Convolution Neural Network Based Emotion Classification Cognitive ModelforFacial Expression

Dr. TanujaPatgar[1], Triveni[2]

[1]Assistant Professor, Dept of ECE, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka, E-mail-tanujaharish13@gmail.com

[2] Assistant Professor, Dept of ECE, Dr. Ambedkar Institute of Technology, Bangalore, Karnataka,E-mail- sg3veni@gmail.com

Abstract

Facial expression is a structured communicative approach in building relationships and interacting with others. It can be easy to focus on sensitivity and emotional content of mental state, personality, behavioral and intention of persons.The human behavior model makes enlighten on automatic facial expression recognition system.In Human-Machine Interaction (HMI), recognition of facial expressions is automated and it is considered as important component of natural communication. The paper proposes Convolutional Neural Networks(CNN) based emotion classification cognitive model for facial expression.The model classifiespositive and negative images which significantly specify regions within an image and network performance is depend on different training options. A rectangular box is drawn around the facial image and output is formatted above the rectangular box. Kaggle facial expression FER-2013 Databasewith seven facial expression labels as happy, neutral, surprise, fear, anger, disgust, and sad is implemented. The evaluation of model shows that accuracy of lab condition testing data set is comparing with proposed model, the highest accuracy for happy emotion with 99%, followed by surprise with 98%, neutral with 96% and least accuracy for fear emotion with 45%. Live validity test is obtained with a webcam resolution of 320x240 and the network input layer is 224x224 with 50 cm distance is maintained between the webcam and face.

Index Terms- Cognitive model, emotional intelligence,Haar classifier, Pooling

1. Introduction

Humans can exchange their emotions either through speech or body gestures.One of the important parts of communication where humans shows their emotions through facial expressions.Though nonverbal communication, emotions are expressed in terms of facial feelings. Facial expressions convey nonverbal cues and they play an important role in interpersonal relations. In Human-Machine Interaction (HMI), recognition of facial expressions is automated and it is considered as important component of natural communication. There are number of researches are carried out for humans recognize facial expressions recognition by machine is still challenge. The advancement in application area such as face detection, feature extraction mechanisms and methodology used for expression classification still under accomplish.

The key feature in human communication system is analyzed using facial expressions and body language. In 19th century, Charles Darwin published globally some facial expressions images. Later, it becomes valuable data set for facial expressions that play an important role in non-verbal communication. In 1971, Ekman and Friesen declared that facial expressions are associated with particular emotions. Even animals also develop similar muscular movements (emotions) related to certain mental behavior state. Hence, if properly modelling of this universality can be very major feature in HMI, well-trained system can understand emotions independent of category.It is understood that facial expressions are not necessarily directly link to emotions or vice versa. Facial expression is additional functioning of mental sequence of events while emotions are also expressed through body language and voice.

Facial Emotion Recognition (FER) system making more attentions in the CNN based research area. Emotions are classified from facial expression images using filter banks and Deep CNN.It gives high recognition accuracy rate.FER can be also performed using image spectrograms with deep convolutional networks.The research proposed an approach based on Convolutional Neural Networks (CNN) for facial expression recognition. The input data is image and CNN are used to predict facial expression and label them whether facial expression is related to anger, happiness, fear, sadness, disgust and neutral. In neural networksConv-Nets is widely used as classifier in many real time applications such as images recognition, images classifications, Objects detections, face recognition etc.

Generally deep learning CNN models are trained with pre-defined data set. Each input image will pass through series of convolution layers with filters (Kernals), Pooling, fully connected layers (FC).Apply Soft max function to classify an object with probabilistic values between 0 and 1.When input images are large then pooling layers, technique is used to reduce number of parameters. Most of neural network prefers Spatial pooling technique. The technique which retains original information but reduces the dimensionality of each input map. It is also called subsampling or down sampling.

Visual features of an image are examined and some of the classifier techniques are discussed which is helpful in further inspection methods of emotion recognition. The predictions of future reactions from images based on recognition of emotions using different classes of classifiers are gaining the much more attention. The advanced classification algorithms such as Random Forest, K-Nearest Neighbor are adopted to classify emotions based on facial expression. In neural network,Deep RNN like Bi-directional LSTM and LSTM are modeled and used for audio visual features which arises tremendous attempts to solve real time problems in data science also.

2. Related Work

The growth of available computational power on consumer computers in the beginning of 21th century gave boost to development of algorithms used for interpreting pictures. In the field of image classification, two approaches make attempt to improve feature extraction concept.The

pre-programmed feature extraction and self-learning neural network are object classification algorithm. Pre-programmed feature extractors are used to analytically break down several elements in picture and self-learning neural networks where system itself develops rules for object classification by training upon labeled sample data.

At beginning of 21st century, Fasel and Luettin made an extensive research on analytical feature extractors and neural network approaches using facial expression recognition system. It is concluded that both said algorithm work approximately equally well. However, the performance of neural network-based models becomes significantly improved with availability of training data and computational power. Some recent achievements arelisted below.

The automated image classifier for human visual cortex using Deep neural network methodology is explained clearly in publication [1]. CIFAR-10 dataset consists of self-developed labeled collection of 60000 images in 10 classes. The model is designed to classify objects from received images.The visual of proposed model segregating network filter is another important outcome of the research.

Another research is proposed CIFAR-10 dataset where network is configured with 4 convolution layers with 300 maps each, 3 max pooling layers, and 3 fully connected output layers[2]. Deep network architecture is implemented with GPU support to decrease training time. Prior state-of-art results improved significantly by using datasets such as MNIST handwritten digits, Chinese characters, and CIFAR-10 images.Human emotions performance is achievedextremely low error rates where GPU used training time for many days.

In 2010, another research proposed Image Net LSVRC-2010 data set. The network is configured with 5 convolution layer, 3 max pooling, and 3 fully connected layers is trained with 1.2 million high resolution images [3]. The introduction of yearly Imagenet challenge boosted research on image classification and data set of labeled data is used in public domain.This techniquesis reducedover fitting, low network size, minimum number of layers and better performance are promising when compared to previous works.

Japanese Female Facial Expression (JAFFE) and Extended Cohn-Kanade (CK+) databases are used for facial expression recognition in [4].The algorithm is implemented using deep neural network. The most notable feature of network is hierarchical face parsing concept whereinput image is passed through network many times.The system will detect face first then eyes, nose, mouth, and finally all emotion. Using other methods like Support Vector Machine (SVM) and Learning Vector Quantization (LVQ) the results are compared with accuracy obtained by other methods on same database.

