2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON) | 978-1-6654-9382-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/REEDCON57544.2023.10151406

Human Emotion Recognition Models Using Machine Learning Techniques

Aftab Alam
Department of Electrical Engineering
Jamia Millia Islamia
New Delhi, India
aftabrzaattari@gmail.com

Shabana Urooj
Department of Electrical Engineering
College of Engineering, Princess Nourah
Bint Abdulrahman University, Riyadh-KSA
shabanabilal@gmail.com

Abdul Quaiyum Ansari
Department of Electrical Engineering
Jamia Millia Islamia
New Delhi, India
aqansari@jmi.ac.in

***Abstract*— Researchers have always been curious if a computer can detect human emotions precisely and accurately. Many research publications have been reported on human-machine interaction systems. The emotion classifiers using machine learning techniques are developed using the feature dataset extracted from physiological and non-physiological parameters. Emotion recognition can be done either by using facial, speech or audio-visual data paths or using physiological signals like ECG, EEG, EMG, GSR and Respiration signals. Many have explored facial recognition techniques for emotion recognition but facial expressions can be masked. A sad person can pretend to have a smiling face and vice-versa. Physiological signals like ECG, EEG, GSR and respiration signals are non-maskable due to their involuntary source of generation. There are many datasets available publicly for researchers to use and develop an efficient emotion classifier system. In this work, the publicly available datasets of EEG, ECG and GSR recorded while watching emotional video are utilized to develop emotion classifiers using machine learning techniques. Here three physiological feature datasets named LUMED-2 (EEG+ GSR), SWELL (HRV), and YAAD (ECG+ GSR) are used to train models and classify emotions. The deep learning classifiers used are Random Forest, SVM, KNN, and/or Decision Tree. The maximum average classification accuracy achieved is close to 100% at least for one classifier in each dataset. It is observed that physiological signals like EEG, ECG, and GSR possess differentiable emotional features which can be used to detect the emotional state of a person precisely using the trained machine learning models.**

Keywords*— Human-Computer Interaction, Electrocardiogram (ECG), Electroencephalography (EEG), Galvanic Skin Response (GSR), Emotion Recognition, Machine Learning Techniques.

I. Introduction

Emotion recognition is an important aspect in the development of intelligent human-computer interaction systems like chatbots, AI-based assistants, smart home appliances etc. Human emotion recognition systems utilize physiological and non-physiological signal features to classify the emotional state. Non-physiological parameters involve features of speech, facial expression, and audio and video signals of the subject. Such parameters can be masked because of their voluntary expression and control of the subject. Moreover, it has variability with subjects depending on their extent of expressive nature. The physiological parameters like ECG, EEG and GSR are involuntary in origin, hence can't be masked or controlled by the subject. The features extracted from EEG, ECG and GSR signals have the potential to classify emotions more precisely. Such systems can be used to identify the emotional states of a person, hence it can be used to diagnose mental illness and required preventive care can be ensured. It can help us to understand the specific orientation and interest of a person towards a particular field in career counselling, to automate smart appliances in homes as per the mood of the owner, to analyze the shopping preferences of customers and even for lie detection systems in criminal investigation procedures.

