ISSN: 0974-5823 Vol. 7 No. 1 January, 2022

*International Journal of Mechanical Engineering*

# Relevance of Thermal Imaging and Respiration Signals in Recognizing Human Emotions

C. M. Naveen Kumar[1], G. Shivakumar[2]

[1]*Assistant Professor, Department of E & I Engineering, Malnad College of Engineering, Hassan, affiliated to VTU, Belagavi, India*

[2] *Professor, Department of E & I Engineering, Malnad College of Engineering, Hassan, affiliated to VTU, Belagavi, India*

*Abstract - Emotions plays a vital role in almost every phase of the human life. Emotion Recognition grabbed researchers' attention especially from last two decades as it has impact in several areas like healthcare, polygraph, and crime detection etc. In recent years, physiological signals and thermal imaging features are widely used to analyse their correlation between subjective emotions. This chapter discusses the novel approach towards primary emotion classification where features of thermal images are combined with Electrocardiogram and Respiration signals. Empirical Mode Decomposition is used to extract the features from Electrocardiogram and Respiration signals (RS). Gray Level Co-occurrence Matrix (GLCM) approach is used for thermal images to extract features like energy, correlation, homogeneity, and contrast. Physiological data and Thermal Images are obtained from the subjects who are undergraduate students. Fusing the features at feature level improvises the recognition rate of the system as bimodal approach dominates unimodal approach.*

*Keywords:* Emotions, Electrocardiograph, Respiration Signal, Empirical Mode Decomposition, Grey Level Co-occurrence Matrix.

## I. Introduction

This document is a template. An electronic copy can be downloaded from The Research Publication website. For questions on paper guidelines, please contact the Research Publication as indicated on the journal website. Before submitting your paper, check that the format conforms to this template. Specifically, check the appearance of the title and author block, the appearance of section headings, document margins, column width, column spacing and other features.

Emotions play an importance role in human's daily life as it directs the human behavior and thoughts. Studies on Human emotion suggest that at least one emotion influence most of the life experience. Recent advancements in neurological research concluded that Emotional brain dominates rational brain because every decision is having the influence of an emotional state. Rational brain and emotional brain work together to respond for incoming stimuli from outside world. Amygdala is an important part of emotional brain reacts to the emotional input. Both rational and emotional brain is connected together with huge number of neurons. Emotion Recognition gained significant importance in the fourth Industrial Revolution especially for developing Artificial Intelligent Agents. Generalized Emotion classification system need to be developed which has its impact in industrial applications as well as in the society. Human emotions cannot be detected directly but they are related to internal factors of the human body like physiological signals , speech signals, physical gestures, facial expressions and also thermal signatures of the face[1].

Emotions can affect the way a person reacts to a situation, the way he talks, the way he talks and many other parameters. To deduce an emotional state, there are many variables that an experimenter can sense and link the patterns extracted to affective computing. Recognition of facial expression, vocal inflection and text inputs are normally used in affective computing. The definition of the emotions is not accepted by most of the Emotion theorists. Normally emotions are designated as elementary discrete classes and are represented as locations in two dimensions namely 'Arousal' and 'Valence'. The arousal dimension refers to the activation or excitement of the emotion and valence dimension refers to the criteria of intensity and polarity of pleasing or displeasing of the emotion [3].

Developing a genuine data set is highly challenging task as it involves breaching privacy and ethics as genuine emotions are elicited in the situation of personal importance and in the real world that is capturing the elicitation in a hidden recording. Attaching recording devices or sensors make the subject's aware about their expressions so efforts need to be done in the area of devices that can capture the data in a hidden fashion [2]. The data recording should consider both external expression and internal feelings because they are interrelated and overlapped as one can improve other. There are two approaches for emotion recognition they are: (i) person dependent where data is acquired for analysis by a single subject over different time interval. (ii) person independent where data is acquired in a short duration from various subjects. The main advantage with the first approach is possibilities of getting same sort of interpretation for a label. The disadvantage is each subject's interpretation and pattern of emotion differs hence arriving at generalized approach is difficult.

The sensors and tools which are integrated in wearable such as watches, shoes, belts etc. can be source of data acquisition systems to sense affective patterns. These systems are advantageous as it can sense the signals and understand underlying affective patterns and infer the decisions to the wearer about their health issues related to stressing or negative emotions, liking or disliking about certain things [4].