Gabor filtering for image processing and SVM for classification based on Cohn-Kanade database is explained in paper [5].The gabor filter is mainly suitable for pattern recognition in images and is claimed to mimic the function of human visual system. The emotion recognition accuracies are

for anger 88% andfor surprise 100%. But disadvantage of this approach where precise pre-processing of data is required before feeding it into classifier.

One of the most recent work proposed neural network based on Facial Expression Recognition Challenge (FERC-2013) data set. The deep network is implemented with 3 convolutional layers, 3 max pooling layers and 1 fully connected layer [6]. The result of emotion classification shows average accuracy of 67%. It describes neural network able to recognize race, age, gender, and emotion from pictures of many faces. By surveying with many literatures, the most promising concept for facial expression analysis is use of deep convolutional neural networks. Hence analysis is made based on previous state of art in adjusting the network size, pooling, and dropout.

The organization of the paper is as follows. Section (3) explains the 7 emotion such as happy, neutral, surprise, sad, fear, anger and disgust classification based on facial expression are proposed.The architecture of emotion recognition Cognitive model which is used for statistical features and CNN features classification is depicted in section (4). Section (4.1) describes Haarcascade-basedface detection technique. The multi modularity of convolutional layers supporting for feature extraction is explained in section (4.2).The explanation in section (4.3) is about design constraints of pooling layer.Data flow analytical diagram for emotional classification methodology is describes in section (5). Experimentation and result analysis is depicts in section (6) and the subtitles section (6.1)explain design constraint of facial expression recognition-2013 database experiment and section (6.2) briefs about training constraint of neural network model. The real time validation test is explained in section (7).Finally, section (8) depicts concluding remarks.

3. Wheel of Seven Emotion Recognition Model

The seven-emotion such as happy, neutral, surprise, sad, fear, anger and disgust classification based on facial expression are proposed and is depicted in fig.3.1. The emotionsare classified based on feeling state withsequence of on-going events are simulated would trigger feelings which results in actions. People to people emotions are exchanged in day-to-day interactions. Emotions are reflected in many forms such as face, voice, hand and body gestures. The face expression is most basic form of non-verbal communication. Understanding emotions and knowing how to interact with people's expressions greatly enriches the interaction.
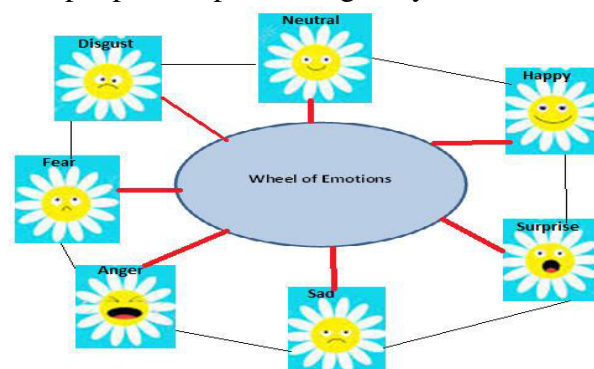

Fig.3.1Wheel of Seven Emotion Recognition Model

By knowing the user emotion, system can adapt to the user emotions. Sensing to user's emotional culture will be perceived as more natural, persuasive, and trusting. All impression of other people is highly dependent on their expression.Challenge is to classify an unknown expression into one of seven classified emotions. A base of affective computing is recognition of human expression.Goal is to introduce natural ways of communication in person-to-machine interaction.

The facial expressions are culturally variable. The vision is first step in the image recognition process.Once input image is captured, next step is pre-processing to reduce noise and improves contrast.Features are extracted and areas of interest are detected. Finally, high-level processing used to capture motion of objects due to relative motion between object and observer.The model should identify face and successfully classifies seven emotions such as Neutral, Happy, Surprise, Sad, Angry, Disgust and Fear respectively. The trained CNN is verified by pooled region of concerned image of face. A rectangular box is drawn around the facial image and output is formatted above the rectangular box. By focusing on interested area on the face it will capture facial expressions as shown in fig.3.2.
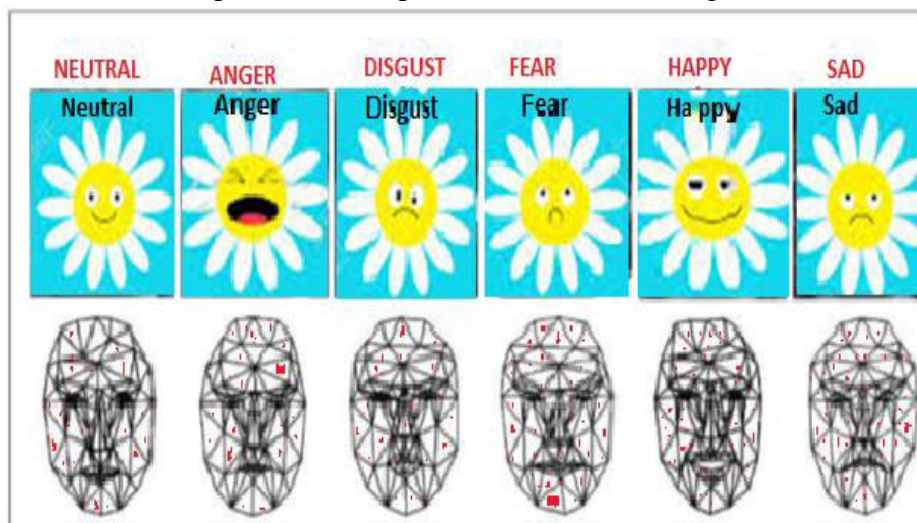
Fig. 3.2 Mapping Pattern of facial expression with neural network