Many physiological and non-physiological signals-based unimodal and multimodal emotion classifier systems have been reported by researchers working in the field. Taiba Majid Wani *et al.*, have reviewed speech recognition-based techniques and tabulated publicly available speech databases. A comparison table after an extensive review of almost more than 30 Speech Emotion Recognition (SER) based reported systems in literature highlighted limitations like the dependency of such systems on voluntary expression, the possibility of concealing and acting by intention, issue of noise correlation etc [1]. Md. Rabiul Islam *et al.*, have reviewed EEG-based emotion recognition systems including the subjects, stimuli types, data acquisition and processing methods. They have listed reported systems details with machine learning-based techniques with recommendations for researchers to enhance accuracies [2]. Tengfei song *et al.*, have illustrated a multimodal emotion recognition classifier (SVM and KNN) using EEG, ECG, respiration and GSR with A-LSTM to enhance the effectiveness of discriminative features and achieved a maximum average classification accuracy of 72% [3]. Muhammad Anas Hasnul *et al.* reported a brief review of ECG-based emotion classifiers using almost 65 papers to highlight the demand for ECG-based emotion recognition systems and their popularity and relevance in healthcare technologies [4]. Amna Waheed Awan *et al.*, have reported a multimodal physiological feature-based emotion recognition system on the AMIGOS dataset using SVM, RF and LSTM to achieve a maximum classification accuracy of 94.5% [5]. R. Subramanian *et al.*, have commented on the relationship between emotional state and personality traits by analyzing EEG, ECG and GSR with facial features to establish a non-linear correlation between them [6]. An emotion recognition system with multiple modalities like Facial expression, GSR and EEG using LUMED-2 and DEAP datasets to classify seven categories of emotions with maximum accuracy of 81.2% has been reported [7]. A. Albraikan *et al.*, have reported a hybrid sensor fusion approach to develop a user-independent emotion recognition system. WMD-DTW (a weighted multi-dimensional DTW), and KNN have been used on E4 and MAHNOB datasets and achieved maximum accuracy of 94% [8].

In this work, one of the three popular publicly available datasets has been used to train emotion classifier models. The physiological signals considered in this study are EEG, ECG and GSR. The databases were acquired experimentally using specific sensors and are made available for researchers to develop an efficient emotion classifier system. The MATLAB R2020b software is used to train classifier models named Random Forest, SVM, KNN and/or Decision Tree. The classifiers have been trained and cross-validated which offers almost 100% classification accuracy for the databases they were trained with. The functional block diagram of a typical emotion recognition system is shown in Fig. 1

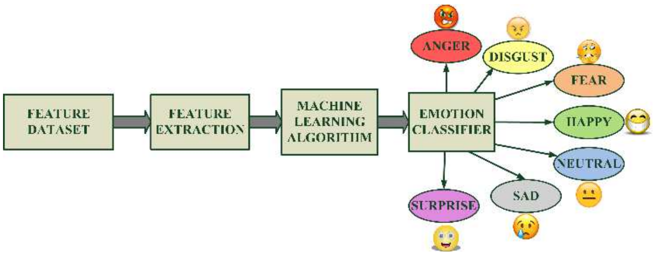

Fig.1 Typical Block Diagram of Emotion Recognition System

II. Methodology

Three publicly available datasets were acquired and features have been selected for emotion classification. The classes were allotted to the feature data based on the description and stimuli. The classifier learner models (RF, SVM, KNN and Decision Tree) were trained to classify the emotions precisely. The process flow chart is shown in Fig.2.

A. Datasets

LUMED-2, SWELL and, YAAD datasets are used in this work to develop highly precise emotion classifier models. These are publicly available multimodal feature data of physiological signals recorded using specific sensors while inducing emotions using audio-visual stimuli.

1. LUMED-2 dataset

Loughborough University Multimodal Emotion Database-2 (LUMED-2) is a multimodal database recorded for 13 subjects using emotional video stimulus by researchers of Loughborough University, UK and Hacettepe University, Turkey [9]. EEG and GSR signals data were recorded using ENOBIO 8-Channel wireless EEG device and Bluetooth-powered EMPATICA E4 wristband respectively. The data is available publicly for researchers to develop emotion recognition systems. The CSV files containing EEG and GSR data for 13 subjects are used in this work for the development of emotion classifier models [7].

2. SWELL dataset

The Smart Reasoning for Well-being at Home and at Work (SWELL) dataset is a publicly available multimodal dataset experimentally collected using body sensors to assess the stress at work on 25 subjects (Mean age: 25) [10]. The subjects were put into various interruptions, stress and time constraints conditions and HRV data was recorded. The features of HRV data were extracted for the subjects under various work pressure conditions like writing reports, interruptions through email, time pressure etc. to analyze their psychological behaviour [11]. The CSV files containing HRV feature data are used to classify the emotions into three categories for which the data was collected.

