Emotion Recognition Using Fused Physiological Signals

Diego Fabiano and Shaun Canavan

Department of Computer Science and Engineering, University of South Florida, Tampa, FL, USA

dfabiano@mail.usf.edu, scanavan@usf.edu

Abstract—In this paper, we propose a new representation of human emotion through the fusion of physiological signals. Using the variance of these signals, the proposed method increases the effect of signals that contribute to the recognition accuracy, while decreasing the effect of those that do not. The new representation is a powerful approach to recognizing emotions. We investigate this by comparing against emotion recognition results from non-fused physiological signals. Both the fused and non-fused signals are used to train feedforward neural networks to recognize a range of emotion. We show that the fused method outperforms each individual signal across all emotions tested. We test the efficacy of the proposed approach on two publicly available datasets, namely BP4D+ and DEAP, showing state-of-the-art results on both. To the best of our knowledge this is the first work to present emotion recognition results using physiological signals on all subjects from BP4D+.

Index Terms—fusion, physiological, affect, emotion recognition

I. INTRODUCTION

Affective Computing has been an exciting and growing field in the past two decades, due largely in part to the seminal work from Rosalind Picard [21]. The field has important applications in artificial intelligence, as being able to recognize emotion is an important part of human intelligence [24]. The ability to recognize emotion has broad impacts for real-world applications in fields as diverse as medicine, defense, entertainment, and retail. Some of these applications include pain recognition [41], customer feedback [4], and educational video games [16]. To move forward with developing these applications, we need to understand the foundation of autonomy, as well as advance interfaces between human and machines. To do this, we must first understand the role of emotion, including what exactly emotion is. This is a difficult problem as there are currently around 100 definitions of what emotion is [22].

In the past two decades, there has been lasting and notable work in analyzing emotions. Most notably these works have focused on using 2D [1], [37], [38] and 3D faces [7], [9], [17], thermal data [23], [32] and audio signals [10], [42] for this task. Along with these modalities, physiological data is another interesting modality. Lisettit and Nasoz [19] showed they could recognize emotions using the min, max, mean, and variance of physiological signals. Motivated by this, we propose a method for fusing physiological signals for emotion recognition. We hypothesize that the fusion of high variance signals will increase the performance of emotion recognition. Considering this, the proposed method increases the effect of high variance signals and decreases the effect of low variance signals (Fig. 1).

In contrast to previous methods, that have used physiological data for emotion recognition [14], [18], [20], [28], we fuse multiple signals into a new representation of emotion for each subject. As we will show, this new representation can be used to increase the overall emotion recognition accuracy when using physiological signals. As wearable devices become more and more commonplace, the proposed research has the potential to extend some of the broad, real-world applications to include real-time lie detection, analysis of stress levels, and prediction of autism in children. The main contributions of the proposed research can be summarized as follows:

1) We propose a method for fusing physiological signals that can be used for emotion recognition.
2) We validate the utility of the proposed approach by comparing against non-fused physiological signals for emotion recognition.
3) We detail an application of the proposed fusion method for pain recognition.
4) We show the superior performance, of the proposed method, compared to the state of the art on DEAP [14] and BP4D+ [44].

