



Emotion Recognition System for Visually Impaired

E. Kodhai, A. Pooveswari, P. Sharmila, N. Ramiya

Abstract: Machine learning is one of the current technologies that use computers to perform tasks similar to humans. It is adopted in many applications like face recognition, Chabots, self-driving cars etc. This work focuses on emotion recognition which is part of computer vision technology. Emotion recognition is mainly used in cybersecurity, online shopping, police investigations, and interview process and so on. In this paper, an emotion recognition system is built for the visually impaired people. The blind people cannot recognize the facial expressions of the person interacting with them. These people can be provided with a device that can recognize the emotions of people through a camera and conveys the kind of emotion via headphones. The system will be made using a Raspberry Pi computer to perform the entire task and it is portable for the user. The emotion recognition model will be trained using convolution neural network (CNN) with the fer2013 dataset that contains more than 30,000 images. The human face is detected using the OpenCV library and some features like Histogram of Oriented gradients (HOG) are also passed with input images for better accuracy. The recognized emotion is then converted to a speech using a python library Pytsx3 that make use of eSpeak engine.

Keywords : Facial Expressions, FER2013, Raspberry Pi, Convolution neural network(CNN).

I. INTRODUCTION

The Eye is one of the vital organ for humans. Eye enables humans to perform various daily tasks and to learn about the world. Our human eye has a lot of features like a three dimensional, moving image, normally colored in daylight. It also extracts some important features from various images so that the decision can be taken for what the image is all about. Nowadays, the computer is being trained in such a way which attains all features of the human eye that can predict some specific result by taking images as input which work like the human visual system. This refers to computer vision

technology [13]. Although a machine can recognize the objects, it analyses the objects as a binary data whereas human beings will see the objects as it is. Computer vision is recently used for creating applications such as face recognition, drowsiness detection, and anomaly detection systems. The face detection system is a recent trend in every device like mobile phones and personal computers. It is used as an authentication tool for unlocking the system, as a way for person verification and identification in many organizations. Emotion recognition is also a recent technology in computer vision. Identifying the facial expressions of a person when communicating is a complex task for a visually impaired person. However computer vision is used here to help such people recognize the emotions of the confronting person. The proposed system is developed for the visually challenged people. The visually impaired person can be provided with an emotion recognition system that detects the facial expressions of the people communicating with them. The expressions like sad, happy, angry, disgust can be identified. The system can be made handy and portable using Raspberry pi [19] so that the visually challenged person can carry wherever they travel. The works related to emotion recognition and raspberry pi are explained in the following section. The Section III contains the explanation of the proposed system, system architecture and modules. The conclusion of the proposed system is given in section IV.

II. RELATED WORKS

Tuhin Kundu et.al[1] used facial analysis and ML models to recommend developments and recent trends in emotion recognition. In this paper, human emotions such as disgust, happiness, drowsiness, surprise, normality are classified using artificial neural networks (ANN) and Support Vector Machine (SVM) [17]. Whereas Ma Xiaoxi et.al [3] proposed Facial Emotion Recognition on different learning methods, and has implemented several methods: Support Vector Machine (SVM) and Deep Boltzmann Machine (DBM) for facial emotion recognition. Both Tuhin kundu and Maxoami used The FER 2015for training. Training can be done in two ways. First, it could transform the information conveyed by the mouth and eye facial regions into a combined new image and use it as an input to a neural network feed forward. The second method was the extraction of texture information that showcases the use of Oriented Fast and Rotated (ORB) on a single frame of imagery to, and the classification is completed using SVM [17]. Using facial imagery by pattern classification, the drowsiness detection is automated to play a revolutionary role in preventing the road accidents which was created by the lethargic activities of drivers. Some of them used an attentional convolutional network for emotion recognition.

Revised Manuscript Received on March 17, 2020.

* Correspondence Author

E. Kodhai*, Associate Professor, Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India. Email: kodhaij@gmail.com

A. Pooveswari, Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India. Email: pooveswariarul@gmail.com

P. Sharmila, Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India. Email: sharmi2898@gmail.com

N. Ramiya, Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India. Email: ramiya1998@gmail.com

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

In that case, Shervin Minaee, Amirali Abdolrashidi [7] proposed face expression recognition using an attentional convolutional network. SIFT, HOG and LBP are used as traditional approaches for this problem. It focuses on the required parts of the human face, and achieves a great improvement from recent models on various datasets, including JAFFE, CK+, FER-2013, FER2015, FER-G.