3. YAAD dataset

Young Adult's Affective Data (YAAD) is a publicly available ECG and GSR dataset for 12 subjects (Ages: 8 to 25 years) in three sessions acquired experimentally with sensors using video stimuli [12]-[13]. In this work, The CSV files containing ECG raw data are processed and features have been extracted to classify the emotions in seven categories for which emotions were induced. Ten features of ECG signals *Mean, Standard Deviation, Variance, Minimum value, Maximum value, First-degree Difference, Kurtosis, Skewness, Heart Rate, and* RR_{mean} are calculated for the ECG database and used to classify seven categories of emotions such as Anger, Disgust, Fear, Happy, Neutral, Sad and Surprise.

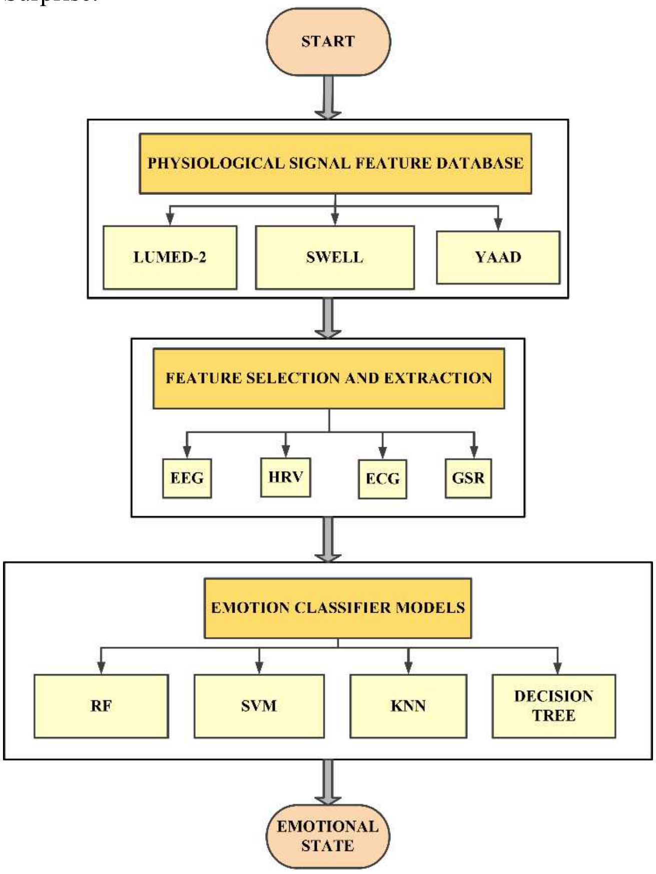

Fig. 2 Process Flow Chart

B. Deep Learning based Classifier Learners

Machine learning techniques are used to train models with feature datasets and the trained models are validated for any new input dataset. In this piece of work, various classifier learners have been used to develop emotion classifier models to classify human emotion precisely using physiological signal feature datasets. The datasets were used to train models and using the 5-fold cross-validation technique the classification accuracy and prediction speed have been recorded to arrive at the best suitable model offering the highest accuracy to classify all sets of considered emotions.