Thermal Imaging has come into the front end of the research on affective computing. It is non-invasive and contactless technique for the evaluation of various emotional states. Basically, temperature maps in the facial thermal images have shown significant correlation with the emotional labels. Visible image-based technique is affected by illumination; skin color of the subject, brightness etc. but thermal image-based technique overcomes all these and can be implemented round the clock. Various emotion labels bring changes in the flow of blood which causes heat distribution signature on the face. These signatures and its topographic distribution have proved their relationship with all affective states [5].

## II. RELATED WORK

Human Robot Interaction (HRI) and Human Computer Interaction (HCI) have gained focus of many on-going technological research developments. Emotional Intelligence has significant importance as a part of Social intelligence which is the basic necessity in HRI and HCI especially in the applications which demands collaboration of humans with machine. The main target of automatic affective states recognition is to enhance effective interaction of humans with a machine so that the machine can better understand humans and reflect the same with convincing actions. The recent research on Affective computing provides multiple modalities through which emotion recognition can be done those include the approaches which are non-invasive and invasive [7].

In recognizing human emotions, physiological data plays an important role. To acquire the data, experimenter should know what is correct data? How the data need to be acquired? The sensor or electrode placement, motion artifacts affect the quality of the data. The main hurdle is to get the data corresponding to emotional state and to validate the label of the emotion for the data acquired. Identifying a real label for the data can be challenging which depends on mood and situation of the subjects and also the subject's rapport with the experimenter. If the facial expression or voice is of interest, experimenter can conclude the label if the interest is physiological signal then subject should validate the label as it is less influenced by cognitive and social issues. Emotion elicitation depends on five factors viz (a) Emotion elicited by subject or it is elicited by external event. (b) Location of the data acquisition like laboratory or real world. (c) Whether importance is given to external expression or internal feeling. (d) Data is recorded by the subject knowingly or unknowingly. (e) The awareness of purpose of the study [2][3].

Earlier studies focused on facial expression-based recognition because it clearly highlighted the aspect of valence dimension as positive and negative. Later it was combined with vocal parameters because it clearly highlighted the arousal dimension of emotion. Emotional Intelligence plays an important role in human's day to day interactions as it builds his/her personality in society. Emotional skills are important part social skills and an individual need to possess these. An automated emotion recognition ability of a computing system supplements any human being to know and analyze his behavior and to modify the same in order to build effective personality. The main requirement for a system with this ability is to learn from the positive and negative feedback of the user so that the user should not feel frustrated with the suggestions of such systems like a stress pill notification for a person under pressure [3].

Most of the Facial feature based research works gave emphasis mainly on the variety of features from the entire face, but now the research is diverting towards features from individual regions of the face. In [9], features from various regions of the face are considered which is a novel approach to improve the results when compared to recognition based on features from each region in a face. In this work cross validation of the results is done through comparison of the results from one region with other region with the help of majority voting method. The method based on multiple regions is a burden on computation but the performance of the system for accuracy is much better. It is necessary to recognise emotions from face especially during post pandemic for people wearing masks on their face.

The sensors need to be robust for motion artefacts and must be comfortable for the users or the subjects to wear. Any sort of irritation of sensors will make the users conscious about the situation and natural emotion elicitation will be distracted [4].

Empirical Mode Decomposition (EMD) decomposes frequency-modulated and amplitude-modulated waveforms using non-stationary and non-linear time series. Here Wrist Pulse Signal is decomposed into number of Intrinsic-Mode functions (IMFs). The extraction of non-linear and statistical features is done using EMD to classify various emotional states. Both EMD and Hilbert Transform are used for emotion recognition. The decomposition technique employed here decomposes signal adaptively making it a state-of-the art method. Totally nine IMFs were considered here individually and in combination for the qualitative analysis of the WPS. Each IMF possesses vital information about the individual emotion label, in this analysis fourth and fifth IMFs are correlated to boredom and anxiety [8].

The emotions elicited from patients with Autism Spectrum Disorder (ASD) uses features like RR interval, various frequency components decomposed using Discrete Wavelet Transform is used in [10]. Elicitation of emotions using video stimuli use to elicit multiple emotions rather than single emotion is a challenging task and need to be addressed as it may create misclassification of emotions.

