A Comparative Study of Machine Learning Techniques for Emotion Recognition from Peripheral Physiological Signals

Sowmya Vijayakumar
Department of Computer and Software Engineering
Athlone Institute of Technology
Athlone, Co. Westmeath, Ireland
https://orcid.org/0000-0002-7304-8342

Ronan Flynn
Department of Computer and Software Engineering
Athlone Institute of Technology
Athlone, Co. Westmeath, Ireland
rflynn@ait.ie

Niall Murray
Department of Computer and Software Engineering
Athlone Institute of Technology
Athlone, Co. Westmeath, Ireland
nmurray@research.ait.ie

***Abstract*—Recent developments in wearable technology have led to increased research interest in using peripheral physiological signals for emotion recognition. The non-invasive nature of peripheral physiological signal measurement via wearables enables ecologically valid long-term monitoring. These peripheral signal measurements can be used in real-time in many ways including health and emotion classification. This paper investigates the utility of peripheral physiological signals for emotion recognition using the publicly available DEAP database. Using this database (which contains electroencephalogram (EEG) signals and peripheral signals), this paper compares eight machine learning models in the classification of valence and arousal emotion dimensions. These were applied to the peripheral physiological signals only. These models operate on three groupings of the peripheral data: (i) the raw peripheral physiological signals; (ii) individual feature sets extracted from each peripheral signal; and (iii) a fusion data set made of the combined features from the individual peripheral signals. The results indicate that support vector machine, linear discriminant analysis and logistic regression give the best recognition results on all three data groups considered. The feature fusion data set, which is made up by fusing all the features from the peripheral signals, gives the best recognition accuracy on both valence and arousal dimensions. In addition, subject dependency for emotion classification from peripheral signals is examined and significant individual variability is observed. The recognition rate varies between each participant from 10% to 87.5%.**

Keywords—peripheral physiological signals, emotion recognition, machine learning, wearables, classification

I. Introduction

Emotion has a central role in human experiences associated with perception, the cognitive process, behaviour and decision-making. The last two decades have seen a significant increase in emotion recognition research within the field of affective computing. The human-computer interaction (HCI) community has adopted the concept of pleasurable experience. Systems should be designed such that they improve the emotional experience during human-system interactions. However, to achieve this, the challenge of recognising and synthesising the emotional responses [1] needs to be addressed.

Emotions in humans produce physical and physiological changes. Physiological signals are widely studied in emotion recognition because they are involuntary and, unlike physical signals such as facial expressions, speech, and gesture, they cannot be easily controlled [1], [2], [3], [4]. Electroencephalogram (EEG) measurements derive physiological signals related to the brain activity directly from the central nervous system (CNS). Peripheral physiological signals such as electrocardiogram (ECG), electrooculogram (EOG), electromyogram (EMG), blood volume pressure (BVP), galvanic skin resistance (GSR), skin temperature and respiration are induced by the activities of the peripheral nervous system (PNS). Of course, these types of signals and responses are highly individualised. However, emotion recognition research has typically produced generalised recognition models using subject-independent approaches. Some authors report that the high inter-participant variability inherent to emotions affects the classification results [1], [5], [6], [7].

Emotion recognition using EEG signals, combined with machine learning, pattern recognition [8] and deep learning algorithms [4], [9], [10], [11] has emerged as a powerful approach to the recognition of emotions since it can extract the cognitive state of the user from the CNS. Studies have also employed multimodal approaches, fusing EEG with peripheral signals, which have improved the accuracy of emotion classification [1], [2], [12]. However, the invasive nature of the EEG measurement is an issue, particularly over long periods of time when it can lead to feelings of discomfort, which makes it unsuitable for real-time emotion recognition applications [13], [14]. On the other hand, the non-invasive nature of peripheral physiological signals enables comfortable long-term assessment through wearables in the form of activity trackers and health monitors as listed in [5]. These devices are becoming more and more integrated into human activities, providing better usability and hands-free experiences. Peripheral physiological signals are a good choice in applications such as automatic emotion recognition [4], healthcare systems [14], HCI [10] and Quality of Experience (QoE) [15], [16]. For example, closely related to emotion recognition, these signals are used in evaluating the user QoE in augmented reality (AR) and virtual reality (VR) technologies [15] measuring electrodermal activity (EDA), heart rate (HR) and skin temperature. Another study [16] measures these signals for QoE evaluation in VR applications. Hence, studies to understand the user are shifting their focus from EEG-based solutions to the use of peripheral signals in emotion recognition [3], [5]. The DEAP [1] database has been widely studied, but only a few of the studies have involved peripheral physiological signals only [13], [14], [17], [18].