**4. Proposed Emotional Intelligence Cognitive Model Trait Frame Work**

The frame work of emotion recognition cognitive model which is used for statistical features and CNN features is proposed in fig.4.1The proposed frame work uses camera to stream video and capture frames. It acquires a video stream from webcam with HD resolution (1920 × 1080, 25 fps). From every frame, positive and negative images are detected and classified for further processing. It largely divided into Cascade classifier, Convolution layers, Max pooling layers, fully connected layers and softmax classifier. The proposed model consisting of 2*2 convolution layers, 2*2 pooling layers, 300 nodes fully connected layers and SoftMax regression layer.
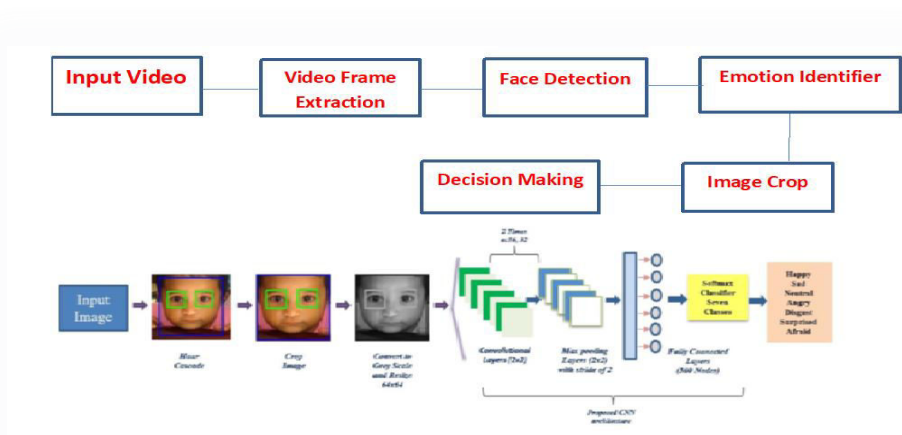
Fig 4.1 Proposed Emotional Intelligence Trait Frame Work

4.1 HaarCascade Based Face Detection Technique

The machine learning based object detection algorithm is to identify objects location in video or image.From bundle of positive and negative images the cascade function is trained. The algorithm has defined with on-going four levels namely Haar Feature Selection (HFS), Creating Integral Images (CII), Ada Boost Training (ABT) and Cascading Classifiers(CC).To train the classifier initially algorithm required a greater number of faces with positive images and negative images without faces.Later extract interested features from it. Haar feature selects adjacent rectangular regions at interested location in detection window, add all pixel intensities in every region and calculates the difference between these sums.
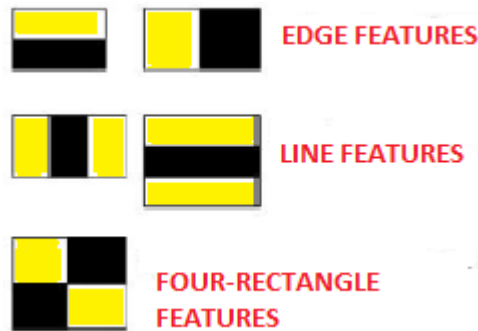
Fig.4.2 Cascade Classifier Labelling

The cascade classifier consists of number of levels. Each stage is designed with weak learners are said to be decision stumps. All levels are trained using boosting technique. Generally, it provides the ability to highly accurate classifier training simply considering weighted average of each decisions made by weak learners. Each level of classifier tags the region targeted by current location of sliding window as either positive or negative images. Positive images show the existence of an object and negative images indicates no objects existence. If tag is negative classification of interested region is complete, and detector slides window to next location. If tag

is positive, classifier passes region to next stage. when final level classifies the interested area as positive by neglecting less area of interest in negative images.

The levels are designed for less interest for negative images. The assumption is considered because most of current windows do not reflect the existence of object in input image. For verification of whether it is true/false positive or negative occurrence when positive or negative samples are correctly classified.Some of classification technique are summarized as
• when positive sample is classified correctly true positiveoccurs.
• when negative sample is classified positive by mistake false positiveoccurs.
• when positive sample is classified negative by mistake false negativeoccurs.

For good classification process each level should have low false negative rate and high false positive rate. If any level labels incorrectly an object as negative then classification process stops and cannot make mistake correctly.If classifier labels incorrectly non-object as positive, then correct the mistake in that level only. Adding more levels can reduces the overall false and true positive rate. Cascade classifier training requires full data set of positive and negative image samples.By providing set of positive images with regions of interest specified to be used as positive samples. With the help of Image Labeler, objects of interest with rectangle bounding boxes are identified. The Image Labeler outputs table to use for positive samples. By providing set of negative images from which function generates negative samples automatically. To achieve most acceptable design the parameter are to be modified for detector accuracy, data set number of stages, feature type, and other function parameters.

## 4.2 Multi Modularity of Convolutional Layers Supporting for Feature Extraction

Conv Nets is used for real-time classification and detection such as images recognition, images classifications, Objects detections, recognition faces etc. In neural network, image classification will find great process where processing input imageand classify it under certain categories.The input image consists of array of pixels and it depends on image resolution. Based on image resolution the mapping is done based on Height* Width*Dimension (h x w x d). Technically, deep learning CNN models to train and test, each input image will pass it through series of convolution layers with filters (Kernals), Pooling, fully connected layers (FC) and apply Soft-max function to classify an object with probabilistic values between 0 and 1. The complete flow of CNN to process an input image and classifies objects based on values is depicted in Fig.4.3
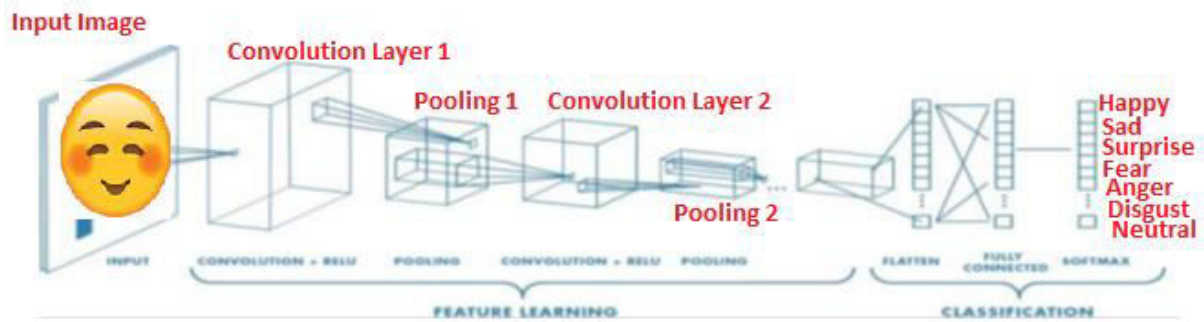
Fig. 4.3Multi modularity of Convolution Neural Network Model

To extract features from an given input image convolution is used.It preserves relationship between pixels by learning image features using small squares of input data. It is mathematical operation that takes two inputs such as image matrix and filter or kernel.
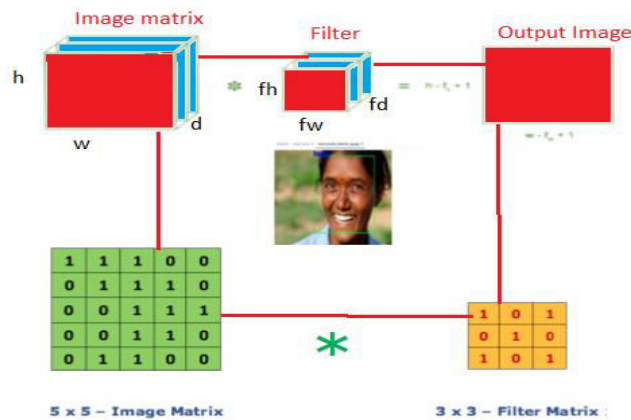
**Fig. 4.4 Matrix methodology for Convolution Neural Network**