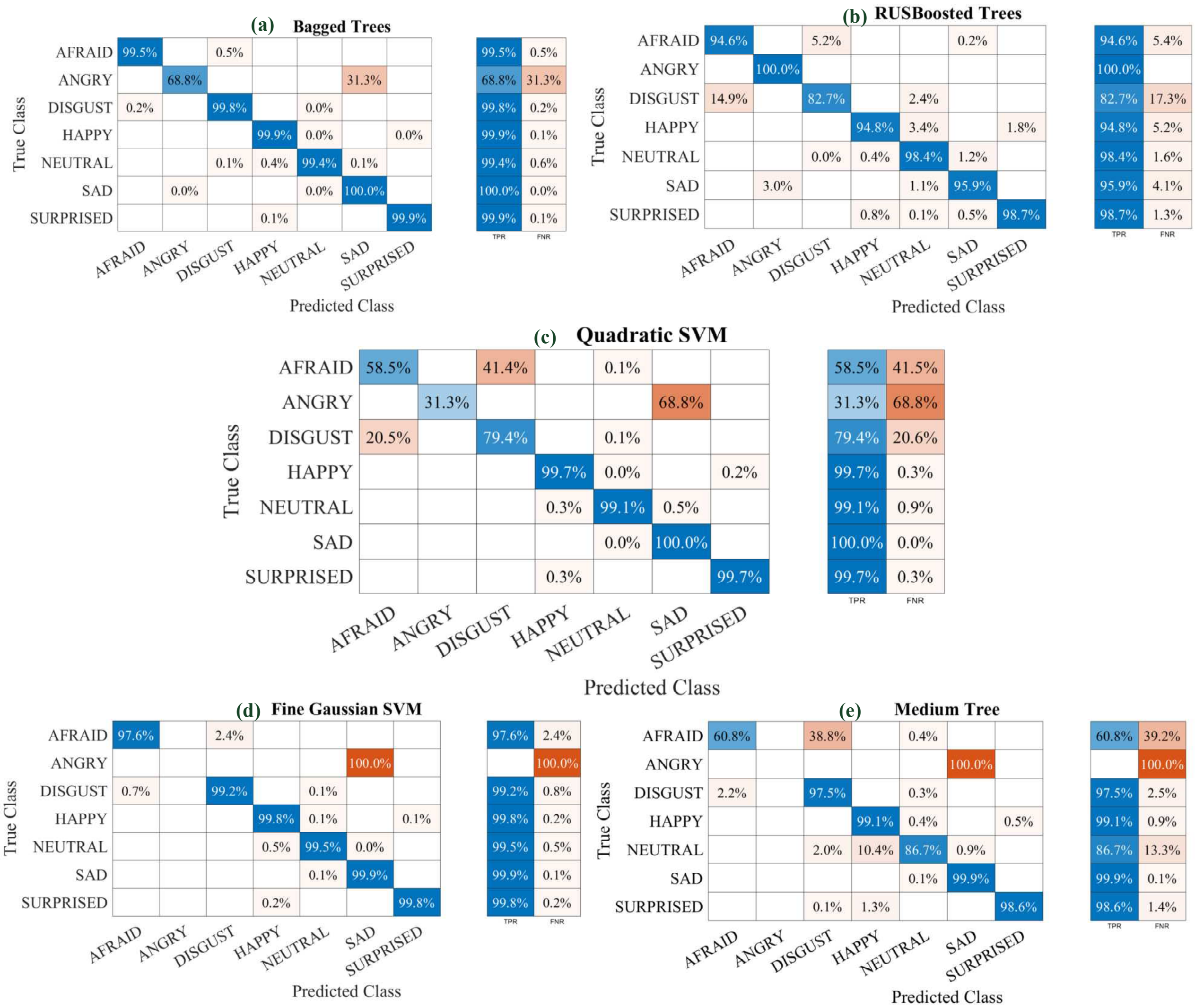

Fig. 3. Confusion matrices for one subject using LUMED-2 dataset (a) Ensemble bagged Tree (b) Ensemble RUSBoosted Trees (c) SVM Quadratic (d) SVM Fine Gaussian Tree (e) Medium Tree

1. Random Forest (RF)

It is one of the robust techniques to classify the data more accurately and precisely. Ensemble Bagged Tree and RUSBoosted Trees are two classifier learners used from this category. The Ensemble bagged tree divides the data into subsets randomly with replacement to train and reduces the variance of the decision tree. Each subset is used to train the collection of decision trees and provide benefits for all decision trees. Hence, it is an ensemble of many decision trees operated on randomly divided subsets and bagged the average response from all the decision trees to arrive at the most efficient classification models. Random forest, an extension of ensemble bagging uses a random selection of features also to grow trees. RUSBoosted reduces the errors while training models with previous simple decision tree classifiers to arrive at the most accurate classification.

1. Support Vector Machine (SVM)

It is one of the most efficient supervised classifier learners. It divides the data into support vectors and keeps one vector of the data at a time for validation, trains the rest vectors and this process repeats till the learner finds the best suitable decision boundary (hyperplane) to classify the data precisely and accurately. In this study, two non-linear SVM classifiers have been used to develop emotion recognition models. It offers different kernel functions among which Quadratic and Fine Gaussian are used in this research work.