This research presents a detailed evaluation of the relevance of peripheral physiological signals for classifying emotions using the publicly available DEAP dataset. The recognition rate depends on the selected features, classifier and the target dataset [3]. Eight machine learning methods were compared in the classification of emotional arousal and valence. The methods used were support vector machine

978-1-7281-9418-9/20/$31.00 ©2020 IEEE

(SVM), K-nearest neighbours (KNN), random forest (RF), decision tree (DT) logistic regression (LR), Gaussian naïve Bayes (GNB), linear discriminant analysis (LDA) and multilayer perceptron (MLP) classifier. As part of the valence and arousal classification, features were extracted from the peripheral signals. Three different data combinations were used: fusion of raw dataset consisting of all the DEAP peripheral signals; the individual feature datasets extracted from the peripheral signals; a feature fusion set consisting of the combined features extracted from all the peripheral signals. The results presented serve as a baseline for comparison for future research into the use of peripheral physiological signals for emotion classification. In addition to emotion recognition performance, individual participant dependency on the classification accuracy from peripheral physiological signals is investigated by observing the variations in classification accuracy between individuals. Few studies have observed these differences in multi-subject emotion recognition [6], [7]. One of the approaches adopted to tackle this is group-dependent recognition study [7] where the subjects are clustered into groups based on the characteristics of the physiological signals and the results show improved performance in recognising four emotions. In the current study, the subjects are grouped based on their gender. The variations in the classification accuracies of valence and arousal recognition are observed between male and female groups.

The rest of this paper is structured as follows. Section II gives a brief explanation on the pre-processing of the DEAP peripheral physiological signals, features extracted from these signals and the experimental setup of classification models. Section III presents the results from model comparison, performance evaluation and subject-dependent classification, then discussed in Section IV, and the paper conclusions are given in Section V.

II. Experimental Method

This section describes the DEAP dataset and extraction of statistical features, and the experimental setup for classifying emotions.

A. DEAP Database

The publicly available DEAP dataset is widely used in the research community in the study of emotion recognition. This dataset consists of 32 EEG channels and 8 channels from six peripheral physiological signals recorded from 32 participants, aged between 19 and 37. The signals were recorded as participants watched 40 music videos designed to arouse emotions in all four quadrants of the valence/arousal space. The participants rated the videos in terms of valence, arousal, dominance and liking. Single-channel peripheral signals (GSR, skin temperature, respiration, and BVP) and dual-channel peripheral signals (EOG (horizontal and vertical) and EMG (zygomaticus major and trapezius muscles)) were recorded. The data is available in the raw and preprocessed format.

For the experiments presented here, the preprocessed dataset in python-format is used. The data is downsampled to 128Hz from 512Hz, filtered, segmented, and artefacts removed. Since the focus of this study is peripheral physiological signals, only 8 peripheral channels out of 40 channels from the dataset are considered, while ignoring the remaining 32 EEG channels. Each channel has 8,064 samples. The structure of the data array in the study is 32 participants x 40 video/trial x 8 channel x 8,064 samples. The emotions are categorised in a multi-dimensional space. It includes an array of 40 video/trial x 4 label (valence, arousal, dominance and liking). Each label is rated on a continuous scale from 1 to 9. Only the valence and arousal labels are used here. Based on other classification studies [1], [12], [14], [17] each arousal and valence labels are divided into 2 classes, high if greater than 4.5 and low if less than or equal to 4.5.

B. Feature Extraction

In a recent review paper [3], the features adopted in related studies for emotion classification are studied. These include time, frequency and wavelet domain features. In [17], feature set attributes are used for feature extraction, while [3] suggest specific features for certain physiological signals. For the work presented here, features extracted are based on [1], [3], [13], [17] and are summarised as standard or non-linear as shown in TABLE I.

