

Effective Facial Emotion Recognition using Convolutional Neural Network Algorithm



MalyalaDivya, R Obula Konda Reddy, C Raghavendra

Abstract: This paper presents the idea related to automated live facial emotion recognition through image processing and artificial intelligence (AI) techniques. It is a challenging task for a computer vision to recognize as same as humans through AI. Face detection plays a vital role in emotion recognition. Emotions are classified as happy, sad, disgust, angry, neutral, fear, and surprise. Other aspects such as speech, eye contact, frequency of the voice, and heartbeat are considered. Nowadays face recognition is more efficient and used for many real-time applications due to security purposes. We detect emotion by scanning (static) images or with the (dynamic) recording. Features extracting can be done like eyes, nose, and mouth for face detection. The convolutional neural network (CNN) algorithm follows steps as max-pooling (maximum feature extraction) and flattening.

Keywords: Convolutional Neural Network, Face Detection, Face Detection, Feature Extraction, Image processing, Opencv, Tensorflow.

I. INTRODUCTION:

Emotion plays a crucial role to understand the feeling of each person's figures clearly about a person's personality. Efficient face recognition detection is an advanced technology that helps us to identify individuals across the globe. Here, taking a trained dataset that constitutes static images and captured images from a video or in life. The dynamic image classifier looks simple as we think but it is very sensitive for facial emotions appearance for each individual. Multi-model human-computer interaction (MMHCI) lies in several investigation areas such as computer vision, psychology, artificial intelligence, and many other things. Human intelligence is very effective to reinforce our feelings to the next person. Emotion is manifested in a multi-modal way. Human communication interaction (HCI) regularities can sense the activity levels for each person like emotion, attention and many things. Each feature has its features such as detect eyebrows, detects mouths, and eyeballs. Besides they differentiate two types of classification schemes as static and dynamic images [4].

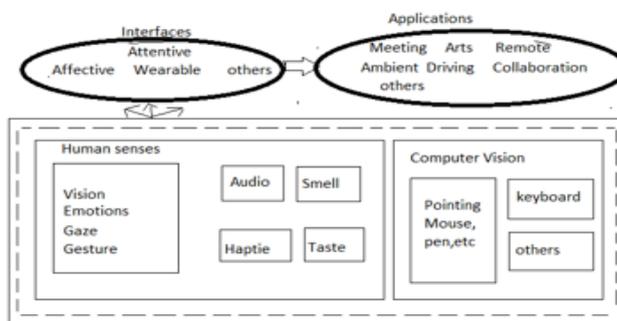


Fig 1. Multi-mode human-computer interaction (MMHCI)

Happy emotion can be identified by following some features like lip corners and cheeks raised, occurred wrinkles and muscle become tight around the eyes. Similarly, angry features can be stated with upper lids pulled up, lips may be tightened, and lower lids pulled up. Sad features are lip corners pulled down, inner corners of eyebrows raised. Disgust features are eyebrows pulled down, upper lip pulled up, nose wrinkled. Surprise features are eyelids pulled up, mouth hangs open, and entire eyebrow pulled up. Neutral features are all features placed at their respective positions. Fear features are eyebrows pulled up, upper eyelids pulled up, and mouth stretched [7][8][9].

1.1 Terms and Methodology:

Haar features extracts edge features, line features, center-surrounded features, and four-rectangles features (Example which used for eye detection). It uses the AdaBoost algorithm for the best features out of 160000+ features. The window will pass all stages in a face region for all haar features for recognition of each feature. OpenCV comes with a trainer as well as a detector, if a project requires its own classifiers for any object such as car we can train our own. OpenCV contains inbuilt trained classifiers for face eyes, smile, and many more features for living and non-living beings. Cascade classifier raises its own positive and negative images and with its feedback. It trains cascade object detector that stores the data as an XML file in the form of stages and that can also create its own cascade classifier.

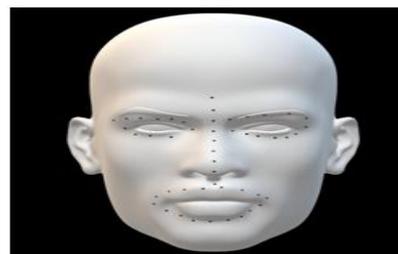


Fig 2: Face Feature Extraction

Manuscript published on November 30, 2019.