Consider 5 x 5 image matrix whose image pixel values are (1 1 1 0 0), (0 1 1 1 0), (0 0 1 1 1),(0 0 1 1 0), (0 1 1 0 0) and filter matrix 3 x 3 [1 0 1], [0 1 0], [1 0 1] as shown in fig. 4.4 Convolution of an image with different filters can perform operations such as edge detection, blur and sharpen by applying filters.Stride is the number of pixels shifts over input matrix. When stride is 1 then move filters to 1 pixel at a time. When stride is 2 then move the filters to 2 pixels at a time and so on. Sometimes filter does not fit perfectly fit the input image. The correct solution is padding picture with zeros (zero-padding) so that it fits anddrop some part of image where filter did not fit. This is called valid padding which keeps only valid part of the image.

4.3 Design Constraints of Pooling Layer

When input images are large then pooling layers,technique is used to reduce number of parameters. Most of neural network prefers Spatial pooling technique. The technique which retains original information but reduces the dimensionality of each input map. It is also called subsampling or downsampling.Spatial pooling can be of different types such as Max Pooling,

Average Pooling and Sum Pooling. Max pooling takes the largest element from rectified feature map. Fig. 4.5 depicts the selection of pooling layers required for convolution neural network
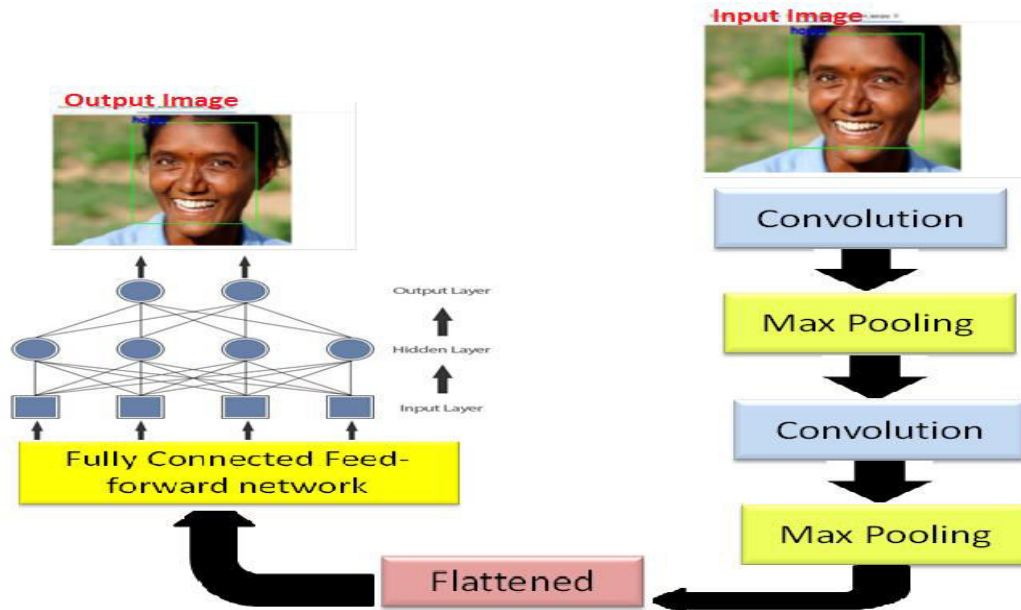
Fig4.5   Selection of Pooling Layers

Taking the largest element could also take average pooling. Sum of all elements in the feature map called as sum pooling. The fully connected layer flattened designed matrix parameter into vector parameter. Let us consider fig 4.6 where matrix mapping is carried in connection with fully connected layer and maximum pooling layers respectively.
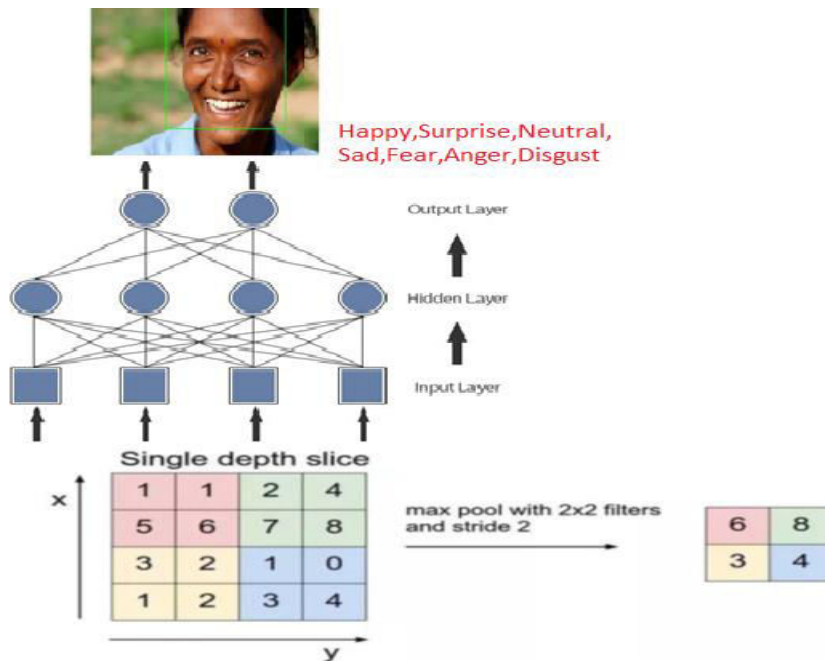
Fig.4.6 Mapping Matrix for Maximum Pooling Layer

Here feature map matrix will be converted as vector (x1, x2, x3, …). With fully connected layers, combined these features together to create model. The purpose is to introduce non-linearity in proposed Conv Net. Soft Max classifierRectified Linear Unit (ReLu) is non-linear operator. The output is f(x) = max(0,x).Since, real world data would want Conv Net to learn would be non-negative linear values. There are other nonlinear functions such as tanh or sigmoid that can also be used instead of Re LU. Most of the data scientists use ReLU since performance wise ReLU is better than the other two.

Logistic Regression (LR) is converted generalization of binary form Soft max classifier.The mapping function f in hinge loss or squared hinge loss is defined.The dot product of input data set x and weight matrix W to map them according to output class labels. It is defined as

$$f(x_i, W) = Wx_i$$

The objective of model is to predict and understand human emotions and to express using facial expression.