The sample entropy measures the complexity based on approximate entropy. Lyapunov exponent measures total predictability, lyap_r and lyap_e estimates largest Lyapunov exponent and a whole spectrum of Lyapunov exponent respectively, whose positive exponents indicate chaos and unpredictability. The Hurst exponent is used to find out if the time series follows a similar pattern as in previous steps. The detrended fluctuation analysis (dfa) measures the Hurst parameter H, which is very similar to the Hurst exponent used for non-stationary processes [19].

C. Peripheral Physiological Features

For fusion data analysis, 16 standard statistical features are extracted from each channel. This resulted in a total of 128 features, concatenated into a single feature space from the 8 channels (those without EEG in the DEAP database). In DEAP [1], 106 features are extracted from peripheral physiological signals. In addition, specific features have also been estimated for individual peripheral signals as listed in TABLE II. EOG measures eye blinking rate related to anxiety, while EMG measures muscle activity. Skin temperature and respiration signals vary with different emotional states [1]. The blood volume in the participant's thumb is measured by plethysmograph from which heart rate and heart rate variability can also be measured. GSR also referred to as electrodermal activity (EDA) can be comfortably assessed for longer durations measuring skin's electrical resistance connected to the sympathetic nervous system [17] related to the level of arousal [1].

TABLE I. FEATURES EXTRACTED FROM PERIPHERAL PHYSIOLOGICAL SIGNALS

Standard Features
Mean, min, max, standard deviation, median, skewness, kurtosis, power spectral density, average of the derivative, mean of the normalised signal, min and max of the normalised signal, mean of the absolute values of first and second differences of raw and normalised signals.
Non-linear Features
Sample entropy, Hurst exponent, detrended fluctuation analysis (dfa), Lyapunov exponent (lyap_r and lyap_e)

TABLE II. LIST OF FEATURES EXTRACTED FROM PERIPHERAL PHYSIOLOGICAL SIGNALS

Modality	No. of features extracted	Features extracted
EOG	16	Standard features
EMG	16	Standard features
GSR	18	16 standard features and 2 non-linear features - Hurst component and sample entropy
BVP	16	Standard features
Respiration	17	14 standard (except skewness and kurtosis) and 3 non-linear features – dfa, lyap_r and lyap_e
Temperature	16	Standard features
ALL	128	Standard features

Relaxed and aroused state is linked to the respiration rate [1]. Detrended fluctuation analysis (dfa) and Lyapunov exponent (lyap_r and lyap_e) are extracted for the respiration signal as they have shown improved recognition rate [3]. Skewness and kurtosis are not used as respiration features.

D. *Classification Models*

The studies on emotion recognition using EEG signals suggest several classification models. A GNB classifier is used in [1], along with F1-scores to deal with unbalanced classes, to evaluate recognition performance with a leave-one-out cross-validation technique followed by decision fusion using a weighting scheme. A new peripheral feature space was created in [13] using canonical correlation analysis (CCA) and an SVM classifier was used. A minimum redundancy maximum relevance (mRMR) algorithm is adopted in [14] for feature selection from skin temperature, respiration and BVP signals; the feature data is then fed to RF, LR and SVM algorithms for classification. The RF classifier produced the best results; it was also used in [17] to classify emotions via GSR for its capability in handling high dimensional data. A list of classification models suitable for emotion recognition can be found in [3].

The models employed in these experiments, informed by the referenced research, are SVM, KNN, RF, DT, LR, GNB, LDA and a single-layer neural network (NN) using a multilayer perceptron (MLP) classifier. Since the optimal classification algorithm depends on the target data and various conditions such as selected features and fusion techniques [3], the goal is to find the effective model to be used with the DEAP dataset to classify valence and arousal using peripheral physiological signals.

E. *Platforms*

The experimental setup includes implementation of machine learning algorithms in the Python programming language (v3.7) on the Spyder integrated development environment. The NVIDIA GeForce GTX 1080 graphics card was used. The imported packages are: scikit-learn for machine learning; numpy and pandas for fundamental operations and mathematical computations; scipy for signal processing; nolds [19] for extracting non-linear features; and matplotlib for plotting graphs. The parameter settings for the classifiers are: 5 nearest neighbours for KNN; radial bias function (RBF) kernel in SVM; and 100 estimators in RF. Other classifiers are implemented with their default settings in the scikit-learn package. The default settings for the MLP classifier are adam solver, relu activation, constant learning rate and alpha = 0.0001.