Raspberry is a portable device which is used for many machine learning projects. T.Anuradha et.al [2] proposed Embedded system on Raspberry Pi was mainly focuses on the study of the methods existing in the present scenario for detection of motion and faces using various kinds of algorithms, techniques and discuss about their area of application by implementing with the help of on the on-board miniature Raspberry Pi. Open CV [26] functions are optimized to the specific platform and can be further used as a surveillance system which is embedded on the computer. The open CV libraries contain all the functions that can be operated and optimized on the Raspberry Pi platform.

The emotion recognition model can also be implemented in Raspberry Pi. Prof. V.D. Bharate et.al [5] proposed emotion detection using raspberry-pi. The system is in the field of Human Computer Interaction (HCI). The entire project is subdivided into the following steps: Face detection, facial feature extraction and classification. Face detection has been done using the Haar Cascade frontal face algorithm in the initial stage. Aswathy, Ashok [10] built a system for helping the low vision people to recognize the emotions of the confronting person and hence enabling them to have better communication. There are five steps followed for building the model. They are Image acquisition, face detection, feature extraction, emotion recognition and express the emotion as an audio to the blind person.

The captured images are checked for verifying whether a person is present in the images using Haar features[27]. Once a person is detected the feature extraction is performed. The feature points are used to extract the patches in the images. A database is used to store the facial points, similar and distinct images. Using the Support Vector Machine (SVM) algorithm [17], the emotion of the person is recognized. An audio device is used for conveying the recognized emotion as a speech to the blind person. The deep learning model will use a neural network that performs all the visual tasks and classify the emotion as a whole. The dataset they used was Japanese Female Face Expression (JAFFE) which has facial expressions of only females. Since it uses SVM, it is not very accurate. It is a static system like local PC or laptop that cannot be carried by the visually challenged person everywhere they travel. In order to overcome these issues, we created a model that uses convolutional neural network for classifying emotions. The system is modified to use a portable device like Raspberry pi [15], so that the visually impaired person can carry it along with them.

III. PROPOSED SYSTEM

A Raspberry-pi [24] device is connected with a camera and headphones. The emotion recognition model is trained using TensorFlow framework. Fig.1 demonstrates the overview of the system. The camera sends input images captured to the computer. The image is then used by the Haar Cascade files for detecting the human face in the image. The convolutional neural network is used for recognizing the emotions. Once the

emotion is classified, it is converted into an audio. The text-to-speech can be done using the [25] Pyttsx3 package. The resulting audio can be heard by the user via headphones. Raspberry pi can be used for this purpose. It can process machine learning models programmed through Python. The Raspberry Pi [19] is a tiny sized, low cost computer that functions similar to a CPU. The keyboard and mouse are plugged to it for providing inputs.

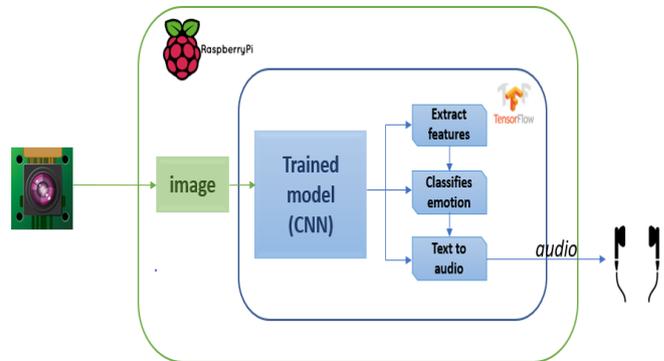


Fig.1. Proposed System Architecture

For displaying the outputs it plugs into a computer monitor or TV through a HDMI port. It can be programmed using Python, Java, C++, Scratch or Ruby. It uses SD cards as hard drives. It requires a low power supply of about 3 Watts for its working.

This paper demonstrates a real-time application for emotion recognition by using a Raspberry Pi 3 board with a camera module. The camera module and Raspberry Pi board can be interfaced with each other and extracts information related to input images by using a computer vision library [13]. The camera captures the images of the person who is communicating with the visually impaired person. The frames of images are individually analyzed for detecting a human face. The human face can be detected by [27] Haar Cascade files. The detected face is processed to get the exact emotion through a convolutional neural network (CNN).