5. Data Flow Analytical Diagramfor Emotional Classification Methodology

Data flow analytical diagram for emotion classification using number of procedure steps is explained in fig.5.1.The process is sumarized   and respective algorithmasare described thoroughly. It started with applying input video sequence, Haar cascade classifier for face detection, feature extraction using convolution neural network, maximum poolingand finally emotional classification using facial expression.

Step 1: Consider the input image (Face)
Step 2: Use Haar Cascade classifier to detect the face from image.
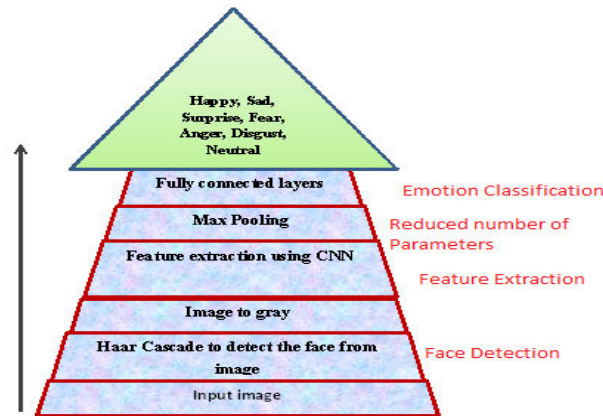Step 3: Convert the image to gray scale.
Step 4: Feature extraction is done using convolutional neural network.
Step 5: Convolution and max pooling are done multiple times.

Step 6: Then flatten layer converts it to one-dimension vector.

Step 7: Emotion classification is done in fully connected layers.
Step 8: Hence the emotion is obtained along with image successfully.

Fig. 5.1   Data flow Analytical Diagram

Algorithm 1–Training Steps to Create a Haar-like Classifier:

- Collection of positive and negative training images
- Marking positive images using objectmarker.exe or Image Clipper tools
  Creating a .vec (vector)
- File based on positive marked images using createsamples.exe
- Training the classifier using haartraining.exe
- Running the classifier using cv Haar DetectObjects ()

STEP 1: Collecting Image Database

- Database take 100 sample images for training. Add more positive and negative images by recording some sequences in HAKA1 or adding more public images from Internet resources.
- Selecting the positive images, that contain object (e.g. face or eye), and negatives images not contain the object.
- Select a greater number of positive and negative (back ground) images will normally cause a more accurate classifier

STEP 2: Arranging Negative Images

- Put input background images in folder…\training\negative and run the batch file create_list.bat dir /b *.jpg >bg.txt
- Running this batch file, will get a text file each line looks as below: image1200.jpg image1201.jpg image1202.jpg ..
- Later, need this negative data file for training the classifier.

STEP 3: Crop & Mark Positive Images

- Create a data file (vector file) that contains the names of positive images as well as location of objects in each image. It can create this file using two utilities: Object marker or Image Clipper.
- Select Object marker
- In folder ..\training\positive\raw data put positive images
- In folder ...\training\positive there is a file objectmaker.exe that for marking objects in positive images. In order to correctly run objectmaker.exe two files cv.dll and highgui.dll should also exist in the current directory.
- Before running the objectmaker.exe make sure to carefully mark and crop tens or hundreds of images.
- The next step is packing object images into vector-file

STEP 4: Creating Vector of Positive Images

- In folder ...\training\ there is a batch file named samples_creation.bat
- Content of the bath file is: createsamples.exe
- -info positive/info.txt, -vec vector/facevector.vec,-num 200,-w 24,-h 24
- Main Parameters: -info positive/info.txt Path for positive info file
- -vec vector/facevector.vec Path for the output vector file
- -num 200 Number of positive files to be packed in vector file -w 24 Width of objects,-h 24 Height of objects
- Batch file loads info.txt and packs object images into vector file with the name of e.g. face vector.vec
- After running batch file, will have the file facevector.vec in folder...\training\vector

STEP 5: Haar-Training

- In folder ..\training , modify the haartraining.bat :
- haartraining.exe, -data cascades, -vec vector/vector.vec, -bg negative/bg.txt,-npos 200,-nneg 200,- nstages 15, -mem 1024,-mode ALL,-w 24 -h 24,–nonsym -data cascades Path and for storing the cascade of classifiers
- -vec data/vector.vec Path which points the location of vector file
- -bg negative/bg.txt Path which points to background file
- -npos 200 Number of positive samples = no. positive bmp files
- -nneg 200 Number of negative samples (patches) = npos
- -nstages 15 Number of intended stages for training
- -mem 1024 Quantity of memory assigned in MB
- -mode ALL Look literatures for more info about this parameter -w 24 -h 24 Sample size
- -nonsym Use this if your subject is not horizontally symmetrical

6.Experimentation and Result Analysis

Numbers of experiments are carried out to enhance the emotions classification based on facial expression using CNN.The deep learning network is trained,test and validated using CNN. It is concluded that in real time the selected area in image will efficiently be classified using CNN.

6.1 Design Constraint of Facial Expression Recognition-2013 Database Experiment

Pierre-Luc Carrier and Aaron Courville has created an open-source dataset which is shared publicly for Kaggle competition during ICML 2013. The proposed network is first trained using database made available for Facial Expression Recognition Challenge. The dataset consists of 35.887 Gray scale, 48x48 sized face images with various emotions. Changes are made to both VGG-16 model and database to make them compatible for training. VGG-16 model's input image layer is changed from [224, 224, 3] to [48, 48, 3] along with last 3 layers of model. The input image dataset is made changes of Grayscale [48, 48, 1] to [48, 48, 3]. The pre-trained network model has been constructed into CNN feature extraction network in integration with pooling layer and convolution layer. The data set used for proposed work is depicted in fig.6.1

6.1 Facial Expression Recognition-2013 Database

6.2 Training Constraint of Neural Network Model

There are a greater number of deciding factor when a neural network is trained. The requirements include Augmented Image Data Generator (AIDG), 2D Convolutional Layers, Spatial Data for 2D with Max Pooling operation and Sequential model are trained using Keras and Tensor Flow.

a) Augmented Image Data Generator (AIDG)

In Keras images are implemented using Image Data Generator API.It generates batches of image data with real-time data augmentation. To train deep neural network, the efficient codes to create and configure AIDG are as follows.

```
Datagen = Image Data Generator()
Datagen.fit (train)
X_batch, y_batch = Datagen.flow (X_train, y_train, batch_size=batch_size)
model.fit_generator(datagen, samples_per_epoch=len(train), epochs=epochs)
```