III. RESULTS

This section compares the results obtained by applying the classification models to the peripheral physiological signals. The performance evaluations of three data combinations are presented followed by the results of the subject-dependent classification.

A. *Comparison of Classification Models*

In this section, 8 machine learning algorithms applied to the peripheral physiological signals from the DEAP database are evaluated. The experiment is conducted on three data combinations: raw peripheral physiological signals; feature sets extracted from individual peripheral signals; feature fusion set obtained by fusing all the peripheral features extracted from the individual signals. In this study, the preprocessed data provided with the DEAP database is used as a raw signal for the experiments. The classification performance is validated by 10-fold cross-validation and the achieved accuracies are reported.

For the raw data analysis, first, the peripheral physiological signals from eight channels were separated from EEG data and then integrated into a single dataset before classification. The data array has 8,064 data samples per channel and the total number of data points from 8 peripheral channels, each with 32 users watching 40 videos, is significantly large. To simplify model training, the dimensionality of the data is reduced using principal component analysis (PCA) [3], [13] as a preprocessing step before being input to the classification algorithms. For feature fusion analysis, as mentioned in section II.C, 128 features from 8 channels were extracted and concatenated to form a peripheral feature space. PCA of 10 and 20 components was applied to the feature sets, however, the recognition accuracies were slightly better without PCA.

The average accuracy and F1 score obtained by each classifier on two of the data set combinations (fusion of raw data and feature fusion) are shown in Fig. 1 and Fig. 2 respectively. Across these sets, the RF classifier achieved an accuracy of 64.45% for valence recognition, which is slightly higher than SVM with an accuracy of 63.12% (with configuration settings of 128 features without PCA as input to the classifier). Moreover, the results show that SVM outperforms other classifiers on arousal recognition for the same configuration on the feature fusion set. The best accuracies and F1 scores on the raw dataset are achieved by the LDA, LR, SVM and MLP classifiers for both valence and arousal classification. KNN did not perform very well for the proposed features and classification settings while DT and GNB gave the poorest accuracy results for both data set combinations.

For individual signal analysis, the classifiers were applied to the individual feature sets extracted from each peripheral signal. Prior to classification, for each modality and label (arousal or valence), three combinations of PCA were selected: no PCA, 5 and 10 components. The best configuration results from each signal are shown in Fig. 3, plotting the accuracy and F1 score while comparing the 8 classification models used to recognise valence and arousal. Respiration and GSR yielded better results with PCA of 5 components while all other modalities (EOG, EMG, BVP and skin temperature) performed well without any PCA reduction applied to the features. The multiple subplots showed in Fig. 3 compares the classifiers for each of the peripheral features.

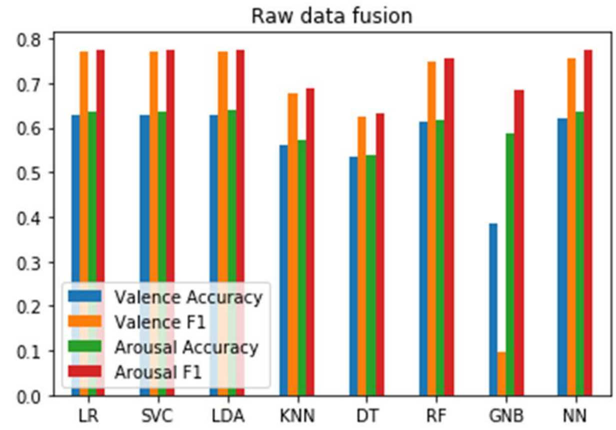

Fig. 1: Comparison of ML models on raw data combination.

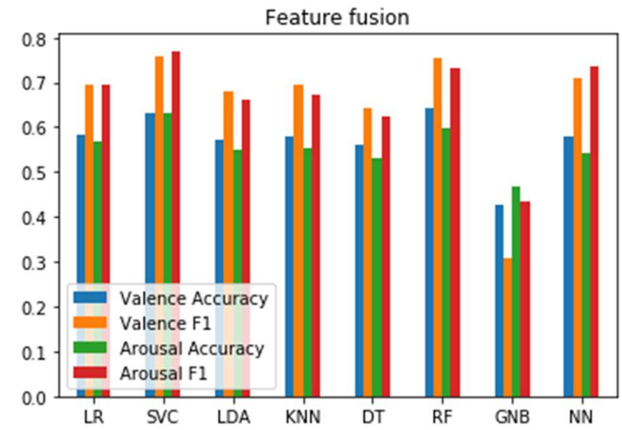

Fig. 2: Comparison of ML models on feature fusion data.