II. RELATED WORKS

In recent years, there has been interesting and exciting work done using physiological signals for emotion recognition. Koelstra et al. [14] developed a heterogeneous dataset of EEG and peripheral physiological signals, as well as subject self-rating. Using this data, they classified arousal and valence, and like and dislike ratings (based on self-rating). Rozgic et al. [29], proposed a method for emotion classification using EEG signals where they extract features from overlapping sequences of the signals from the DEAP dataset. Learning features from a deep belief network, Li et al. [18] extracted high-level EEG features for classification with an SVM. Wagner et al. [34] collected physiological data using four-channel biosensors. Emotion was elicited using a music induction method. They extracted hand-crafted features from the signals including breathing rate and amplitude of the signals. They found that it is easier to distinguish between emotion along the arousal axes compared to the valence axes. They experimented with multiple classifiers including multilayer perceptron and k-nearest neighbor. Yin et al. [40] trained an ensemble deep learning model with physiological signals for
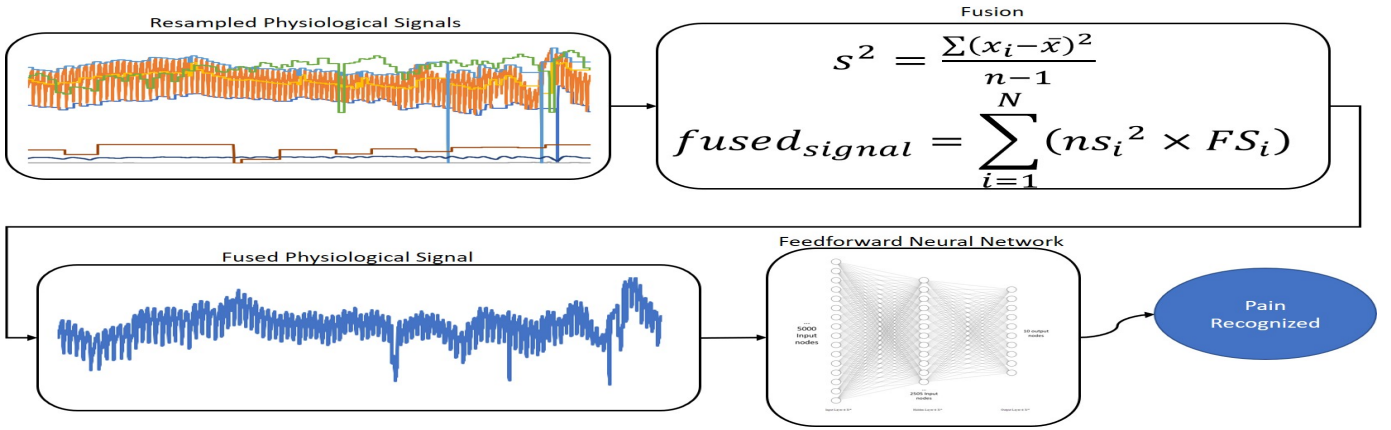
Fig. 1. Overview of proposed method. Fusion of 8 physiological pain signals (BP4D+ [44]) correctly recognized as pain..

emotion recognition. Using a stacked autoencoder approach, they derive stable feature representations. A separate deep model is then used for the stacked autoencoder ensemble. They found that this ensemble-based approach can lead to higher generalization capability compared to other shallow methods. Using a music-based method, Kim et al. [13], investigated changes in physiological signals for the task of emotion recognition. They collected data over multiple weeks to extract features from domains that include time, geometric, and the sub-band spectra. These features were used to classify four musical emotions along the arousal axes using an extended linear discriminant analysis.

Over the last decade, there has been a great deal of interesting work also done in the medical field using physiological signals, especially with deep learning methods [8]. Yang et al. [39], used a recurrent neural network to detect anomalies in heart sounds from acoustic physiological signals. They proposed a method to augment the signals by using Discrete Fourier Transform, where they include the variance from the window with the acoustic signals. Tan et al. [31], detected seizures using EEG signals. They developed a 13-layer Convolutional Neural Network (CNN) that was able to detect normal, preictal, and seizure classes with 88.67%, 90%, and 95% accuracy, respectively. They also used electrocardiography (ECG) signals to identify coronary artery disease by training a stacked CNN and Long Short-term Memory (LSTM) network. Using the proposed approach, they achieved a diagnosis accuracy of 99.85% from 47 subjects (7 with CAD, 40 normal).

Ragot et al. [27] investigated emotion recognition accuracies of lab sensors (Biopac MP150), vs. wearable sensors (Empatica E4). Their investigation showed similar accuracies between the two devices, showing emotion recognition is feasible in a real-world setting. Chen et al. [5] proposed a wearable healthcare system that collected physiological data and sent it to a cloud-based architecture to analyze the users health, along with their emotional state. This system was designed to help with emotional care deficiency (e.g. seniors quality of life, and empty nest syndrome [30]). Another study using cloud-based technology along with wearable devices was conducted by Zhang et al [43]. They proposed using the cloud to create a system for patient-centric healthcare. Using a robotics-assisted interface, they collected user data such as temperature and heart rate, to identify health risks, which in turn can be used to create a personalized health plan.