* Correspondence Author

M Divya*, Dept. of CSE, Institute of Aeronautical Engineering, Dundigal, Hyderabad, India.

Dr. R Obula Konda Reddy*, Professor, Dept. of CSE, Institute of Aeronautical Engineering, Dundigal, Hyderabad, India.

C Raghavendra, Asst. Professor, Dept. of CSE, Institute of Aeronautical Engineering, Dundigal, Hyderabad, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Image Processing:

Image processing is used to extract essential features from static images where useful information will be gathered. An image is a two-dimensional figure where x and y spatial coordinates of function $f(x,y)$ are represented through a matrix with rows and columns in the form of pixels.

The steps for images processing are

1. Reading the image as input
2. Pre-process (Noisy removal and shaping) of the image
3. Extracting features on the input image (segmentation, description)
4. The result is an image classification after manipulation

We have two types of image processing:

1. Analog image processing
2. Digital image processing (computer vision)

Some of the important techniques for image processing are preprocessing, enhancement and, information extraction. An image is converted into grayscale to perform a specific task whereas to display by converting into RGB (color) for user view.

II. EXISTING METHOD:

Facial Action Coding System (FACS) developed by Paul Ekman & Wallace Friesen with forty-four animation units (AU) and AU is a feature true points, it is a technique for the measurement of facial moments and expressions. In 2002, they completed an automatic based system for representing the movement for each person. Facial expressions described by various combinations of animations units. It gives a sketch of gesture, gaze, and effective interaction such as facial expression recognition, and emotion recognition while capturing [3] Emotion Recognition through wireless signal through EQ-radio uses both machine learning concept and wireless transmission which comes under analog image processing [1]. Facial Emotion Recognition is done through OpenCV Machine Learning technique SVM (Support Vector Machine) and OpenCV libraries [2]. Principal Component Analysis (PCA) [3] through face recognition plays various roles in real-time smart car lock security systems by PCA and image processing techniques. As biometric is having small drawbacks, so to enhance they implemented face recognition attendance system through OpenCV and face recognition techniques [6].

2.1 Convolutional Neural Network (CNN)

The convolutional network named CNN or ConvNet is a feed-forward artificial neural network and also called multi-layer perceptrons (MLPs). The main objectives are image recognition and image classifications which are widely used in the present generation. It starts with reading the input image where computer vision will be in the form of arrays of pixels and it refers to the number of pixels called image resolution. Then it converts into three dimensional-h*w*d (h-height, w-width, d-dimension) with RGB in 6*6*3 and grayscale in 4*4*1.

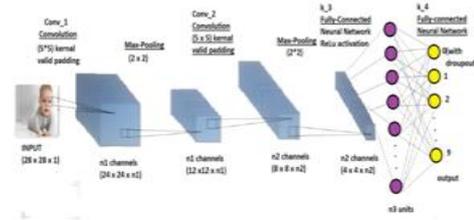


Fig 3: CNN Process

Convolution Layer extracts features from an input image from the training dataset where the kernel is a small part of an input image for feature classification. It performs some mathematical operations on image matrices and kernels such as RELU (Rectified Linear Unit). Which refers to the Rectifier unit to classify the object as 1 ('Yes') or 0 ('No') with a probabilistic value which lies between 0 and 1 to avoid negative pixels we use activation function[10]. Pooling is the next layer, here it performs extracting the maximum features for the input and classifies the object through max pooling, average pooling, and global pooling. Flatten layer used for reshaping from 3D to 2D and it is placed between the convolutional layer and fully connected layer[11]. Flattening is a two-dimensional matrix converted to features into a vector to a fully connected neural network [12].

III. PROPOSED METHODOLOGY:

The design starts with the initializing CNN model by taking an input image (static or dynamic) by adding a convolution layer, pooling layer, flatten layers, and dense layers. Convolution layers will be added for better accuracy for large datasets. The dataset is collected from CSV file (in pixels format) and it's converted into images and then classify emotions with respective expressions. Here emotions are classified as happy, sad, angry, surprise, neutral, disgust, and fear with 34,488 images for the training dataset and 1,250 for testing. Each emotion is expressed with different facial features like eyebrows, opening the mouth, Raised cheeks, wrinkles around the nose, wide-open eyelids and many others. Trained the large dataset for better accuracy and result that is the object class for an input image. Based on those features it performs convolution layers and max pooling. These are the seven different universal emotions with the following expressions below.