In proposed case, data generator generates batch of 9 augmented images with rotation by 30 degrees and horizontal shift by 0.5.

```
Datagen = ImageDataGenerator(rotation_range=30, horizontal_flip=0.5)
Datagen.fit(img)
i=0
for img_batch in datagen.flow(img, batch_size=9):
for img in img_batch:
plt.subplot(330 + 1 + i)
plt.imshow(img)
i=i+1
if i >= batch_size:
break
```

b) 2D Convolutional Layers Constraint

Let us consider input image of 3-D with three color channels (RGB). Input image is passingthrough filter called convolution kernel.At a time, inspecting small window of pixels over the input image.For image of 3×3 or 5×5 pixels, task is to moving the window till full image being scanned. The convolution operation proceeds with dot product of pixel values in current filter window with pre-defined weights.Using keras. layers. Conv2D () function we can create 2D convolutional layers in Keras.As defined in TensorFlow Conv2D process, in Kerasno need to define variables or separately construct the activations and pooling.It does automatically and code sample creates 2-D convolutional layer.

```
keras.layers.Conv2D(filters, kernel_size, strides=(1, 1),
padding='valid',
data_format=None,
dilation_rate=(1, 1),
Activation=None,
Use_bias=True,
kernel_initializer='glorot_uniform',
 bias_initializer='zeros',
kernel_regularizer=None,
Bias_regularizer=None,
Activity_regularizer=None,
kernel_constraint=None,
Bias_constraint=None).
```

c) Spatial Data for 2D with Max Pooling Operation

The input image is represented by considering either maximum or minimum value using pool size for each dimension along features trend. For this input image should undergone downsampling or up sampling. The window is moved by strides in every dimension features.Finally the result shows output window using "valid" padding option which has number of rows or columns called shape of:

$$output\_shape = (input\_shape - pool\_size + 1) / strides)$$
The resulting output shape when using the "same" padding option is:
$$output\_shape = input\_shape / strides.$$

d) Sequential model for stack of Layers:

For plain stack of layers, it is designed witheach layer should have one input tensor and one output tensor respectively.By passing list of layer instances to the constructor, a sequential model is constructed.

```
from keras.models import Sequential
from keras.layers import Dense, Activation
model = Sequential(
Dense(32, input_dim=784),
Activation('relu'),
Dense(10),
Activation('softmax'))
```

In the experimental verification on Facial Expression Recognition-2013 Database, the paper proposes seven emotions using different facial expression.

Case1 – Emotion Analyzer for Image with Single Face
In the testing phase, various OpenCv functions and Keras functions have been implemented. In the initial stage image and video is stored in frame object.Haar cascade classifier is used to detect facial expression layout. The image frame is converted in to grayscale and resized for further processing. The resized image is loaded with keras model function and maximum argument is output. A rectangular box is drawn around the facial image and output is formatted above the rectangular box.
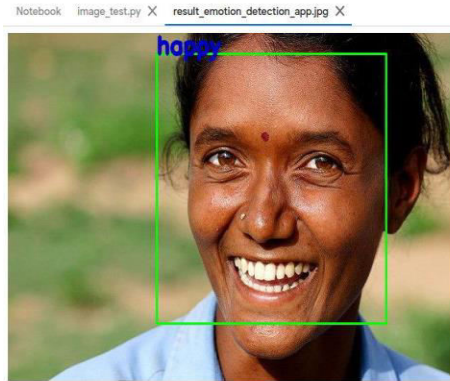

Fig.6.1 Single Face Image

Case 2 Image with Multiple faces
Fig. shows seven emotions with different facial expression.
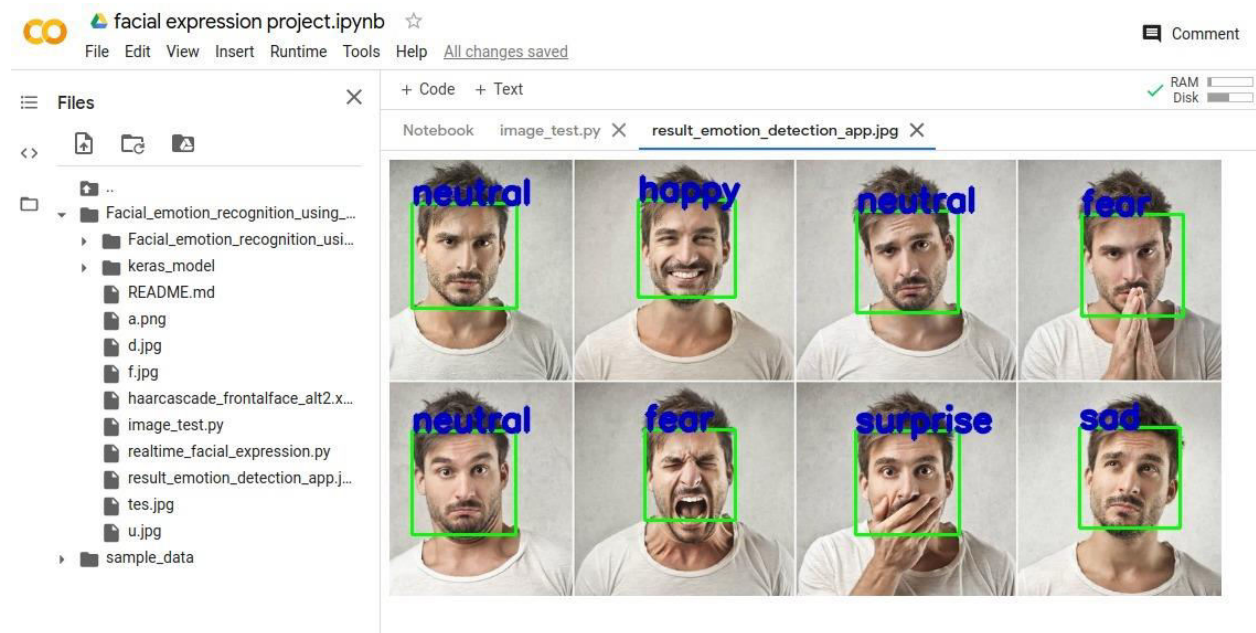

Fig. 6.2 Multiple Face Images

Table1 shows the design parameters using FER 2013 data set and proposed CNN training set.