A Wavelet Transform based emotion detection for thermal images yields high accuracy [11]. The validation of the results is done effectively using feedback from a therapist who knows about the true label of emotion as subjects are patients with ASD. The negative valence emotions can be well detected using non-contact and non-invasive based systems.

## III. METHODOLOGY

The flow diagram for the proposed Human Emotion Recognition System is displayed in the figure 1.

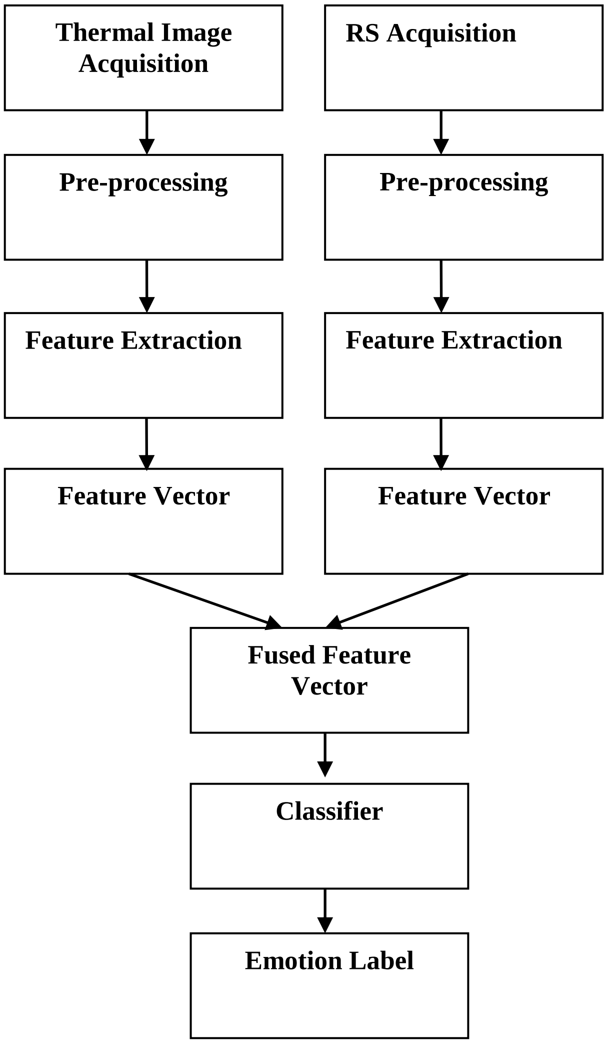


Fig. 1 Proposed system flow diagram

*A. DATA ACQUISITION:*


Fig. 2 Thermal Image depicting the Data Acquisition from subjects.

The data collection for the task like emotion recognition is a herculean task. Here data for a particular emotion is collected by eliciting emotions from the subjects through an external event that is by playing three minutes videos. The videos for the purpose are taken from www.youtube.com. Those videos are selected by conducting a formal survey in Hassan, Karnataka by investigating several subjects from various age groups. The Subjects use to watch the video and enter their feedback about the video like the label of the emotion for which the video belongs to and also the intensity of the particular emotion label experienced by that video [6].

The data for the experimental analysis is acquired by playing the selected videos from the survey so that the subjects start reacting to the videos by displaying the particular emotion label. The reaction to the videos makes the modification in respiration signal and distribution of temperature in the face. The modification in the respiration signal is measured using piezoelectric transducer embedded in the Velcro belt.

The Thermal Imager FLUKE TiS20 is mounted on a tripod and subjects are allowed to sit in front of the imager which is placed at 3 feet distance. The imager has auto capture mode which is used to capture the images.

This signal is acquired using a Velcro belt which contains piezoelectric transducer. Piezoelectric transducer is placed on the chest so it senses the expansion and contraction of chest which in turn catches respiration signal. The transducer is interfaced with a smart microcontroller. The code for respiration data acquisition is in in MATLAB 2018a. Data is stored in notepad after acquiring the data.

*B. PRE-PROCESSING:*

The signal acquisition through the piezoelectric transducer has certain limitations like degradation of the respiration signal due to motion artefacts and error in the measurement due to the displacement of the transducer due to compression and expansion of the chest. A high pass filtering is used to remove high frequency noise. The desired signal is dominated by the frequency components related to artefacts which may occur due to the activity of the muscular tissue. The filtering through Median filter provides the more convincing data that is not impacted very much by sample which is not related in the neighbourhood [6]. The below figure 2a and 2b shows signal before and after pre-processing.