Fig 4: Different Human Facial Emotions

The feature will be extracted through the max-pooling method by creating the model with .h5 extension and then compile the model with loss and optimizer. Here we import haar cascade for face recognition which is in XML format.

OpenCV:

Open Source Computer Vision Library provides a common infrastructure for computer vision applications for humans and computer vision which contains 2500 optimized algorithms. These algorithms used for face detection, identification of objects for training and detecting objects.

TensorFlow:

TensorFlow is a second-generation arrangement for the google company and the implementation and deployment of large-scale machine learning projects. It is flexible enough to be used both in research and product invention. It creates large scale neural networks, used for creation, classification, discovery, prediction, prescription[15]. The main applications of Tensorflow are the voice to text and text to voice, recognition while capturing video, audio, image, and time series, and text-based applications.

Keras:

Keras is an open-source neural network in python, which is used for the preprocessing, modeling, evaluating, and optimization. It is used for high-level API as it handled by backend. It is designed for making a model with loss and optimizer function, and training process with fit function. Keras does not support low-level graphs and computations as it handled by the backend engine[14]. For backend, it designed for convolution and low-level computation under tensors or TensorFlow.

Importing the below python libraries are used for preprocessing, modelling, optimization, and testing.

```
#import keras libraries
from keras.models import Sequential
from keras.layers import Dense
from keras.layers import Conv2D
from keras.layers import MaxPooling2D
from keras.layers import Flatten
from keras.preprocessing.image import ImageDataGenerator
from keras.models import load_model
import numpy as np
import cv2
from keras.preprocessing.image import img_to_array
```

Fig 4: Python Libraries

Applications:

Airports: Observe the pilot’s psychological condition before take-off.

Hospitals: Can be performed on a psychological disorder patient by a psychiatric doctor.

Crime Department: Used as a lie detector.

Social websites: Feedback depicted through the face in the absence of written feedback or rating.

Social Welfare: Gathering information would be profitable in the case of deaf and dumb people.

Driver Monitoring: Monitoring driver facial expressions while driving.

Result Analysis:

The emotion percent varies for each second, it displays each and every universal emotion. It displays a number of faces

detected and at last, it will display emotion which having a maximum percentage[13].

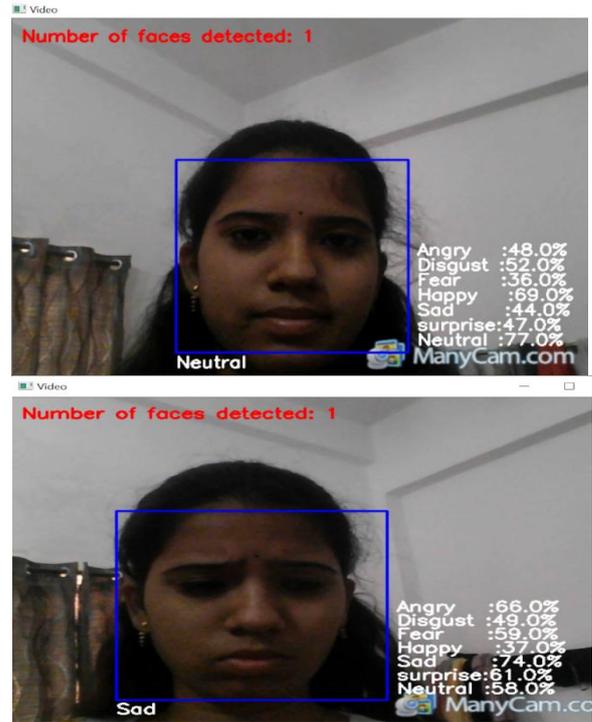


Fig 5: Sample Results

IV. CONCLUSION AND FUTURE SCOPE:

The outcome of the project is to display the percentage of every emotion classifier to each individual i.e., 66% of accuracy is achieved. Future scope for this paper by adding an effective dataset and little modifications in techniques will help us to reach better accuracy.