The classifiers SVM, LDA and LR again outperform the other classifiers across all the individual signals. The MLP classifier achieves higher accuracies on EMG signal for valence classification and the BVP signal for arousal classification. The highest accuracy rate of 64.92% is achieved by the SVM classifier on the EOG signal for valence classification. Results show that for all 3 groups the SVM, LDA and LR classifiers perform equally well, giving better recognition results on both valence and arousal compared to the other classifiers, while MLP classifier showed comparable results on feature fused data and three individual modalities such as respiration, EMG and BVP signals.

B. *Performance Evaluation*

The F1 scores and recognition rates for the classification in different modalities are given in TABLE III. The results show that the raw signals without prior feature extraction achieved the best accuracy of 63.86% for arousal classification. The EOG signal achieved the highest valence classification rate of 64.92% followed by the feature fusion set with an accuracy of 64.45%.

TABLE III. ACCURACY AND F1 SCORES OF AROUSAL AND VALENCE CLASSIFICATION FOR DIFFERENT DATA INPUT MODALITIES

Modality	**Valence**		**Arousal**	
	Accuracy	***F1***	***Accuracy***	***F1***
All raw (PCA = 20)	62.99	0.7709	**63.86**	**0.7757**
All feature (PCA =10)	63.59	0.7658	62.59	0.7638
All feature (PCA =20)	61.25	0.7439	61.56	0.7557
All feature (no PCA)	**64.45**	**0.7588**	63.20	0.7702
GSR (PCA = 5)	62.89	0.7698	63.52	0.7733
EOG (no PCA)	**64.92**	**0.7661**	63.05	0.7688
EMG (no PCA)	63.28	0.7709	63.67	0.7748
Temperature (no PCA)	63.13	0.7720	63.67	0.7748
Respiration (PCA = 5)	62.89	0.7698	63.75	0.7755
BVP (no PCA)	63.28	0.7704	63.52	0.7706