Zamzmi et al. [41], developed a multimodal approach to predict pain in infants. This approach used facial expressions, body movements, and physiological signals that include heart rate, respiration rate, and oxygen saturation levels. In their approach they extracted the mean value of each of these to train multiple classifiers such as KNN, SVM, and random forests. Using leave-one-subject-out cross validation, they achieved a max classification accuracy of 96%, when using the physiological signals to predict Neonatal Infant Paint Scores. They also predicted three states of pain (no pain, moderate pain, severe pain), achieving 82% accuracy.

Motivated by these works, the proposed approach is complimentary to both general emotion recognition, as well as the medical field as we detail results on pain recognition, as well as the prototypical emotions (e.g. sad, happy).

III. FUSION OF PHYSIOLOGICAL SIGNALS

We propose a new method to fuse physiological signals into a new signal that retains relevant temporal information. Given different physiological signals, some of them will have a different frame count, due to difference in data capture (e.g. data from BP4D+ [44]). Considering this, given signals of varying lengths, we first down-sample the signals to the same unit of time (i.e. same number of frames). This allows us to make direct comparisons between each of the signals. To do this, we first compute the ratio of the raw signal compared to the number of frames of data that we want to keep as ($ratio = original\_raw\_frame\_count/new\_frame\_count$). We then create a new frame by averaging ratio number of continuous frames resulting in exactly new_frame_count frames of data. In doing this, we effectively down-sample all signals to have the same sampling rate (Fig. 2).
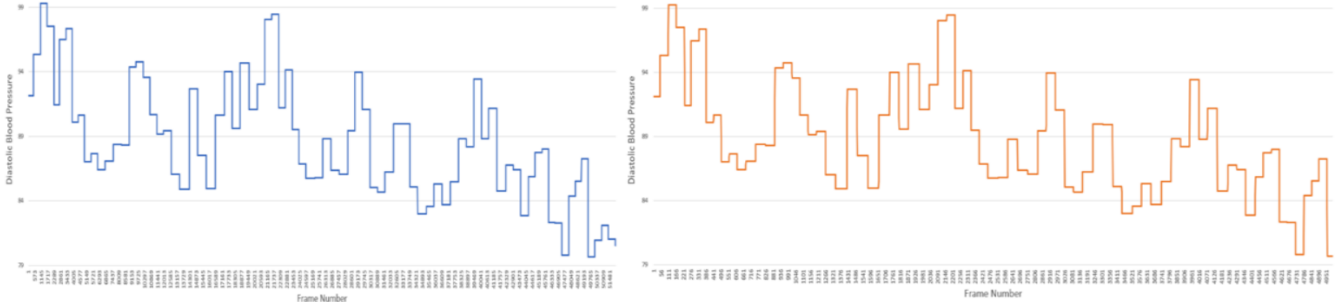
Fig. 2. Left side: original diastolic BP from BP4D+; right side: down-sampled diastolic BP. The original signal (left) has over 50,000 frames of data, while the resampled signal (right) has 5,000 frames of data, however, it still retains the original shape.

Given the down-sampled, raw signals, we then fuse each of them to create a new signal that retains the important emotion-based information from each of the fused signals. Our technique takes each of the signals, from each subject, for each emotion and fuses them (e.g. given a subject, each physiological signal with a Happy emotion is fused to create a new signal that represents Happy for that subject). When fusing the signals, we want to retain information that will result in higher emotion recognition accuracy. To investigate this, we first computed the variance, across all subjects, of each physiological signal as $S^2 = \dfrac{\sum (x_i - \overline{x})^2}{n-1}$. It is important to note that we have experimented with different statistical measures (e.g. entropy), and found no statistically significant difference in the resulting fused signals. We then compared this to the accuracy of each individual signal (non-fused) for emotion recognition (experimental design is detailed in section IV). In doing this, we found a trend that higher variance signals generally correspond to a higher accuracy in emotion recognition (Fig. 3).