REFERENCES:

1. Mingmin Zhao, FadelAdib, Dina Katabi, Emotion Recognition using Wireless Signals, Massachusetts Institute of Technology, 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN)
2. Shivam Gupta, Facial emotion recognition in real-time and static images, Proceedings of the Second International Conference on Inventive Systems and Control (ICISC 2018) IEEE Xplore Compliant - Part Number: CFP18J06-ART, ISBN:978-1-5386-0807-4; DVD Part Number: CFP18J06DVD, ISBN:978-1-5386-0806-7.
3. S.Padmapriya& 2Esther Annlin KalaJames, Government College of Technology, Real Time Smart Car Lock Security System Using Face Detection and Recognition, 2012 International Conference on Computer Communication and Informatics (ICCCI -2012), Jan. 10 – 12, 2012, Coimbatore, INDIA.
4. Alejandro Jaimes,NicuSebeb, an IDIAP, Switzerland b University of Amsterdam, The Netherlands, Multimodal human-computer interaction: A survey. Received 21 September 2005; accepted 30 October 2006, available online 31 January 2007, communicated by Mathias Kolsch.
5. Alejandro Jaimes a, NicuSebe, an IDIAP, Computer Vision and Image Understanding Multimodal human-computer interaction: A survey. The Switzerland b University of Amsterdam, The Netherlands, Received 21 September 2005; accepted 30 October 2006
6. Adrian RhesaSeptianSiswanto, AntoSatriyoNugroho, MaulahikmahGalinium, Information Technology, Swiss German University, Indonesia. Implementation of Face Recognition Algorithm for Biometrics Based Time Attendance System. Center for Information & Communication Technology.

7. Kruti Goyal, Kartikey Agarwal, Rishi Kumar CSE, Amity, University Noida, India. Face Detection and Tracking Using OpenCV. International Conference on Electronics, Communication, and Aerospace Technology ICECA 2017.
8. Hyeon-Jung Lee, Sungkyunkwan Hong College of Information and Communication Engineering University Suwon, 440-746, South Korea, A Study on Emotion Recognition Method and Its Application using Face Image.
9. Leih Luoh, Chih-Chang Huang and Hsueh-Yen Liu, Chung-Hua University, 2010 International Conference on System Science and Engineering Image processing based emotion recognition, 2010 International Conference on System Science and Engineering.
10. A. Kumaravel, S. Sivasubramanian, "Identifying images on moving objects to enhance the recognition", International Journal of Engineering & Technology.
11. S. Sivasubramanian, A. Kumaravel, "Improved image compression using effective lossless compression technique", Cluster Computing, March 2018.
12. C. Nalini, K. Rajendra Prasad, C. Raghavendra, "Comparative Observation and Performance Analysis of Multiple Algorithms on Iris Data", International Journal of Pure and Applied Mathematics.
13. C. Raghavendra, K. Rajendra Prasad, Padakandla Vyshnav, "Intelligent System For Visualized Data Analytics A Review", International Journal of Pure and Applied Mathematics.
14. <https://keras.io/>
15. <https://developers.google.com/machine-learning/crash-course/first-steps-with-tensorflow/toolkit>

AUTHORS PROFILE



Ms. Malyala Divya is pursuing her Bachelor's degree in Institute of Aeronautical Engineering, Hyderabad. She completed internship on artificial intelligence and machine learning. Her interested areas are Data Science, Artificial Intelligence and Machine Learning. She did projects on image processing and artificial intelligence.



Dr. R. Obulakonda Reddy received Bachelor's Degree in Computer Science and Engineering from Sri Krishnadevaraya University, Anantapur, Master's Degree in Computer Science and Engineering from JNTUH, Hyderabad and Ph.D from JNTUA, Anantapur. He is having 12 years of teaching and 2 years research experience. He published and presented research papers in national and international conferences and reputed journals. His area of interest includes Pattern Recognition, Image Analysis, Cloud Computing and Network Systems.



Mr. C. Raghavendra Graduated in B.Tech (CSE) from Jawaharlal Nehru Technological University, Hyderabad in 2008. He received Master's Degree in M.Tech (CSE) from Bharath INSTITUTE OF Higher Education and Research (BIHER), Chennai, in 2011. He is pursuing a Ph.D. in Computer Science & Engineering from BIHER, Chennai. Presently, he is working as Assistant Professor of CSE Dept., Institute of Aeronautical Engineering, Hyderabad. He has published more than 20 Publications in various International Journals and conferences. His research interests are Image processing & Security, Big data and Databases. He is a life member of CSI, IAENG, CSTA professional societies.