Table 1. Chart Showing Pre-Data Set and Training Set

FER -2013 Data   **CNN Training Set**

Pre- defined Emotions	Training Set	Data Properties	Image Dimension	Positive Overlap Range	Negative Overlap Range	Batch size
Happy Neutral Surprise Fear Angry Disgust Sad	28,000 examples	48*48 Pixel Gray scale, 1 Bit	600*800	[0.3 1]	[0 0.3]	1

The recognition rate of seven emotions labels is represented in pictorial diagram. The graph shows comparison between prediction model and proposed model. The recognition rate accuracy is set maximum of 90% for prediction model whereas proposed model shows 95% maximum accuracy level.
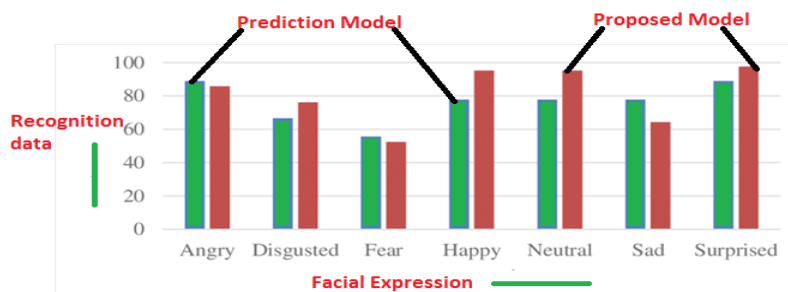
Fig.6.3 Graph Showing Recognition Rate with Seven Labeled Emotions

Table 2 shows the comparison of emotional attributes for prediction model and our proposed model respectively. The prediction model highest accuracy for angry and surprise emotion with 85% followed by least accuracy for fear with 50%. By comparing with proposed model, the highest accuracy for happy emotion with 99%, followed by surprise with 98%, neutral with 96% and least accuracy for fear emotion with 45%.

Table 2. Comparison Chart for Emotional Attributes

Emotion attributes	Prediction Model (%)	Proposed Model (%)	
Happy	79	99	**Highest Accuracy**
Surprise	85	98	
Neutral	78	96	
Fear	50	45	**Lowest Accuracy**
Sad	77	65	
Angry	85	83	
Disgust	65	75	

7.Live Validation Test

Many experimental trails are made to test and verify predictive and proposed cognitive model which are trained on FER-2013 data set. In real–time, the network performed better performance by classifying facially expression through a webcam device by comparing emotional attributes for prediction model and our proposed model respectively. Using trained CNN the following validation designed conditions are to be considered.

- Live input from web-cam.
- Facial expressed emotions only
- Test only for pre-defined seven emotions
- Compare predictive model with proposed model.
- Webcam's resolution is set to 420x220
- Distance 50 cm from the face is maintained
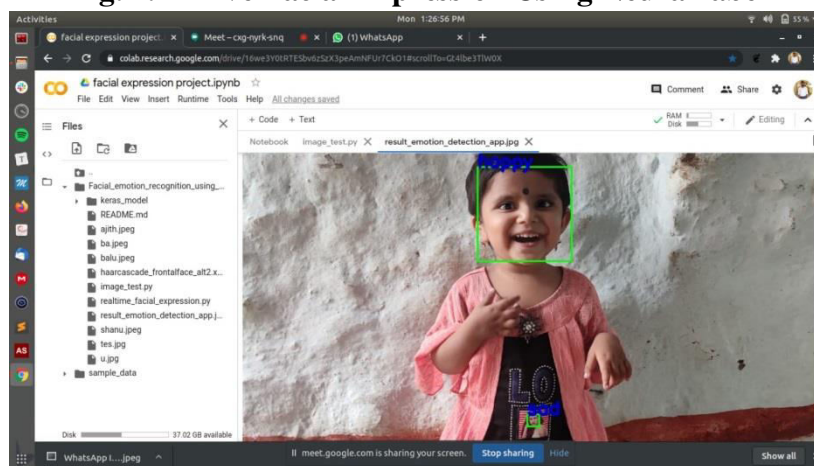
Live Verification

The emotions are classified accurately according to training model using FER-2013 data set. The model should identify face and successfully classifies seven emotions such as Neutral, Happy, Surprise, Sad, Angry, Disgust and Fear respectively. The trained CNN is verified by pooled region of concerned image of face.A rectangular box is drawn around the facial image and output is formatted above the rectangular box. By focusing on interested area on the face it will capture facial expressions.Trained CNN classifies different facial expressions for setting webcam's resolution to low 420x220 and 50 cm distance is maintained from the face. But trained CNN will not detect emotions using facial expressions when webcam's resolution is set to high with distance greater than 50 cm from face. Under this condition no facial expressed emotions are recorded.Finally, it is concluded that for webcam resolution of 420x220 with distance of 50 cm

or less between webcam and face will capture similar face size by the proposed model during the training of CNN.

The proposed model is trained and real time (Live) tested on different facial expressions like happy, neutral and surprise of different people. Fig.7.1,7.2,7.3 shows people who tried to express facial emotions happy and neutral. The model performed excellent in recognizing happy, neutral and surprised faces. Some people made their contribution in facially expressed emotions using selfies.
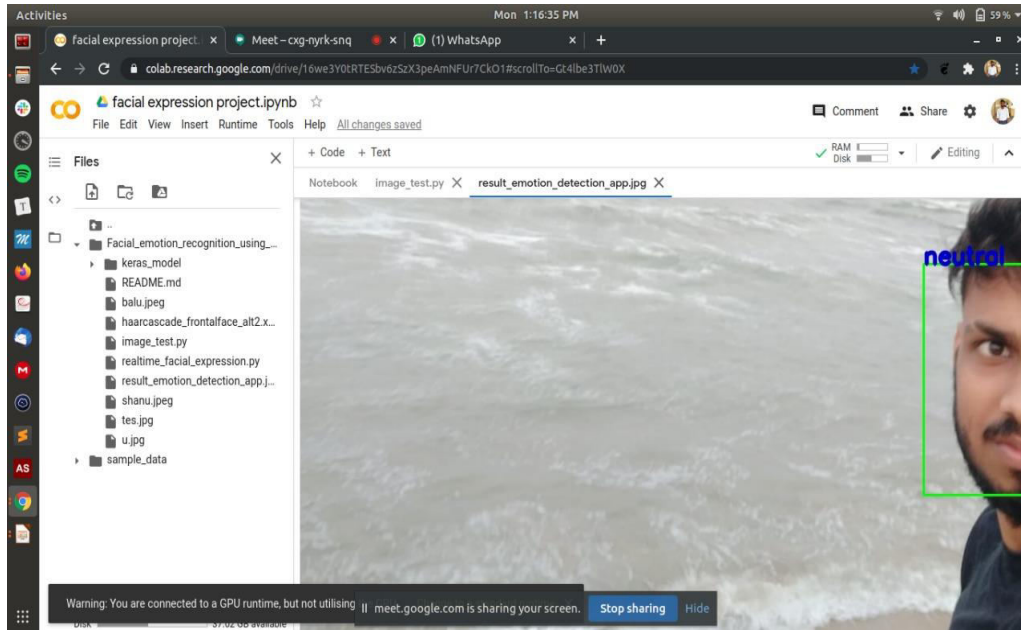
Fig. 7.1   Live Facial Expression Using Neural label