2. KNN

It is one of the simplest, basic and non-parametric supervised machine learning techniques which use K-Nearest Neighbours (KNN) to classify the new dataset. It stores the dataset during training and testing or cross-validation it uses the technique of nearest neighbours distance method to classify the data. Here KNN is used to classify the datasets to analyze if the parametric relationship among features is not assumed. In this work, k was taken as 5 and Euclidean distance was calculated to identify the number of nearest neighbours.

3. Decision Tree

It is also a non-parametric supervised learner which classifies datasets by dividing them into various nodes of a tree like a root node, internal nodes, leaf node and branches. It uses the divide and conquers technique to identify the most optimal splits in the developed decision tree. In this work, a medium tree was used to classify the datasets wherever applicable, as the ensemble decision tree method is void of disadvantages such as overfitting caused by generic and simple decision tree algorithms

III. RESULT AND DISCUSSION

Three publicly available physiological signals datasets for the development of emotion recognition systems have been used to develop emotion classifier models using machine learning techniques. LUMED-2 dataset is an EEG and GSR features dataset recorded using sensors while inducing emotions with audio-visual stimuli.

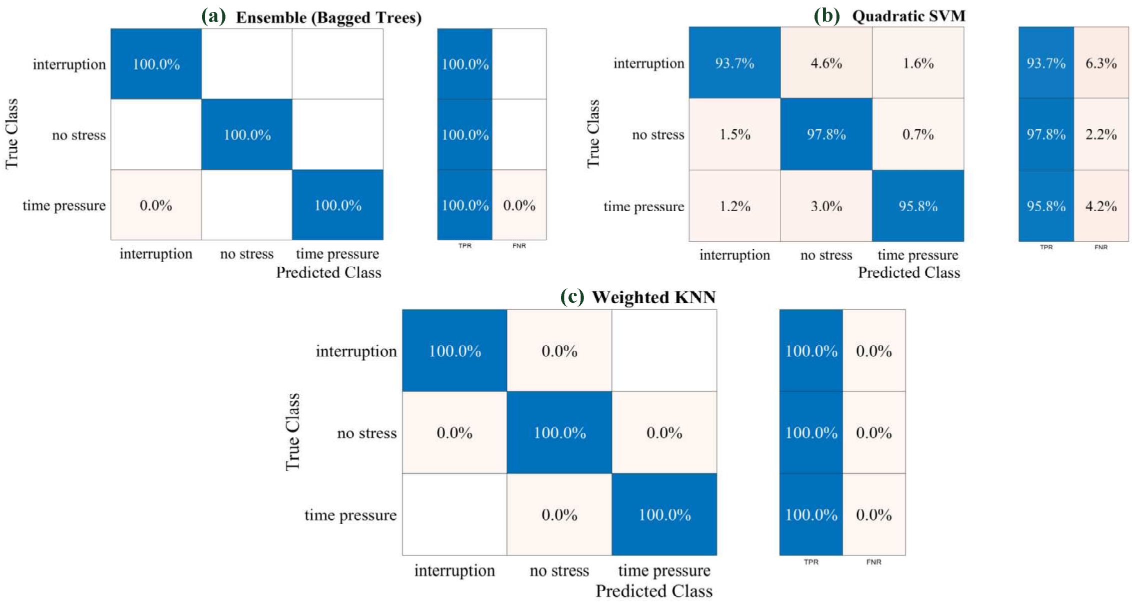

Fig. 4 Confusion matrices using SWELL dataset (a) Ensemble Bagged Trees (b) SVM Quadratic (c) Weighted KNN