The modulus of this proposed work are :

- Data Preprocessing
- Model Training
- Testing Phase
- Text-to -speech Conversion.

A. Data preprocessing

Data processing is the process of converting raw data into cleaned data. Fig.2 demonstrates the following steps:

1. The Fer2013 dataset is downloaded from kaggle and unzipped. The Fer2013 dataset has more than 30,000 images covering almost all the facial expressions of humans.
2. In the raw data the missing values are filled using sklearn package and noise are removed. This process is called data cleaning.
3. The root folder of Fer2013 package is added with the files such as fer2013.csv [28] and shape_predictor_68_face_landmarks.dat.
4. The dataset is then converted into extracted face landmarks and HOG(Histogram of Oriented Gradients) images[20].

5. Then one hot encoding is performed by mapping each emotion to a number, for example happy = 0 , neutral = 1, sad = 2 and so on.

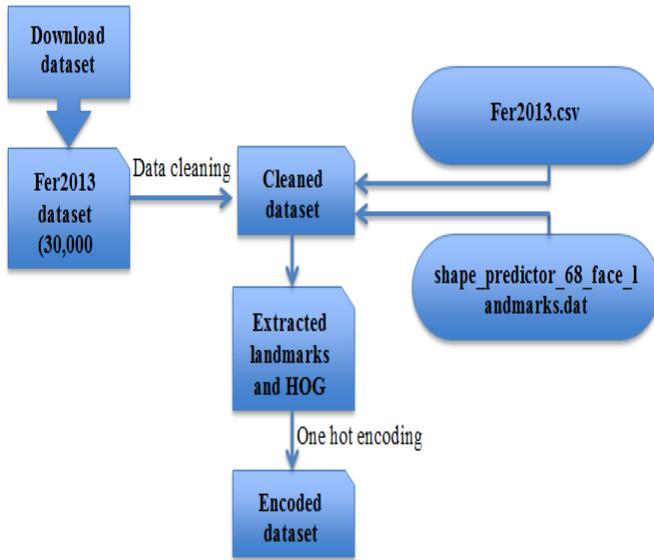


Fig.2. Data Preprocessing

B. Model training

Model training is simply, the learning the properties and behavior of the particular object from labeled examples. In this work, once the dataset is preprocessed and organized. The model uses Convolutional Neural Network (CNN) [22] which interconnects the various weights on the nodes representing the different emotions. The model is trained with the help of TensorFlow background. To compute the CNN model, that needed to include all the necessary modules for TensorFlow and the data set modules. In this work the dataset are collected in csv file .That csv file can be given to the Tensorflow and the CNN model for training, which is shown in fig.3.Finally a trained model that can recognize emotions is obtained.

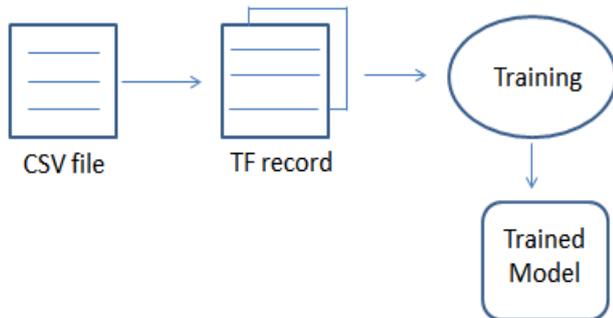


Fig.3. Model Training

C. Testing phase

The system’s local camera is used for testing in real time. The raspberry-pi camera module is activated. The frames of images that are captured and processed against the trained model. Initially the human faces are detected through the Haarcascade Frontal face Xml file. It is a pre-trained model for detecting human faces.

The model will compare the input images with the values that exactly match with already trained data. Finally the emotion that has high accuracy is returned as output. Figure 3 shows the overall testing process.