Fig. 7.2 Live Facial Expression for Happy Label
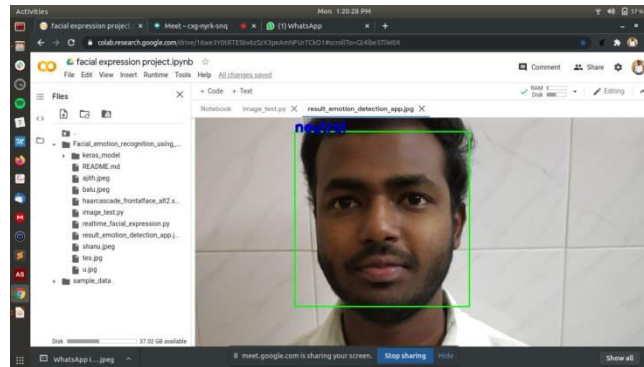
Fig. 7.3 Live Facial Expression for Neutral Label

Conclusion

The emotions are classified accurately according to training results. The model identifies face and successfully classifies seven emotions such as Neutral, Happy, Surprise, Sad, Angry, Disgust and Fear respectively. The human machine interfacing information about facial expression based advanced technology will lead to enrich predictive and proposed model should agree with the level of results.Also, the better trained models can be used to predict emotions with higher accuracy. This model can also be used in predicting happiness indexes and in the health sector.

Many experimental trails are made to test and verify predictive and proposed cognitive model which are trained on FER-2013 data set.In real–time, the network performed better performance by classifying facially expression through a webcam device by comparing emotional attributes for prediction model and our proposed model respectively. The prediction model highest accuracy for angry and surprise emotion with 85% followed by least accuracy for fear with 50%. By comparing with proposed model, the highest accuracy for happy emotion with 99%, followed by surprise with 98%, neutral with 96% and least accuracy for fear emotion with 45%. The proposed model can also be used in predicting happiness index and in health sector. The innovation of proposed system can list as faster,can get input from different cameras. Changing code to be more efficient with clear visualizations, mask loading and recovery rate is high.

References

[1] T. Ahsan, T. Jabid, and U.-P. Chong. Facial expression recognition using local transitional pattern on gabor filtered facial images. IETE Technical Review, 30(1):47–52, 2013.

[2] D. Ciresan, U. Meier, and J. Schmidhuber. Multi-column deep neural networks for image classifification. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on, pages 3642–3649. IEEE, 2012.

[3] C. R. Darwin. The expression of the emotions in man and animals. John Murray, London, 1872.

[4] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, pages 248–255. IEEE, 2009.

[5] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. Journal of personality and social psychology, 17(2):124, 1971.

[6] B. Fasel and J. Luettin. Automatic facial expression analysis: a survey. Pattern recognition, 36(1):259–275, 2003.

[7] A. Gudi. Recognizing semantic features in faces using deep learning. arXiv preprint arXiv:1512.00743, 2015.

[8] Kaggle. Challenges in representation learning: Facial expression recognition challenge, 2013.

[9] A. Krizhevsky and G. Hinton. Learning multiple layers of features from tiny images, 2009.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenetclassifification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012.

[11] O. Langner, R. Dotsch, G. Bijlstra, D. H. Wigboldus, S. T. Hawk, and A. van Knippenberg. Presentation and validation of the radboud faces database. Cognition and emotion, 24(8):1377–1388, 2010. EMOTION CLASSIFICATION FROM FACIAL EXPRESSIONS 2019-2020 Dept. of ECE Page 52

[12] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specifified expression. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on, pages 94–101. IEEE, 2010.

[13] Y. Lv, Z. Feng, and C. Xu. Facial expression recognition via deep learning. In Smart Computing(SMARTCOMP), 2014 International Conference on, pages 303–308. IEEE, 2014.

[14] J. Nicholson, K. Takahashi, and R. Nakatsu. Emotion recognition in speech using neural networks. Neural computing & applications, 9(4):290–296, 2000.

[15] A. Mehrabian, Communication without words, psychology today, vol. 2, no. 4, pp. 53- 56, 1968.

[16] NicuSebe, Michael S, Lew, Ira Cohen, Ashutosh Garg, Thomas S. Huang, "Emotion recognition using a Cauchy naïve bayes classifier", ICPR, 2002.

[17] P. Ekman, W.V. Friesen, "Facial action coding system: investigator's guide", Consulting Psychologists Press, Palo Alto, CA, 1978.

[18] G. Little Wort, I. Fasel. M. Stewart Bartlett, J. Movellan "Fully automatic coding of basic expressions from video", University of California, San Diego.

[19] M.S. Lew, T.S. Huang, and K. Wong, Learning and feature Selection in Stereo Matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 16, no. 9, 1994.

[20] Ira Cohen NicuSebe, Larry Chen, Ashutosh Garg, Thomas Huang, "Facial Expression Recognition from Video Sequences: Temporal and Static modeling Computer Vision and Image Understanding(CVIU) special issue on face recognition.

[21] P.Ekman. Strong evidence of universals in facial expressions: A reply to Russell's mistaken critique. Psychological Bulletin, pp. 268-287, 1994

Author Details

Dr. Tanuja.P.Patgar received B.E degree in Electronics and Communication from Kuvempu university, Karnataka, India in 1996. In 2010, she received M.E in Control and Instrumentation from University Vishweshraya Collage of Engineering, Bangalore, India. She received her PhD in "Performance Analysis of Communication Based Train Control System using WSN" from Visvesvaraya Technological University, Belgaum, India in 2020. Her research field is Wireless Sensor Network, Artificial Neural Network, Machine learning, Deep Learning, Data Science and Computer Vision. Presently serving as Professor at Dr. Ambedkar Institute of Technology, Bangalore, India

Triveni received B.E degree in Electronics and Communication Engineering from V.T.U University, Karnataka, India in 2005.In 2010; she received M.Tech in Digital Communication from V.T.U University, Karnataka, India. Her Research field is Embedded systems, Robotics, Artificial Nueral Network, Machine learning, Data science. Presently serving as Assistant Professor at Dr. Ambedkar Institute of Technology, Bengaluru, India.