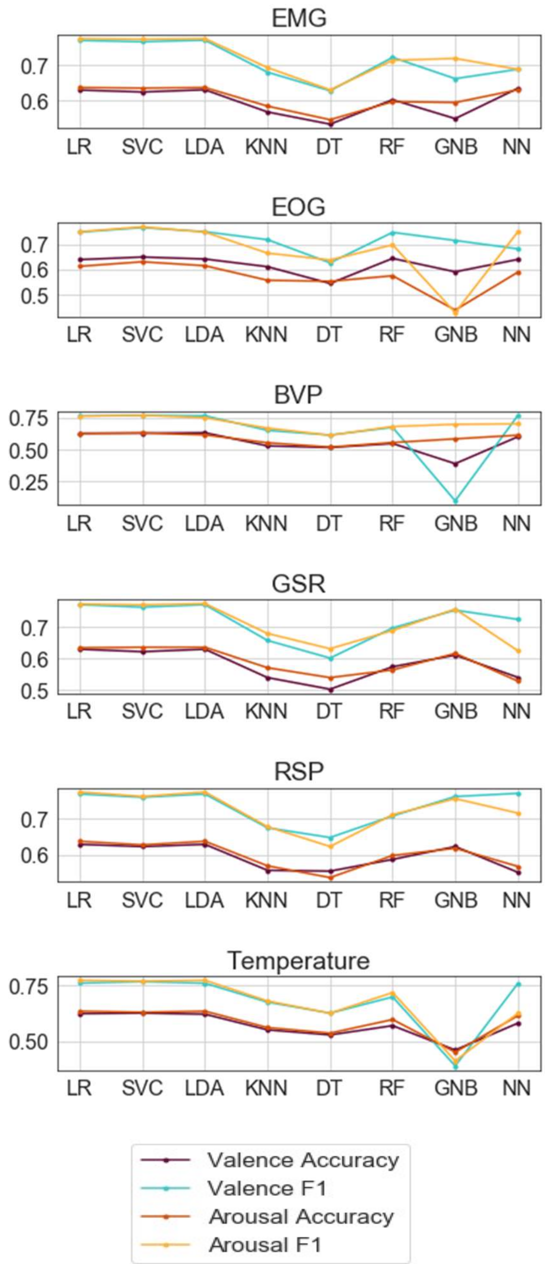

Fig. 3: Accuracy and F1 scores for 8 classifiers on individual peripheral signals, a) EMG (*no PCA*), b) EOG (*no PCA*), c) BVP (*no PCA*), d) GSR (*PCA=5*), e) Respiration (*PCA=5*), and f) Temperature (*no PCA*).

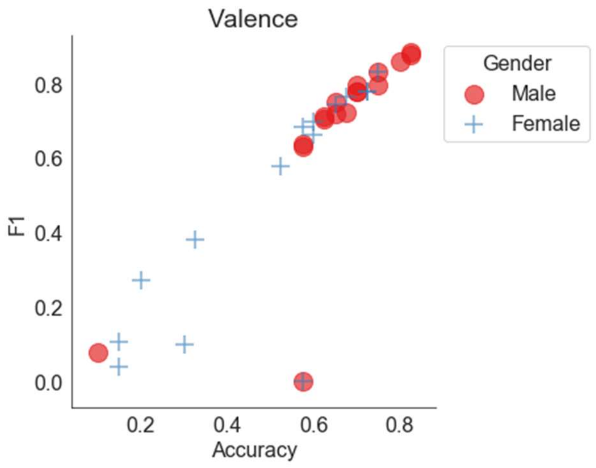

Fig. 4: Scatterplot of 32 subjects and their valence accuracies and F1 scores grouped based on gender

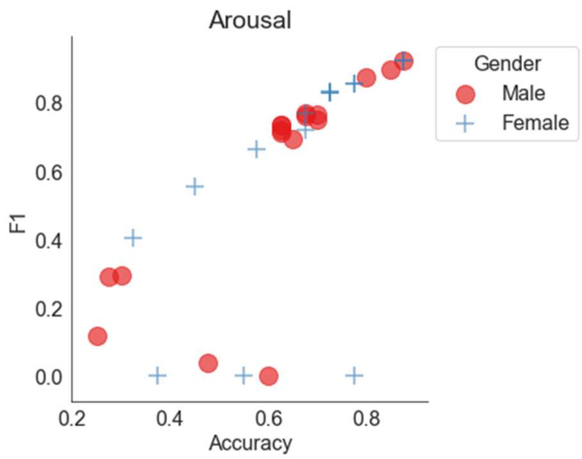

Fig. 5: Scatterplot of 32 subjects and their arousal accuracies and F1 scores grouped based on gender

All the individual peripheral physiological signals considered, and fusion of these signals show the same level of significance in classifying arousal dimensions with 63% approximate accuracy. The PCA has found to give better results for GSR and respiration signals, while the remaining modalities performed well without PCA.

C. *Subject-dependent Classification*

The goal of subject-dependent classification is to investigate the individual variability between subjects on emotion recognition from their peripheral physiological signals. The results from the previous evaluations shown that SVM, LDA and LR give optimal accuracies for the DEAP peripheral physiological signals, therefore SVM is used here. The test is carried out on 6 peripheral physiological signals of each subject separately by running SVM classifier and evaluated using 10-fold cross-validation. This was repeated on all 32 subjects and the resulting accuracies and F1 scores for valence and arousal are plotted in Fig. 4 and Fig. 5 respectively. The correlation between emotions and peripheral physiological signals differ from person to person reflecting on the high variance of accuracy on both valence and arousal. The classification accuracy rate on both valence and arousal vary between 10% and 87.5%.

The metadata of the participants in the DEAP dataset consists of age, gender, education, handedness and consumption of beverages, alcohol, tobacco and other drugs. The age group ranges from 19 and 37, which is too small to observe any variations. Of the 32 participants, there are 15 female and 17 male subjects in the dataset. To explore the effect of gender on emotion recognition from peripheral physiological signals, the classification rate was investigated by grouping the subjects based on their gender. Fig. 4 on valence recognition shows that majority of the male subjects have high classification rates, while the female subjects are distributed across low and high accuracies. On average the male subjects exhibit higher valence classification rates, with an average accuracy of 65.29%, compared to female subjects, with a 49.83% classification rate. However, on arousal classification in Fig. 5, female subjects perform better than male subjects achieving average classification rate of 65.83% and 60.73% respectively.

