points of the area, as shown in above image. The pixel value of region *D* can be calculated by



where *ii* (1) represents the pixel value of region *A*, *ii* (2) represents the pixel value of region *A* + *B*, *ii* (3) represents the pixel value of region *A* + *C*, *ii* (4) represents the pixel value of regions *A* + *B* + *C* + *D*. The eigen values of rectangular features can be calculated by integral graphs of feature endpoints. Taking the edge features as an example, the eigen value calculation can be expressed by Fig 1 (c). The pixel values of point A and point B are:



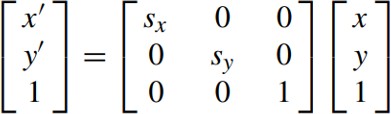
According to the definition, the eigenvalue of rectangular feature is the pixel value of region *A* minus the pixel value of region *B*. According to formula (3) and formula (4), the formula for calculating eigenvalue is as follows.

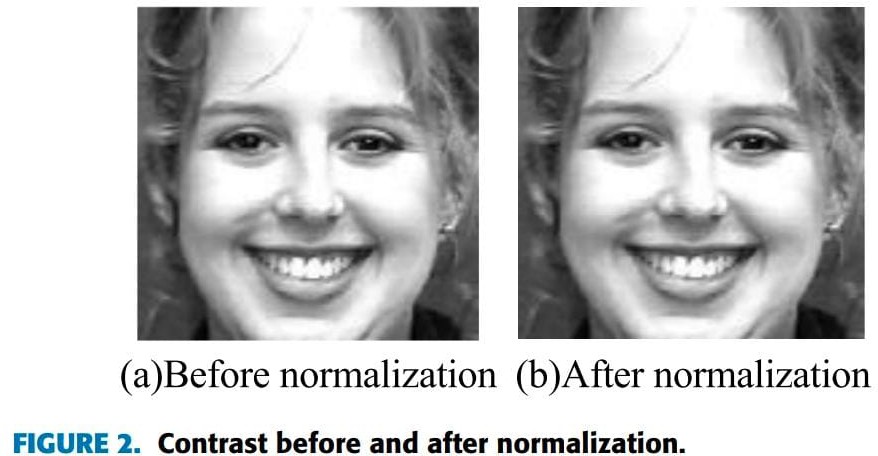


The extracted Haar-like features are used to train the classifier, and the AdaBoost algorithm is used to train the classifier. Finally, the trained classifier is used to extract the face from the image.

### SCALE NORMALIZATION

Because the input of the network is a fixed sized picture, the original picture should be normalized to generate a specific size picture. Through normalization, the input image is scaled to 128 X 128 size. Let point (*x*, *y*) in the original picture be normalized and mapped to point *(x*`, *y*`). The mapping is as follows:





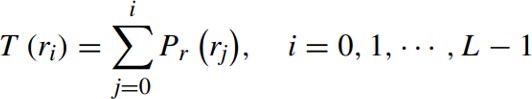
### GRAY LEVEL EQUALIZATION

Uneven distribution of light and shade, which will increase the difficulty of feature extraction. Histogram Equalization (HE) method is used to average the gray level to enhance the contrast. It can be simply done by transforming the histogram of the original graph into a uniform distribution form.

If the gray level of the gray image is *L*, the size is *M* x *N*, and the number of pixels in the *ri* gray level is E, the corresponding probability of gray level occurrence is as follows:

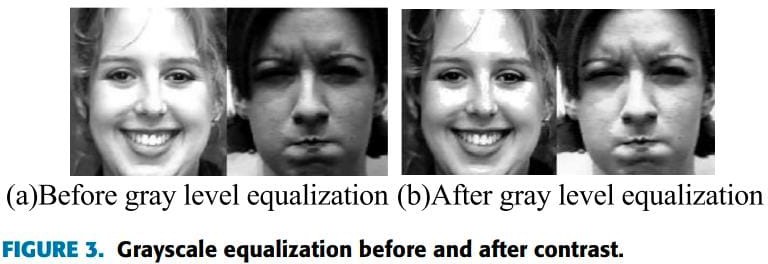


Subsequently, the cumulative distribution function is calculated using the following equation.