Motivated by this trend, we use the variance to weight each signal during fusion. Given the variance for each signal type, we then normalize these values be in the range [min, max], which are the final variance values used to weight each signals importance. Each signal frame is multiplied by the weight, and then each signal is summed together as:

$$fused_{signal} = \sum_{i=1}^{N}(ns_i^2 \times FS_i). \qquad (1)$$

Where ns_i^2 is the normalized variance (i.e. weight), FS_i is the frame of the current signal being fused, and N is the total number of frames to be fused. This weighted fusion effectively boosts the high variance signals while dampening the low variance signals (Figs. 4 and 5). It should also be noted that the proposed fusion method will accurately follow and boost the directional trends of the original non-fused signals. For example, the overall trend of pain overtime (from BP4D+ data) is an increase in the signal. This is intuitive as the task to elicit pain in BP4D+ was for the subject to hold their hand in ice. The longer this task occurs, the more likely the subject is to be in pain. The fused signal takes the general trend of the original signals and boosts it (Figs. 4 and 5).
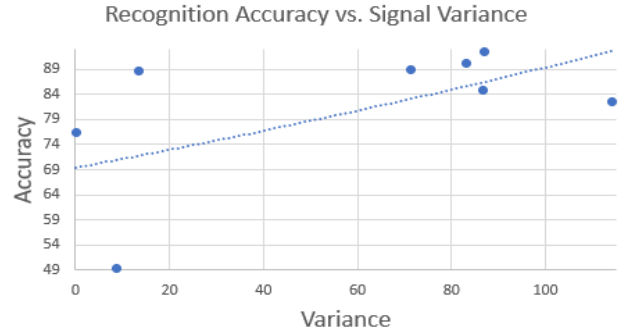

Fig. 3. Emotion recognition accuracy vs. variance of individual signals from BP4D+ [47]. Blue dotted line represents general trend.

IV. EXPERIMENTAL DESIGN AND RESULTS

A. Datasets

For our experiments, we used 2 state-of-the-art emotion-based datasets, namely BP4D+ [44], and DEAP [14]. Details on each of these is given below.

BP4D+ Dataset. BP4D+ is a large-scale, multimodal emotion dataset. It was used in the FERA challenge 2017 [33]. It consists of 140 subjects split between 58 male and 82 female subjects with ages ranging from 18-66. There is a total of 8 physiological signals that include blood pressure (diastolic, systolic, mean, and raw), respiration (rate and volts), heart rate, and electrodermal (EDA). Each subject contains data from 10 target emotions: happiness, sadness, anger, disgust, embarrassment, startled, skeptical, fear, pain, and surprise. The physiological signals from this dataset vary in length, therefore it is necessary for us to down-sample the data (Fig. 2). For our experiments, we fuse all physiological types (8 total), with weights of [0,1] (normalized variance as shown in Equation 1). In using these min and max weights, the signal with the lowest variance is removed, due to a weight of 0 (Fig. 4). We have empirically found that these weights work well for this data.