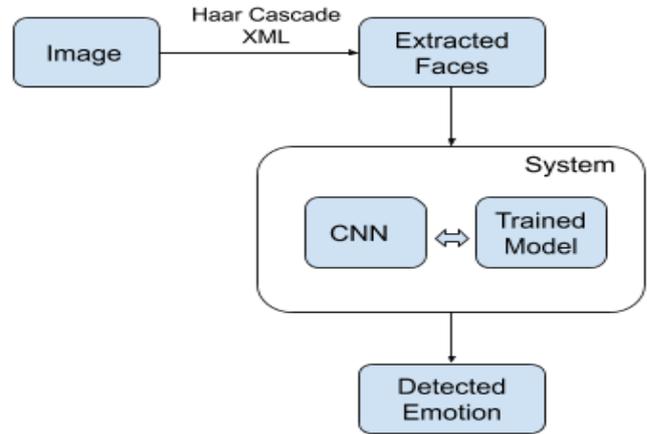


Fig.4. Testing Process

D. Text-to-Speech Conversion

The resulting emotion is converted to a speech using the Pytttx python package which is shown in Fig.5.Pytttx [25] is a cross platform that can support multiple operating systems like Mac OS, Linux, Windows. The corresponding driver for the OS is used to initiate the speech conversion. In this work , the Ubuntu operating system is used. The eSpeak engine is used as the driver for Ubuntu. An audio device attached to the system. The audio is then conveyed to the person through the audio device like Earphones.

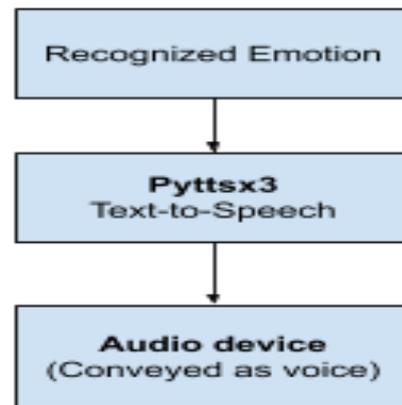


Fig.5. Workflow of Text-to-Speech conversion

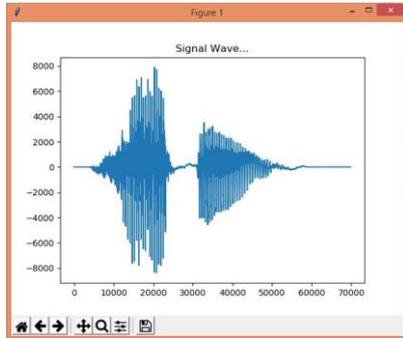
IV. EXPERIMENTAL RESULTS

The system uses Convolutional Neural Network which is more accurate than the Support Vector Machine. Moreover all kind of emotions including happy, sad, angry, fear, surprise, neutral and disgust are recognized perfectly

The system has recognized the emotion for happy and neutral and conveyed through voice. Fig.6 and Fig.7 shows the pictorial representation of the emotion recognized and signal wave of voice.

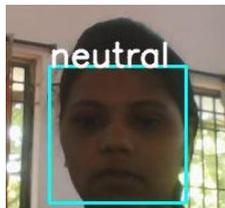


Recognized
Emotion

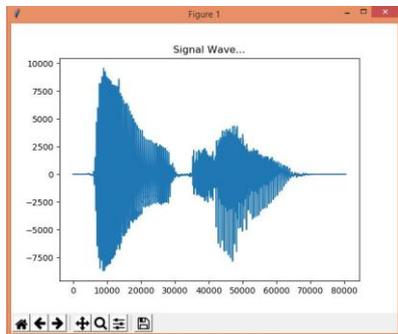


Voice
Graph

Fig. 6. Experimental result for HAPPY



Recognized
Emotion



Voice
Graph

Fig. 7. Experimental result for NEUTRAL

V. CONCLUSION

The proposed system will be built with all the above mentioned functionalities including Convolutional Neural Network. The CNN works better than a shallow network and provides better accuracy. This helps the visually impaired to recognize the Facial expressions of the people communicating with them. From this, the low vision people can have better social interactions and ease of communication. In future the system will be enhanced with face recognition and object identification technologies.

REFERENCES

1. Tuhin Kundu, Chandran Saravanan, "Advancements and recent trends in Emotion Recognition using facial image analysis and machine learning models", 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT).