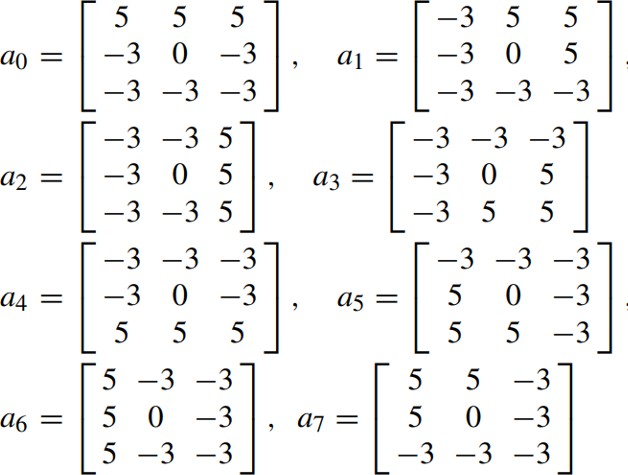
Finally, the image histogram is averaged using the following mapping relations:





### IMAGE EDGE DETECTION

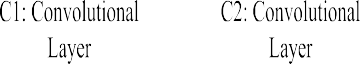
The edge information of an image is often reflected in the area where the gradient information of the image changes dramatically. The edge of the image gives people a stronger visual sense. Kirsch edge operator is used to extract image edge information whose templates are in eight directions.



# APPROACH

### FACE EXPRESSION RECOGNITION NETWORK MODEL BASED ON CNN

The CNN is a feedforward neural network, which can extract features from a two- dimensional image and optimize network parameters by using back propagation algorithm.



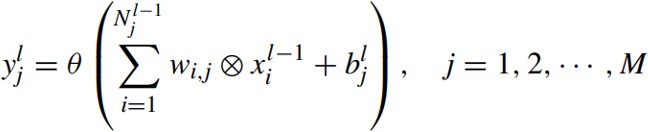


CNN structure for facial expression recognition.

Excluding input layer, the network consists of seven layers, including three convolution layers (C1, C2 and C3), two pooling layers (S1 and S2), one full connection layer and one Softmax layer. The input layer is a 96x96 face pixel matrix. The CNN has three important characteristics: local perception, weight sharing and down sampling. Common sense holds that people’s perception of the outside world is generally from the local to the whole. The adjacent pixels are closely related, while the distant pixels have little correlation. There- fore, neurons only need to perceive the local pixels, and then integrate the local information at the bottom to get the global information at the high level.

### CONVOLUTION LAYER

The convolution layer is the core of CNN, which has the characteristics of local connection and value sharing. The input image and several trainable convolution filters are convoluted to produce the Cl layer of the feature mapping layer. Then the feature mapping map is processed, including summation, weighting and bias operations. Generally, the computational expression of convolution layer is



Finally, these feature maps are rasterized and connected into a set of feature vectors, which are then transferred to the neural network classifier for training. Practice has proved that the network trained with ReLU activation function has moderate sparsity. At the same time, it can solve the problem of gradient disappearance.



### POOLING LAYER

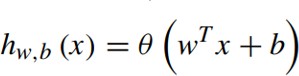
The main purpose of the pooling operation is to reduce the dimension. If we train the Softmax classifier directly with all the features we have learned, it will inevitably bring about the problem of dimension disaster. So pooling layer is used. A pooling window of 2 x 2 step size can reduce the dimension of the next feature map by half which greatly improves the training speed. The general expression of the pooling is



### FULL CONNECTION LAYER

The input of the full connection layer must be a one-dimensional array, whereas the output of the previous pooling layer S2 is a two-dimensional array. First, the two- dimensional array corresponding to each feature graph is converted into a one- dimensional array using Flatten function and then 128 one- dimensional arrays are connected in series to a feature vector of 51200 dimensions (20 x 20 x 128 = 51200)

as the full connection.