The dataset used here contains input features like subject number (to include individual variability), 8 electrodes EEG outputs, and GSR signal output with an output containing 7 emotional states (4 on average). The prepared dataset has been used to classify using five different classifier learners. The confusion matrices for one subject with seven emotions categorized using the selected five learners are shown in Fig. 3 (a)-(e). It can be seen in Fig.3. (e), the medium tree classifier model can classify four emotions (SAD, HAPPY, SURPRISED, and DISGUST) with a minimum average classification accuracy of 97.5%. This classifier has misclassified ANGRY as SAD. In the confusion matrix in Fig. 3 (c), the SVM Quadratic classifier can classify four of seven emotions with an accuracy of more than 99%. The misclassification for the other three emotions is high. The fine gaussian classifier model is capable of classifying all emotions with a minimum average accuracy of 97.6% except the ANGER state was 100 % misclassified as SAD, shown in Fig. 3 (d). The misclassification issue with these classifiers was reduced using Random Forest classifiers shown in Fig. 3 (a) and (b) respectively. The ensemble RUSBoosted classifier can classify all emotions very accurately and minimise the false negative rate (FNR). The ensemble-based classifier is popular for the development of accurate and precise emotion classifier systems [14].

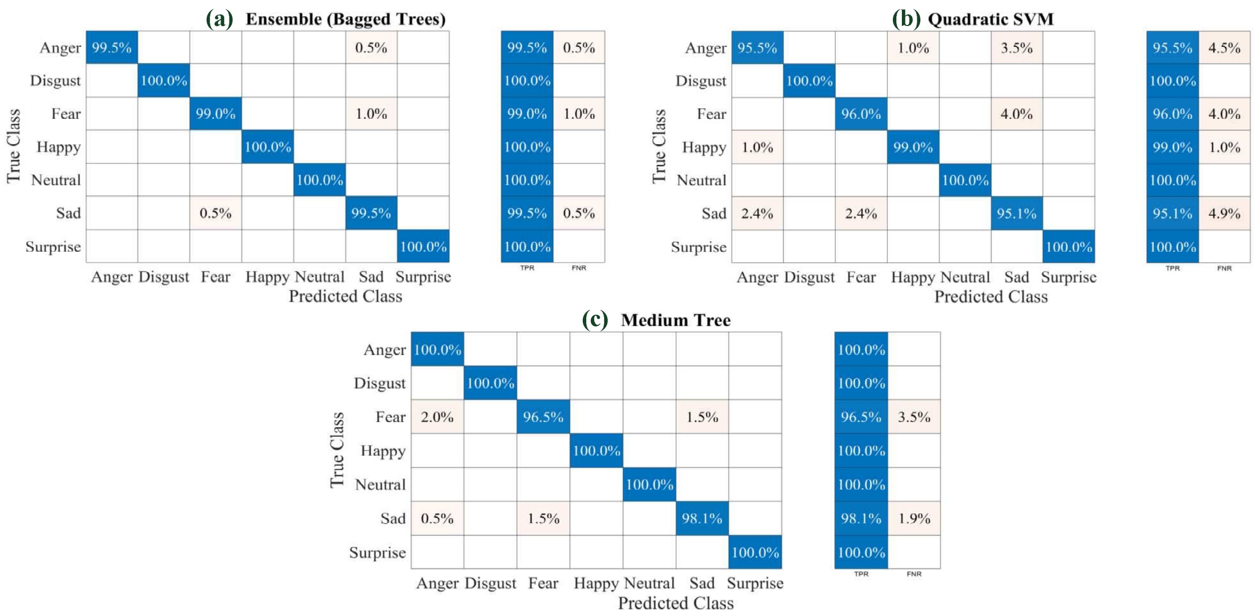

Fig. 5 Confusion matrices for Set-1 using YAAD dataset (a) Ensemble Bagged Trees (b) SVM Quadratic (c) Medium Tree

TABLE I. CLASSIFICATION ACCURACY OF THE TRAINED AND CROSS-VALIDATED EMOTION RECOGNITION MODELS

Dataset Type	No. of Subjects	No. of Features	No. of Emotions	Classifiers	Model Type	Average Accuracy in Emotion Recognition (%)
LUMED-2 (EEG + GSR)	13	10	7 (Four in Average)	*Random Forest (RF)*	*Ensemble Bagged Tree*	99.9%
					Ensemble RUSBoosted	93.4%
				Support Vector Machine (SVM)	*Quadratic*	*92.3%*
					Fine Gaussian	99.5%
				Decision Tree	*Medium Tree*	96.2%
SWELL (HRV)	25	35	3	*Random Forest (RF)*	*Ensemble Bagged Tree*	100%
				Support Vector Machine (SVM)	*Quadratic*	96.3%
				KNN	*Weighted KNN*	100%
YAAD (ECG + GSR)	12	10	7	*Random Forest (RF)*	*Ensemble Bagged Tree*	99.7%
				Support Vector Machine (SVM)	*Quadratic*	*98%*
				Decision Tree	*Medium Tree*	99.2%