A Gaussian filtering is used to smooth the thermal image. The further processing of the image is done by deciding the ROI. The subjects face is the ROI that is manually cropped from the thermal image. [6]

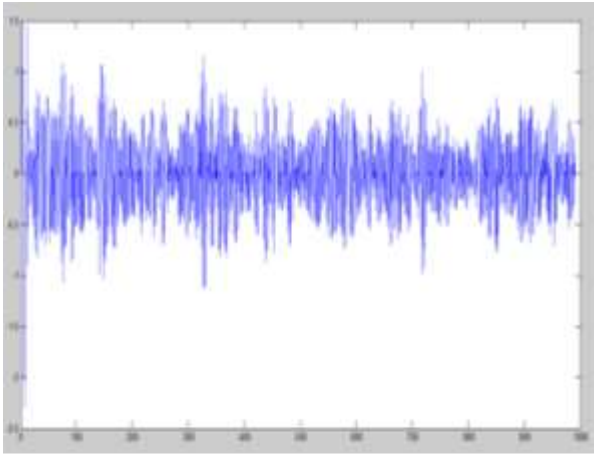

Figure 2 a: Acquired Respiration signal before pre-processing.

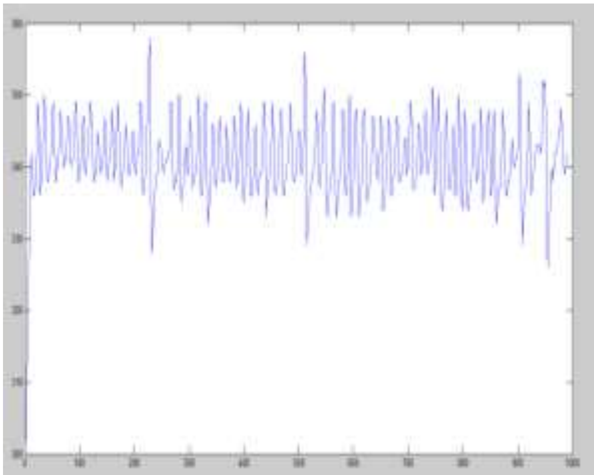

Figure 2b: Acquired Respiration signal after pre-processing.

*C. FEATURE EXTRACTION & SELECTION:*

Empirical Mode Decomposition (EMD) is implemented for the respiration signal where the first two Intrinsic Mode Function (IMF) were included likeIMF1 and IMF2. This is because of the shortcoming in the bandwidth of the respiration. The IMF1 and IMF2 are shown in figure 3a and3b.