IV. Discussion

A comparison of the results obtained with previously published results is given in TABLE IV. The results from this work are better than those obtained in [1], [13], [18]. The DEAP database single-trial classification results achieved a recognition accuracy rate of 62.7% for valence and 57% for arousal [1]. C. Godin et al. [18] achieved similar results as DEAP database study [1] with a few well-selected features. The results from the studies [14], [17] on individual signals show better recognition accuracy than those obtained here. One possible reason for this is the use of a sliding window approach for feature extraction in [14], [17]. This will be an avenue for future work. In [13], the peripheral feature space is constructed using EEG signals, the results are better than [1] for arousal and worse for valence.

For the experiments in [13], it is noted that on the 9-point continuous rating scale the output label is divided into two classes with a threshold value of 5, above which is high, and below or equal is low. When the models were trained on our configuration of data with a threshold value of 5, the results show that the recognition performance decreased significantly by approximately 12.5 percentage points for feature fused data compared to a threshold value of 4.5 (see TABLE V). This was found to be similar for all the data modalities (i.e., raw, individual and feature fusion) in this study. These results show that the accuracy can be improved further by tuning parameters associated with both feature extraction and classification models. The results bear out the relevance of peripheral physiological signals in emotional arousal and valence recognition or classification.

TABLE IV. COMPARISON OF RESULTS OBTAINED WITH PUBLISHED RESEARCH

Modality (after feature extraction)	**Valence**		**Arousal**	
	Accuracy	***F1***	***Accuracy***	***F1***
All [1]	62.70	0.6080	57.00	0.5330
All [13]	58.10	0.5750	62.70	0.5820
All [This work]	64.45	0.7588	63.20	0.7702
GSR [17]	71.04	-	71.53	-
RSP, BVP,Temp[14]	72.18	-	73.08	-
Temperature [14]	67.73	-	69.68	-
BVP [14]	70.23	-	68.59	-
RSP [14]	70.62	-	71.32	-

TABLE V. CLASSIFICATION RESULTS OBTAINED ON FEATURE FUSION DATA WITH THRESHOLDS 4.5 AND 5.0

Threshold	Valence		Arousal	
	Accuracy	*F1*	*Accuracy*	*F1*
5.0	56.01	0.6199	55.46	0.5159
4.5	64.45	0.7588	63.20	0.7702

The results overall show that emotions can be recognised from peripheral physiological signals. Tuning the hyperparameters of classification models can be investigated to further improve the recognition accuracy. Other factors to consider for future work on peripheral signals are investigating the following factors: output label (2-class vs 3-class or continuous label using regression techniques), feature extraction (window size, overlap vs non-overlap), feature selection methods, cross-validation(k-fold vs leave-one-out), including many relevant features based on other studies, decision fusion techniques, various group-based study, and advanced classification models (Deep and Convolutional neural networks). The studies on EEG signals [11] have shown improved recognition rate with advanced models and optimised feature selection.

V. CONCLUSION

Studies on EEG-based emotion recognition have progressed in classifying emotions using advanced learning techniques, very little work has been done on peripheral signals, something that is addressed in this paper. A wide range of standard and non-linear statistical features were extracted from peripheral physiological signals of the DEAP dataset. A total of eight machine learning techniques were evaluated, and the results show that SVM, LR and LDA models give the best performance results for valence and arousal classification achieving better results on all data combinations used in the experiments. While RF achieved the highest valence classification rate of 64.45% on feature fusion set and MLP classifier stands out for valence classification from EMG and BVP signal. The fusion of all the features extracted from peripheral signals is better than individual signals. The results obtained demonstrate an improvement in the recognition accuracy of emotional valence and arousal compared with other published results that use the DEAP database. Subject-dependency was also investigated, and the results obtained showed that male subjects achieved an average valence recognition accuracy that is significantly greater than that obtained by female subjects. Overall, the results demonstrate that emotion recognition is possible from peripheral physiological signals. Future work will look at optimising the feature extraction methods used to improve recognition performance and explore advanced neural network configurations.