DEAP Dataset. DEAP is another multimodal emotion dataset. It contains 32 channels of electroencephalogram (EEG) signals based on the 10-20 system [11], as well as 8 physiological signals from 32 subjects (Fig. 5). The physi-
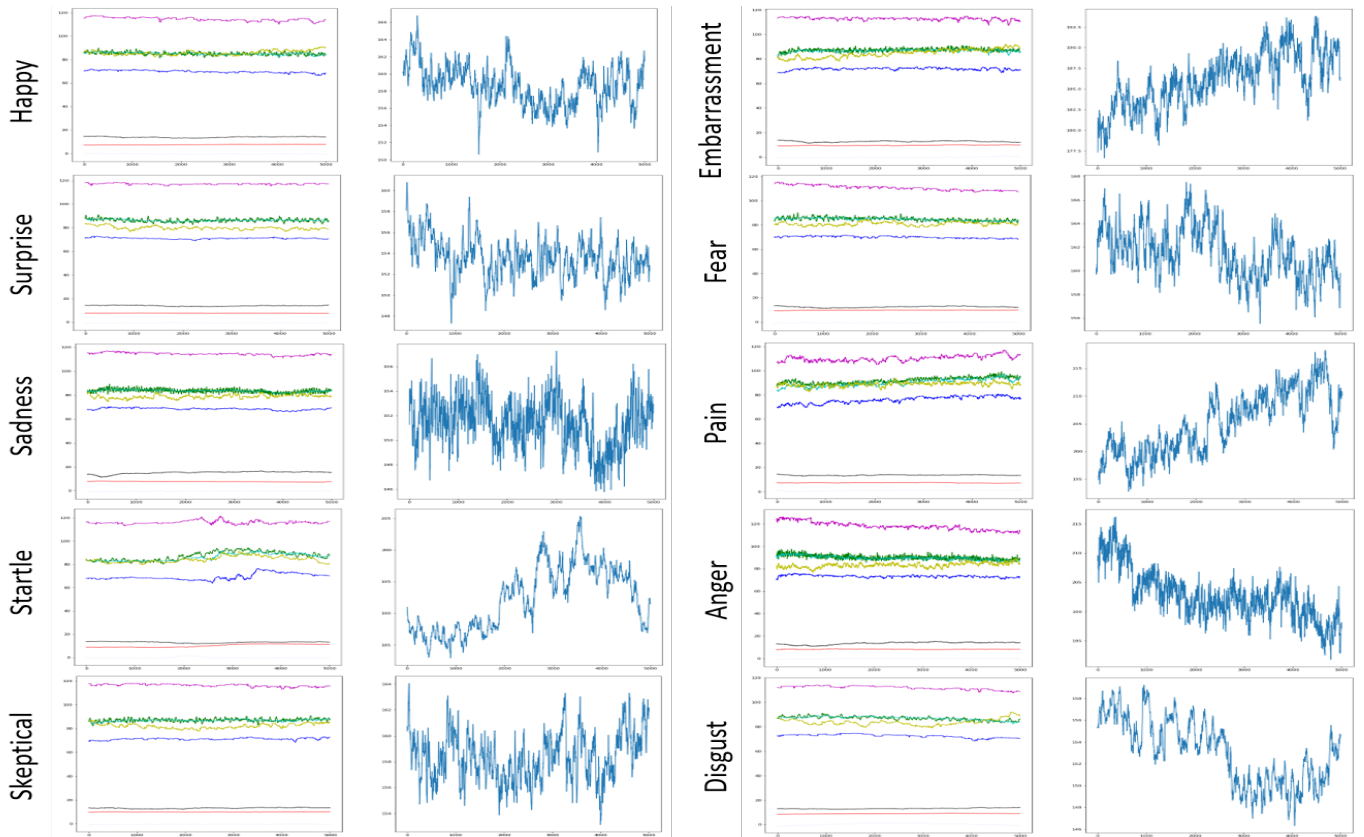
Fig. 4. Left side: Physiological signals, from a subject, for each of the 10 emotions in BP4D+ [44]. Right side: Fused signals from raw signals on left.
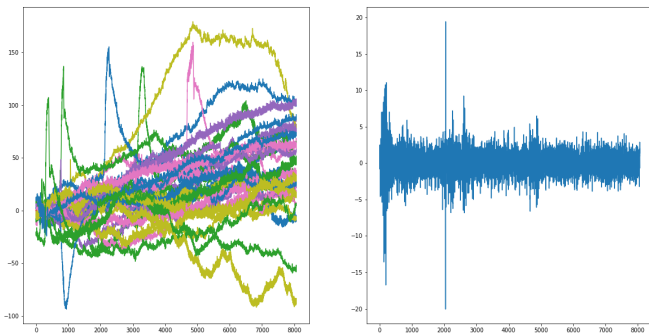

Fig. 5. Left: 32 EEG channels (subject in DEAP [14]); right: fused signal.

ological signals include horizontal Electrooculogram (hEOG), vertical Electrooculogram (vEOG), Zygomaticus Major Electromyogram (zEMG), Trapezius Electromyogram (tEMG), galvanic skin response (GSR), respiration, plethysmograph, and temperature. Along with the physiological signals, the dataset set also contains frontal face videos for 22 of the subjects. Each subject watched 40 one-minute music videos, which were selected using the affective tags that appeared on the last.fm website. For each of the videos, the subjects rated arousal, valence, like/dislike, and dominance/familiarity on a scale from [1-9]. Each of the signals consist of 8064 frames of data, therefore resampling is not needed for the signals in

DEAP. For our experiments we only focus on fusing the raw EEG signals, in this paper.