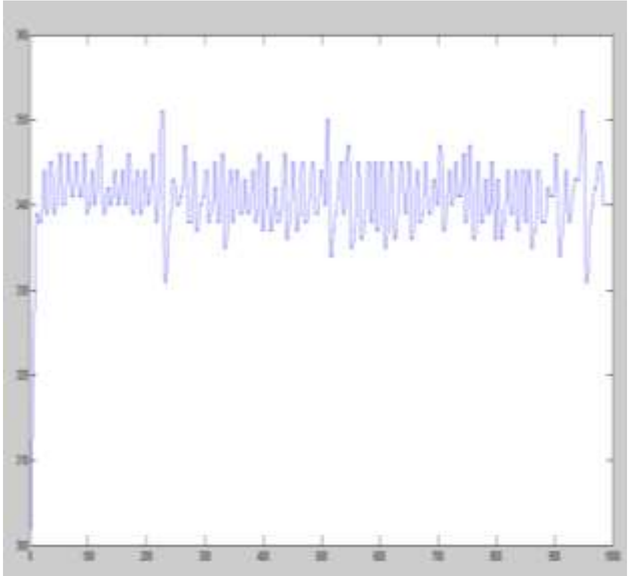

Fig 3a: Intrinsic Mode Function 1 Respiration signal

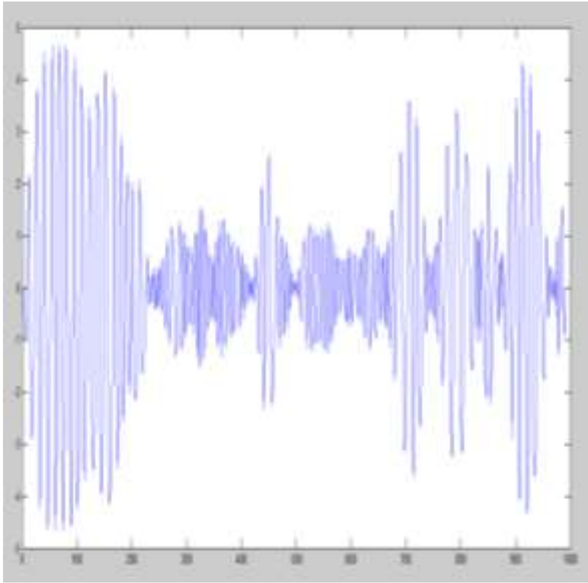

Fig 3b: Intrinsic Mode Function 2 Respiration signal

The Mean, Maximum and RMS value from the first two IMFs were calculated. So finally six features were considered, two from each.

The thermal images were statistically analyzed by performing Gray Level Co-occurrence Matrix (GLCM). The second order statistics of a thermal image is obtained by the GLCM that calculates the textures in the image which reveals the pixels' spatial connection. This is viewed as the histogram in the two dimensional gray level. GLCM provides the occurrence probability among pixel pairs; this computation uses pixel direction and distance. The particular relationship between pair of pixels and the particular value possessed by GLCM image forms are included while classifying image surfaces that are basically dependent on functions derived from GLCM.[6][7]. Sixteen matrices in the order of 4 by 4 are considered for every thermal image. Each matrix is used to derive four features. The following four features are considered here:

**Contrast:** The local variations in the GLCM are calculated using contrast measure. This computation is done by considering the whole image that returns a value which reflects the difference in the intensity with pixels and their neighbors.

$$\sum_{i,j} |i-j|^2 p(i,j)$$

Range = [0 - (size (GLCM, 1)-1)^2]

**Correlation:** A particular pixel pairs are considered to calculate joint probability of the occurrence. This computation yields a vaue that reflects the correlation of a pixel with its neighbors for the whole image.

$$\sum_{i,j} \frac{((i-\mu_i)(j-\mu_j)p(i,j))}{\sigma_i \sigma_j}$$

Range = [-1 1]

**Energy:** Energy parameter gauges the uniformity in an image by computing sum of square of all the pixels which is in a range between zero and one.

$$\sum_{i,j} p(i,j)^2$$

**Homogeneity:** Homogeneity provides the information about the closeness of the element distribution in a GLCM with respect to its diagonal and a value is returned which is in a range between zero and one.

$$\sum_{i,j} \frac{p(i,j)}{1+|i-j|}$$

*D. DATA FUSION:*

A total of 64 features were derived from thermal IR image and six features from respiration signal. The features from two domains are derived independently and are combined at this stage. Fusion results in total of 70 features.

*E. CLASSIFICATION:*

The data corresponding to features from Respiration signal and thermal Image that is relevant to a particular emotion label are utilized for the purpose of classification. The actual affective state is detected first by training the classifier with combined data of the selected features. A non-parametric technique based classifier is used which provides class membership as its output that is K-Nearest Neighbour (KNN) decision principle. The majority in the votes by the neighbours for a specific class that is most common among its k-nearest neighbour. The selected decision rule is perfect if the data input is huge and it is also robust for the noisy data. At this stage, combined data corresponding to selected features converges into six classes of emotions.

## IV. RESULTS & DISCUSSIONS

The analysis of the proposed system is done by considering Recognition rate for all the primary emotions. The Recognition rate which is displayed in the below table I is calculated by considering only Respiration signal and secondly by combining Respiration signal and thermal IR image.