TABLE II. PERFORMANCE PARAMETERS OF THE EMOTION RECOGNITION MODELS

Dataset Type	Classifier Model Type		5 Folds Cross Validation	
			Average Training Time (seconds)	***Prediction Speed (obs/second)***
LUMED-2	Random Forest (RF)	*Ensemble Bagged Tree*	95.601	100000
		Ensemble RUSBoosted	34.452	68000
	Support Vector Machine (SVM)	*Quadratic*	2527.12	29000
		Fine Gaussian	293.47	11000
	Decision Tree	*Medium Tree*	5.1732	1200000
SWELL	Random Forest (RF)	*Ensemble Bagged Tree*	54.112	100000
	Support Vector Machine (SVM)	*Quadratic*	2354.5	10000
	KNN	*Weighted KNN*	396.46	3600
YAAD	Random Forest (RF)	*Ensemble Bagged Tree*	3.9681	5600
	Support Vector Machine (SVM)	*Quadratic*	1.8583	10000
	Decision Tree	*Medium Tree*	0.45005	60000

The SWELL dataset contains the HRV features of the subjects calculated using ECG data. It was recorded to analyze the stress level conditions of working professionals. The dataset includes 36 different features of HRV data as an input with 3 categories (interruption, time pressure and no stress) as an output to understand the psychological condition of working professionals posing them with various stressful tasks. The dataset was used to train 3 different models which are capable of classifying the mental state with 100% accuracy. The confusion matrices for all three selected models are shown in Fig. 4 (a)-(c). It can be seen that the three classes of stress levels are due to the stimuli given to the subjects. From Fig. 4 (b) in which quadratic SVM was used to classify emotional state, the maximum FNR is 6.3% for 'interruption' whereas Fig. 4 (a) and (c) show Ensemble bagged tree and weighted KNN models which are capable of classifying the emotions with 100% TPR. Hence, for the SWELL dataset in which 36 HRV features were taken as input, Ensemble bagged Tree and Weighted KNN can classify emotions with 100% accuracy. YAAD dataset is ECG and GSR raw datasets. The raw ECG dataset was used to calculate statistical features (10 features) and the dataset is prepared. The prepared feature dataset was used to train 3 different classifier learners. YAAD datasets included 3 sets of raw ECG data which were used to calculate 10 statistical features as inputs and corresponding emotional states (a total of seven categories) were allotted as output to have the final feature dataset. The confusion matrices of all three trained classifier models corresponding to one set of feature data are shown in Fig. 5 (a)-(c). It can be seen that the maximum FNR in the SVM quadratic was 4.9 % (for sad), 3.5% in Medium Tree (for fear) whereas the maximum FNR obtained in Ensemble bagged tree is 1% (for fear), the least of all. Hence, Ensemble bagged Tree can classify all seven emotions with a minimum classification accuracy of 99%. These three datasets have been recorded by researchers and shared with others for the development of emotion recognition systems. The performance and accuracies obtained using selected machine learning classifiers are quite high in comparison with available models using these datasets. The trained models using selected datasets and the average maximum classification accuracy achieved in this work are tabulated in Table I. The performance parameters of the trained models like training time and prediction speed are included in Table II.

IV. Conclusion

Three publicly available datasets have been used to develop highly accurate emotion classifier models using EEG, HRV and ECG. The developed models are capable of classifying emotions with almost up to 100% accuracy which is quite higher in comparison to the classifiers developed by researchers who shared these datasets. Emotionally intelligent systems are very much dependent on the accuracy with which they can access human emotions to execute further functions accordingly. Such systems have the potential to be realized into wearable devices capable of assessing real-time emotional state as well as identifying mental illness for early detection to provide required medical attention and care. Emotion detection systems using physiological signals with such precision and accuracy are surely going to help in the development of the digital healthcare infrastructure. In future work, clinical acceptance and end-use product realization will be done to identify patients with depression and autism.