### SOFTMAX LAYER

The last layer of the CNN uses a Softmax classifier. The Softmax classifier is a multi-output competitive classifier. When a given sample is input, each neuron outputs a value between 0 and 1, which represents the probability that the input sample belongs to that class. Therefore, the category corresponding to the neuron with the largest output value is selected as the classification result.

### CNN PARAMETER TRAINING:

The training process of CNN is essentially the process of optimizing and updating network weights. Appropriate initialization of weights has a great impact on the updating of weights. The commonly used initialization methods include constant initialization, uniform distribution initialization and Gauss distribution initialization.

The training of convolution neural network is divided into two stages:

#### Forward Propagation:

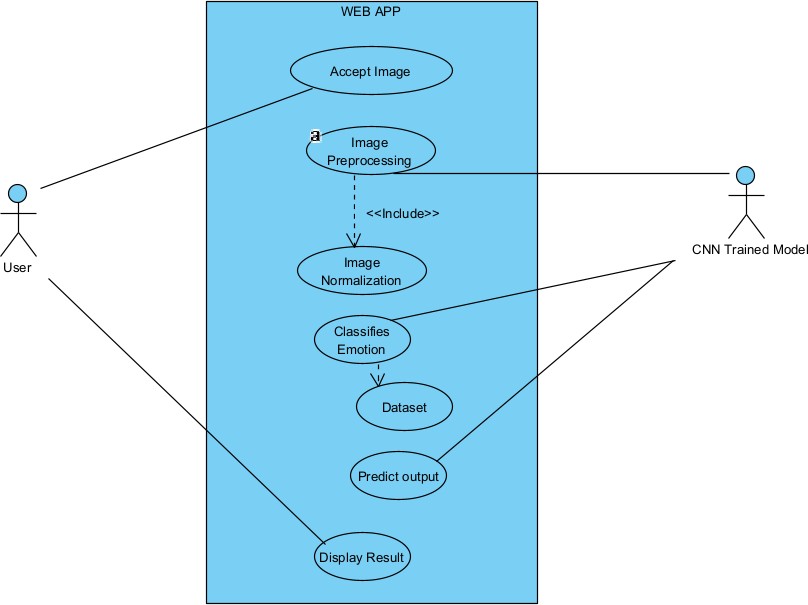
Sample x is extracted from the training sample set. Its corresponding category label is y, y is a 7-dimensional vector whose elements represent the probability that x is divided into different categories. The output of the upper layer is the input of the current layer. Then, the output of the current layer is calculated by activation function, which is passed down layer by layer.