TABLE 1. PRIMARY EMOTIONS RECOGNITION RATE

| | **Recognition Rate** | | | |
|---|---|---|---|---|
| | **ECG [6]** | **Respiration Rate Signal** | **Thermal Images + ECG [6]** | **Thermal Images + Respiration Rate Signal** |
| **Happy** | 61.90% | 61.3% | 85.71% | 85.71% |
| **Sad** | 57.14% | 61.9% | 80.95 | 85.71% |
| **Disgust** | 66.66% | 66.66% | 90.41% | 90.47% |
| **Anger** | 52.38% | 57.14% | 80.95% | 76.19% |
| **Fear** | 66.66% | 66.66% | 85.71% | 80.9% |
| **Surprise** | 52.3% | 52.3% | 90.41% | 76.1% |

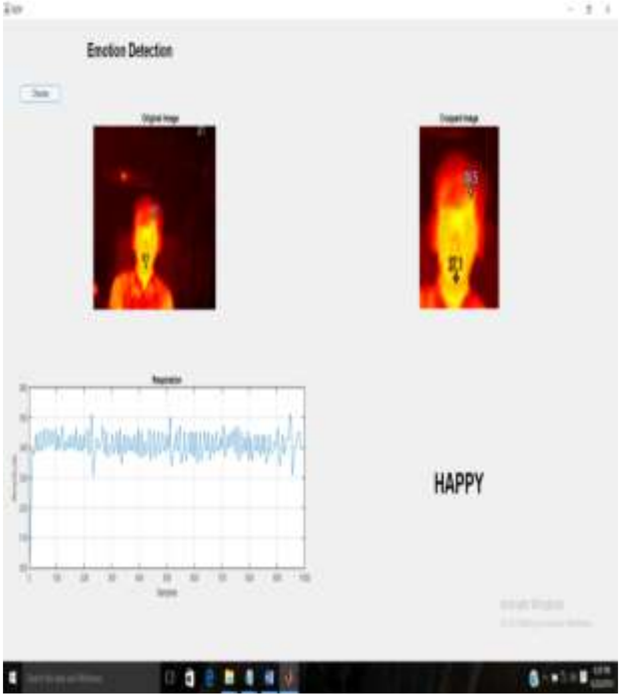


Fig 5.2.1: Screenshot of the output for Happy

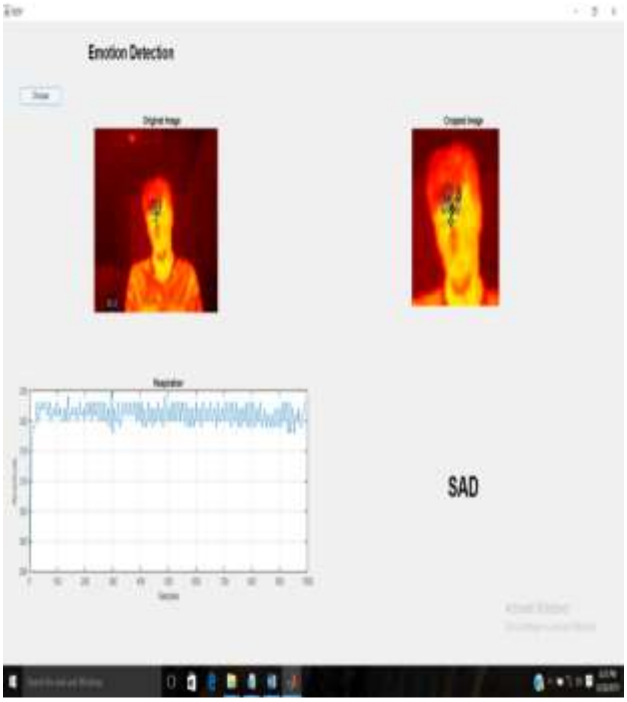


Fig 5.2.2: Screenshot of the output for Sad

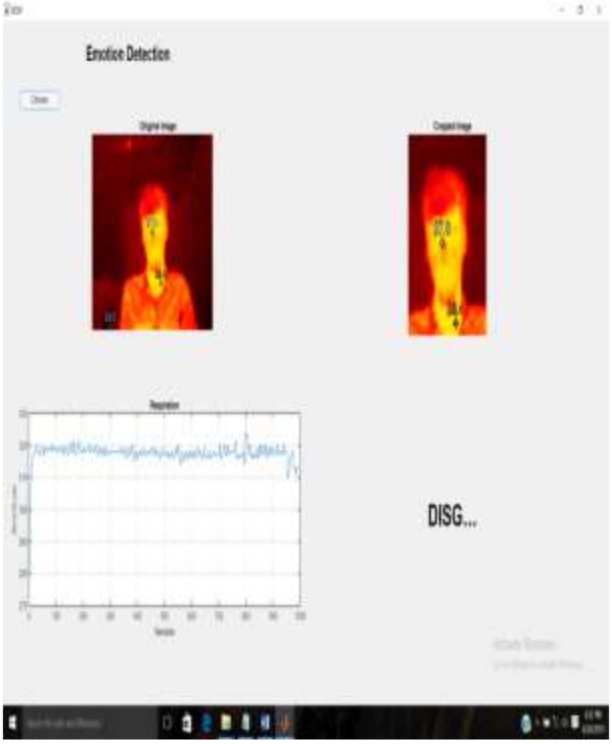


Fig 5.2.3: Screenshot of the output for Disgust

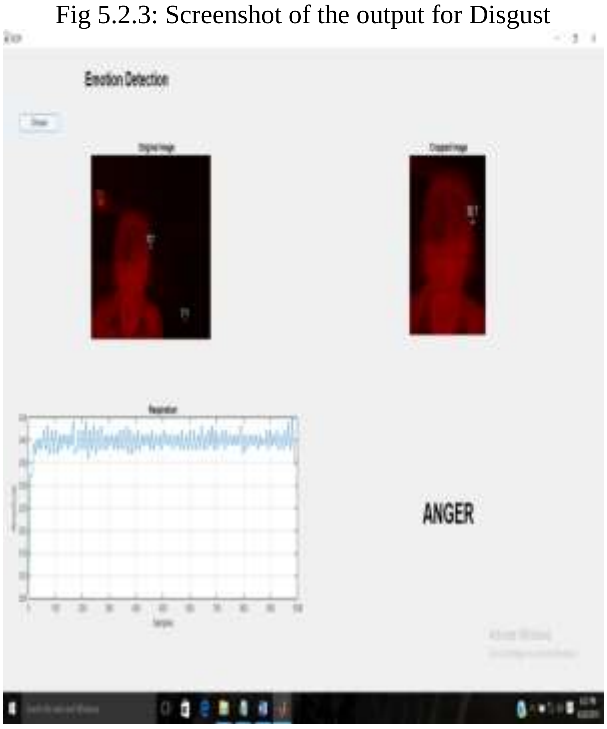


Fig 5.2.4: Screenshot of the output for Anger

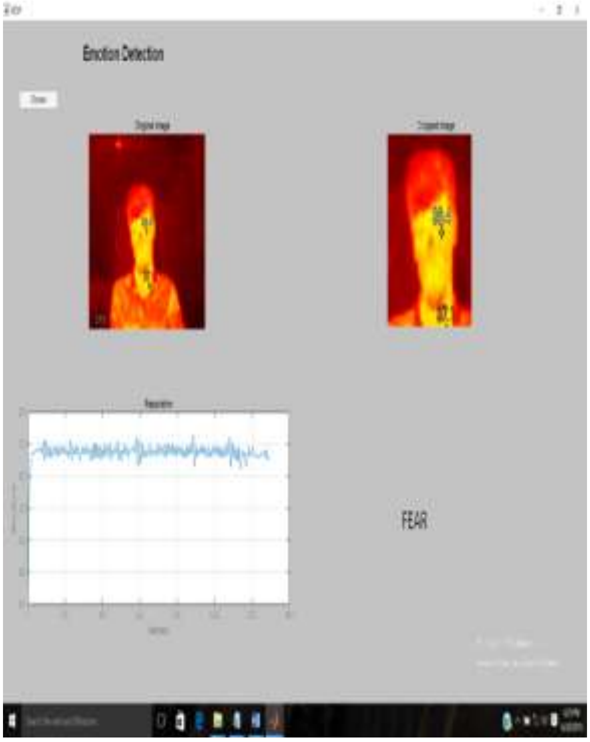


Fig 5.2.5: Screenshot of the output for Fear

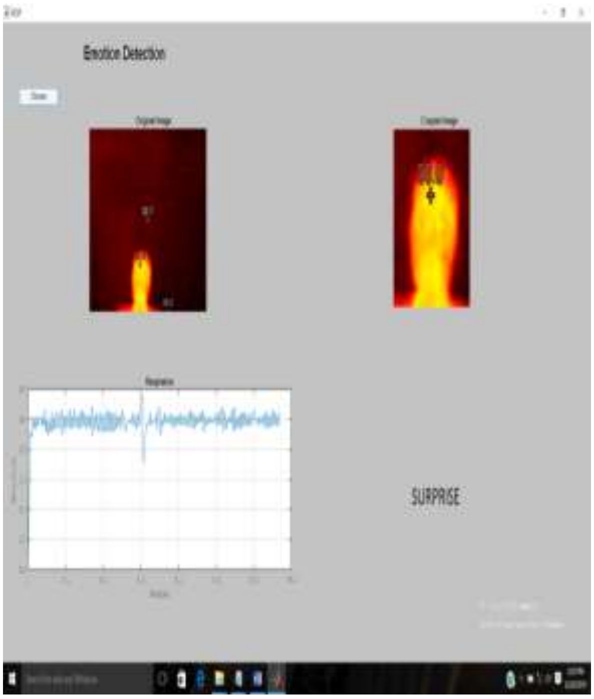


Fig 5.2.5: Screenshot of the output for Surprise

## V. CONCLUSIONS

The emotion recognizer proposed here fuses features from the respiration signal and thermal IR images. 60.99% is the average recognition rate by considering only respiration signal and 82.51% is the average recognition rate by combining features of respiration signal with thermal IR image which clearly indicates the dominance of the bimodal approach over unimodal approach. The drawbacks