ACKNOWLEDGMENT

This publication has resulted from research conducted with the financial support of the Athlone Institute of Technology President's Doctoral Scholarship 2019.

REFERENCES

1. S. Koelstra et al., "DEAP: A Database for Emotion Analysis; Using Physiological Signals," in *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, Jan.-March 2012.
2. Mohammad Soleymani, Jeroen Lichtenauer, Thierry Pun, and Maja Pantic, "A multimodal database for affect recognition and implicit tagging," *Affective Computing, IEEE Transactions on*, vol. 3, no. 1, pp. 42–55, 2012.
3. L. Shu et al., "A review of emotion recognition using physiological signals", *Sensors*, vol. 18, no. 7, pp. E2074, 2018.
4. J. O'Dwyer, "Speech, Head, and Eye-based Cues for Continuous Affect Prediction," *2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, Cambridge, United Kingdom, 2019, pp. 16-20.
5. Stanisław Saganowski, et al., "Emotion Recognition Using Wearables: A Systematic Literature Review Work in progress," arXiv:1912.10528v2 [cs.HC], Dec 2019.
6. E. Kanjo, E. M. Younis, and C. S. Ang, "Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection," *Information Fusion*, vol. 49, pp. 46–56, 2019.
7. J. Chen et al., "Subject-independent emotion recognition based on physiological signals: a three-stage decision method", *Bmc Medical Informatics & Decision Making*, vol. 17, no. Suppl 3, pp. 167, 2017.
8. A. Chatchinarat, K. W. Wong and C. C. Fung, "Fuzzy classification of human emotions using Fuzzy C-Mean (FCFCM)," *2016 International Conference on Fuzzy Theory and Its Applications (iFuzzy)*, Taichung, 2016, pp. 1-5.
9. L. He, D. Jiang, L. Yang, E. Pei, P. Wu, H. Sahli, "Multimodal affective dimension prediction using deep bidirectional long short-term memory recurrent neural networks", *Proc. 5th Int. Workshop Audio/Visual Emotion Challenge*, pp. 73-80, 2015.
10. F. Ringeval et al., "Prediction of asynchronous dimensional emotion ratings from audiovisual and physiological data", *Pattern Recognit. Lett*., vol. 66, no. Supp. C, pp. 22-30, 2015.
11. S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset", *Proc. IAAI*, pp. 4746-4752, 2017D.
12. Jing Chen, Bin Hu, Lixin Xu, P. Moore and Yun Su, "Feature-level fusion of multimodal physiological signals for emotion recognition," *2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Washington, DC, 2015, pp. 395-399.
13. S. Chen, Z. Gao and S. Wang, "Emotion recognition from peripheral physiological signals enhanced by EEG," *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, 2016, pp. 2827-2831.
14. D. Ayata, Y. Yaslan and M. Kamaşak, "Emotion Recognition from Multimodal Physiological Signals for Emotion Aware Healthcare Systems," *J. Med. Biol. Eng*, 2020.
15. E. Hynes, R. Flynn, B. Lee and N. Murray, "A Quality of Experience Evaluation Comparing Augmented Reality and Paper Based Instruction for Complex Task Assistance," *2019 IEEE 21st International Workshop on Multimedia Signal Processing (MMSP)*, Kuala Lumpur, Malaysia, 2019, pp. 1-6.
16. P. Salgado et al., "A QoE assessment method based on EDA heart rate and EEG of a virtual reality assistive technology system", *Proceedings of the 9th ACM Multimedia Systems Conference on - MMSys '18*, pp. 517-520, 2018.
17. D. Ayata, Y. Yaslan and M. Kamaşak, "Emotion recognition via random forest and galvanic skin response: Comparison of time based feature sets, window sizes and wavelet approaches," *2016 Medical Technologies National Congress (TIPTEKNO)*, Antalya, 2016, pp. 1-4.
18. C. Godin, F. Prost-Boucle, A. Campagne, S. Charbonnier, S. Bonnet, and A. Vidal, "Selection of the Most Relevant Physiological Features for Classifying Emotion," *PhyCS 2015*.
19. S. Christopher, "Nonlinear measures for dynamical systems (Version 0.5.2)," Zenodo, Jun. 16, 2019. http://doi.org/10.5281/zenodo.3814723.