2. T Anuradha, S Monika, S V S Prasad, "Embedded Video Processing on Raspberry Pi", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-4S2 March, 2019
3. Ma Xiaoxi, Lin Weisi, "Facial Emotion Recognition", 2017 IEEE 2nd International Conference on Signal and Image Processing.
4. Nithya Roopa.S, Emotion Recognition from Facial Expression using Deep Learning, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8 Issue-6S, August 2019 .
5. Prof. V.D. Bharate, Shubham Sunil Phadatare, Suhas Panchbhai, Vishal Pawar, "Emotion Detection using Raspberry Pi", International Research Journal of Engineering and Technology (IRJET).
6. Ketki R. Kulkarni, "Facial Expression Recognition", 2015 International Conference on Information Processing (ICIP)
7. Minaee, Shervin & Abdolrashidi, Amirali. "Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network", ArXiv 2019.
8. Dr.Priya Gupta , Nidhi Saxena, Meetika Sharma, Deep "Neural Network for Human Face Recognition", Published by MECS Publisher.
9. T. Guo, J. Dong, H. Li and Y. Gao, "Simple convolutional neural network on image classification," 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA).
10. Ashok, Aswathy & John, Jisha, "Facial Expression Recognition System for Visually Impaired", International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018.
11. Hsi-Chieh Lee, Chia-Ying Wu and Tzu-Miao Lin, Facial Expression Recognition Using Image Processing Techniques and Neural Networks, J.-S. Pan et al. (Eds.): Advances in Intelligent Systems & Applications, SIST 21, pp. 259–267.
12. Omkar Pawar , Prathamesh Lonkar , Randhir Singh, Vivek Salunke , Prof. D.M. Ujlambkar , "Door Lock System using Facial Recognition", International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: Volume 7 Issue III, Mar 2019
13. <https://www.techopedia.com/definition/32309/computer-vision>
14. <https://expertsystem.com/machine-learning-definition/>
15. <https://www.raspberrypi.org/>
16. <https://medium.com/themlblog/how-to-do-facial-emotion-recognition-using-a-cnn-b7bbae79cd8f>
17. <https://medium.com/machine-learning-101/chapter-2-svm-support-vector-machine-theory-f0812effc72>
18. <https://elitedatascience.com/model-training>
19. <https://pythonprogramming.net/introduction-raspberry-pi-tutorials/>
20. <https://towardsdatascience.com/feature-extraction-techniques-d619b56e31be>
21. <https://www.simplilearn.com/data-preprocessing-tutorial>
22. <https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148>
23. https://docs.opencv.org/master/d9/df8/tutorial_root.html
24. <https://projects.raspberrypi.org/en/projects/getting-started-with-pi-camera>
25. <https://pypi.org/project/pytsx3/2.7>
26. <https://medium.com/the-andela-way/simple-operations-on-images-using-opencv-d37b26e6e3ab>
27. <https://becominghuman.ai/face-detection-using-opencv-with-haar-cascade-classifiers-941dbb25177>
28. Christopher Pramerdorfer, Martin Kampel, "Facial Expression Recognition using Convolutional Neural Networks: State of the Art"

AUTHORS PROFILE



Dr. E. Kodhai is currently working as Associate Professor in the Department of Computer Science and Engineering at Sri Manakula Vinayagar Engineering College affiliated to Pondicherry University, Puducherry, India. She has completed her M.C.A from Cauvery College for women, Trichy affiliated to Bharathidasan

University, Trichy, M.E. in Computer Science and Engineering from Vinayaka Mission's Kirupananda Variyar Engineering College, Salem and Ph.D in Computer Science and Engineering from Pondicherry Engineering College, Puducherry. She has 19 years of experience in teaching in various engineering colleges. Her Research interests include Software Clones and Software Engineering. She has

published more than 50 papers in international conference and journals.



A. Pooveswari is pursuing her Undergraduate Degree in Sri Manakula Vinayagar Engineering College, Puducherry in the field of Computer Science and Engineering. She is interested in the research works on Machine Learning and Software Testing. Worked in Edge Computing where the product billing system is created for a shopping mall using machine learning algorithms. Also worked in automation testing using the Selenium tool.



P. Sharmila is pursuing Bachelor of Technology in Sri Manakula Vinayagar Engineering College, Puducherry. She is from the Computer Science and Engineering department. Her current research interests are in the field of Data science and Machine learning. Worked in a face recognition project where the age, emotio, and ethnicity of the people are identified. Also done projects on anomaly detection where the abnormal activities taking place in a pedestrian walkway are recognized.



N. Ramiya is pursuing Bachelor of Technology in the field of Computer Science & Engineering in Sri Manakula Vinayagar Engineering College, Puducherry. Her research interest is Computer vision. Worked o machine learning projects like diabetes analyzing and weather forecasting. Currently working on computer vision projects like face verification and object recognition.