with this emotion recognizer are the absence of automatic face detection feature and smaller database which is suggested for future implementations or modifications to the proposed work [9]. Significant change in the thermal values in the face is mainly in the nose, forehead, cheek and chin. The intra-person variability can be considered by considering multiple images and physiological data from the same person by playing different videos corresponding to particular emotional label of same intensity at different point of time. Thermal image with higher resolution certainly contains more data and hence reduces the error in the implementation and also highlights clear thermal variability [6].

**Acknowledgments**
We like to thank our Institution Malnad College of Engineering, Hassan for the encouragement and support in the work.

REFERENCES

[1] T. Roy, et al. "Advancements and Role of Emotion Recognition in the 4th Industrial Revolution.", In: book The Disruptive Fourth Industrial Revolution. Springer, Cham, Vol. 674 pp.179-203, 2020.
[2] Picard et al., "Towards machine emotional intelligence: analysis of affective physical state." IEEE Trans Pattern Analysis and Machine Intelligence (2001): 119-175.
[3] Picard, Rosalind W. "Building HAL: Computers that sense, recognize, and respond to human emotion." Human Vision and Electronic Imaging VI. Vol. 4299. International Society for Optics and Photonics, 2001.
[4] Picard, Rosalind W., and Jennifer Healey. "Affective wearables." Personal Technologies 1.4 (1997): 231-240.
[5] Priya, M. S., and Kadhar Nawaz GM. "Modified emotion recognition system to study the emotion cues through thermal facial analysis.", Biomedical Research, Vol. 28, Issue 20, 2017.
[6] C M Naveen Kumar, G Shivakumar, "Sensor and Feature Level Fusion of Thermal Image and ECG Signals in Recognizing Human Emotions", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Vol. 9 Issue-2S, December 2019.
[7] Latif M. H., et al. "Emotion detection from thermal facial imprint based on GLCM features." ARPN J. Eng. Appl. Sci 11 (2016): 345-350.
[8] Garg, Nidhi, and Gurpreet Kaur. "Exploring Wrist Pulse Signals using Empirical Mode Decomposition: Emotions." IOP Conference Series: Materials Science and Engineering. Vol. 1033. No. 1. IOP Publishing, 2021.
[9] Agrawal, Ekanshi, Jabez Christopher, and Vasan Arunachalam. "Emotion Recognition through Voting on Expressions in Multiple Facial Regions.", Vol. 2, pages 1038-1045, ISBN: 978-989-758-484-8 2021.
[10] SSindhu, N., and S. Jerritta. "Evaluation of Emotion Elicitation for Patients With Autistic Spectrum Disorder Combined With Cerebral Palsy." Advances in Electrical and Computer Technologies. Springer, Singapore, Vol. 672, pp. 737-750, 2020.
[11] Rusli, Nazreen, et al. "Implementation of Wavelet Analysis on Thermal Images for Affective States Recognition of Children With Autism Spectrum Disorder." IEEE Access 8, pp. 120818-120834, 2020.