#### Back Propagation stage:

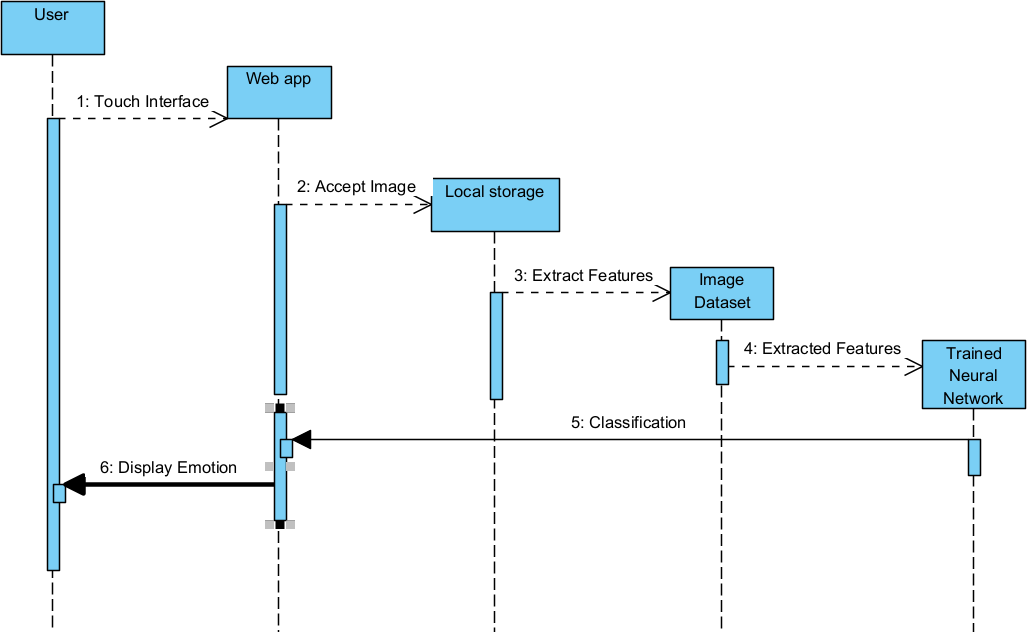
Calculate the error between the output y˜ of Softmax layer and the class label vector y of a given sample , and adjust the weight parameters by minimizing the mean square error cost function.

# UML DIAGRAMS

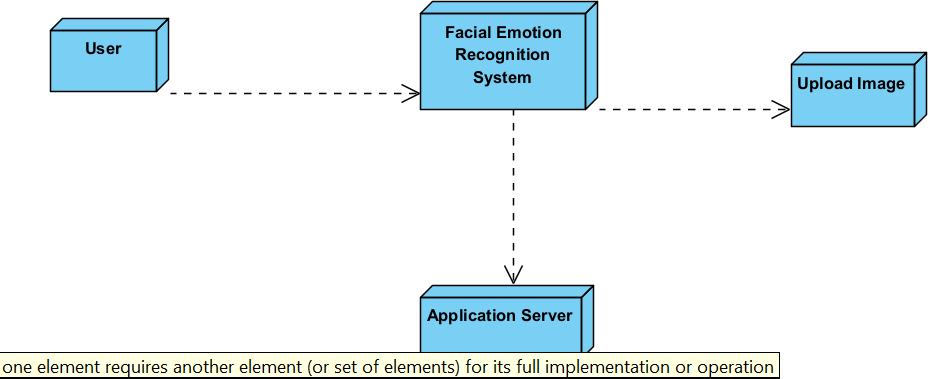
**Use case Diagram**



**Sequence Diagram**



**Deployment Diagram**



# REFERENCES

1. [https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8884205](https://ieeexplore.ieee.org/stamp/stamp.jsp?tp&arnumber=8884205)
2. R. M. Mehmood, R. Du, and H. J. Lee, ‘‘Optimal feature selection and deep learning ensembles method for emotion recognition from human brain EEG sensors,’’ IEEE Access, vol. 5, pp. 14797–14806, 2017.
3. T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang, and Z. Cui, ‘‘MPED: A multi- modal physiological emotion database for discrete emotion recognition,’’ IEEE Access, vol. 7, pp. 12177–12191, 2019.
4. E. Batbaatar, M. Li, and K. H. Ryu, ‘‘Semantic-emotion neural network for emotion recognition from text,’’ IEEE Access, vol. 7, pp. 111866–111878, 2019.
5. Y. Zhang, L. Yan, B. Xie, X. Li, and J. Zhu, ‘‘Pupil localization algorithm combining convex area voting and model constraint,’’ Pattern Recognit. Image Anal., vol. 27, no. 4, pp. 846–854, 2017.
6. H. Meng, N. Bianchi-Berthouze, Y. Deng, J. Cheng, and J. P. Cosmas, ‘‘Time- delay neural network for continuous emotional dimension prediction from facial expression sequences,’’ IEEE Trans. Cybern., vol. 46, no. 4, pp. 916–929, Apr. 2016.

**Signature of Project Supervisor**