Acknowledgement

This research work acknowledges the financial support received as a research fellowship from UGC under UGC-NET-JRF Scheme (R.No-190510901080).

References

[1] T. M. Wani, T. S. Gunawan, S. A. A. Qadri, M. Kartiwi, and E. Ambikairajah, "A Comprehensive Review of Speech Emotion Recognition Systems," IEEE Access, vol. 9, pp. 47795–47814, 2021.

[2] M. R. Islam et al., "Emotion Recognition From EEG Signal Focusing on Deep Learning and Shallow Learning Techniques," IEEE Access, vol. 9, pp. 94601–94624, 2021.

[3] T. Song, W. Zheng, C. Lu, Y. Zong, X. Zhang, and Z. Cui, "MPED: A Multi-Modal Physiological Emotion Database for Discrete Emotion Recognition," IEEE Access, vol. 7, pp. 12177–12191, 2019.

[4] M. A. Hasnul, N. A. A. Aziz, S. Alelyani, M. Mohana, and A. A. Aziz, "Electrocardiogram-Based Emotion Recognition Systems and Their Applications in Healthcare—A Review," Sensors, vol. 21, no. 15, p. 5015, Jul. 2021.

[5] A. W. Awan et al., "An Ensemble Learning Method for Emotion Charting Using Multimodal Physiological Signals," Sensors, vol. 22, no. 23, p. 9480, Dec. 2022.

[6] R. Subramanian, J. Wache, M. K. Abadi, R. L. Vieriu, S. Winkler, and N. Sebe, "Ascertain: Emotion and personality recognition using commercial sensors," IEEE Trans. Affect. Comput., vol. 9, no. 2, pp. 147–160, 2018.

[7] Y. Cimtay, E. Ekmekcioglu, and S. Caglar-Ozhan, "Cross-subject multimodal emotion recognition based on hybrid fusion," IEEE Access, vol. 8, pp. 168865–168878, 2020.

[8] A. Albraikan, D. P. Tobon, and A. El Saddik, "Toward User-Independent Emotion Recognition Using Physiological Signals," IEEE Sens. J., vol. 19, no. 19, pp. 8402–8412, 2019.

[9] E. Ekmekcioglu, "Loughborough University Multimodal Emotion Dataset-2. figshare.Dataset," CIMTAY, Yucel, vol. 5, 2020.

[10] M. A. Neerincx and W. Kraaij, "The SWELL Knowledge Work Dataset for Stress and User Modeling Research Categories and Subject Descriptors," Proc. 16th Int. Conf. Multimodal Interact. 2014 Pages 291–, pp. 291–298, 2014.

[11] K. Nkurikiyeyezu, K. Shoji, A. Yokokubo, and G. Lopez, "Thermal comfort and stress recognition in office environment," Heal. 2019 - 12th Int. Conf. Heal. Informatics, Proceedings; Part 12th Int. Jt. Conf. Biomed. Eng. Syst. Technol. BIOSTEC 2019, no. February, pp. 256–263, 2019.

[12] A. Akram, Muhammad Usman; DAR, MUHAMMAD NAJAM; Rahim, "Young Adult's Affective Data (YAAD) Using ECG and GSR Signals," Mendeley Data, vol. 4, 2022.

[13] M. N. Dar, A. Rahim, M. U. Akram, S. Gul Khawaja, and A. Rahim, "YAAD: Young Adult's Affective Data Using Wearable ECG and GSR sensors," 2022 2nd Int. Conf. Digit. Futur. Transform. Technol. ICoDT2 2022, no. July, 2022.

[14] T. Dissanayake, Y. Rajapaksha, R. Ragel, and I. Nawinne, "An ensemble learning approach for electrocardiogram sensor based human emotion recognition," Sensors (Switzerland), vol. 19, no. 20, pp. 1–24, 2019.