B. Feedforward Neural Network

We are motivated by the work from Han et al. [10], where they successfully used feedforward neural networks for emotion recognition from speech. Considering this, we used one for our experiments. The network is composed of an initial input layer that has the same number of neurons as the input vector, one hidden layer where the number of neurons = ⌊(number of neurons in input layer + number of neurons in output layer)/2⌋, and the final output layer output layer where the number of neurons = the number of classes to predict. The softmax activation function was used, and the adamax optimizer [12] with a learning rate of 0.001.

C. Results on BP4D+

To conduct our experiments on the BP4D+, we used 10-fold cross validation on both the fused signals, and individual signals (e.g. EDA). We randomly created 10-folds where 90% of the data was used for training, and 10% was used for testing. Each fold was used for testing, with it being independent from the training data. The average accuracy across each fold is reported. Along with our experiments on the fused signals, we conducted two experiments on the individual signals to test the validity of fusing the signals. First, we trained one neural network on all 8 signal types (Exp 1). Secondly, we
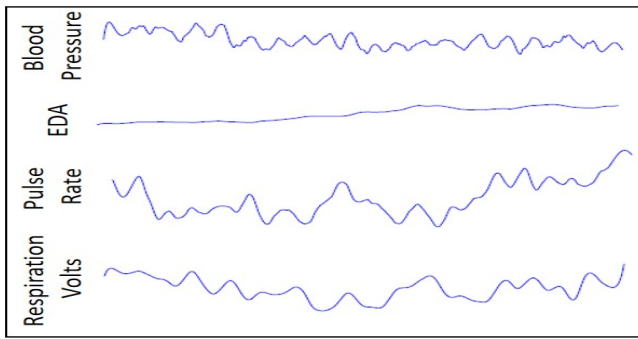
Fig. 6. Visual comparison of signal types (Happy emotion in BP4D+ [41]).

TABLE I
ACCURACY OF FUSED VS. NON-FUSED SIGNALS.

Emotion	Fused Accuracy	Exp 1	Exp 2
Anger	98.44%	81.67%	84.05%
Happy	93.18%	71.96%	79.93%
Fear	92.70%	67.71%	79.84%
Embarrassment	92.08%	62.29%	84.19%
Startle	92.03%	74.85%	84.92%
Pain	91.37%	53.78%	84.23%
Sad	90.78%	49.09%	86.55%
Surprise	90.21%	63.42%	78.21%
Skeptical	90.00%	52.59%	79.93%
Disgust	85.14%	62.06%	75.72%

TABLE II
CONFUSION MATRIX OF 10 EMOTIONS FROM BP4D+ [44]. KEY- HA:
HAPPY; SU: SURPRISE; SA: SADNESS; ST: STARTLE; SK: SKEPTICAL;
EM: EMBARRASSED; FE: FEAR; PA: PAIN; AN: ANGER; DI: DISGUST.

	HA	SU	SA	ST	SK	EM	FE	PA	AN	DI
HA	**.931**	.015	0	.008	.008	.015	0	0	.008	.015
SU	.007	**.902**	.021	0	.007	.014	.014	.007	0	0
SA	.014	.007	**.907**	.022	.014	.022	0	.014	0	0
ST	.007	0	.007	**.92**	.007	0	.015	.007	.015	.022
SK	.029	.007	0	0	**.9**	.014	.022	0	.014	.014
EM	0	0	.014	.008	.014	**.92**	.008	.014	.014	.008
FE	.021	0	.007	.014	.014	0	**.93**	.007	0	.007
PA	.008	.022	0	.008	.014	.021	.014	**.913**	0	0
AN	0	.008	0	0	0	0	0	.008	**.984**	0
DI	.028	.014	.028	.006	.006	0	.006	.021	.04	**.851**

TABLE III
COMPARISON OF FEEDFORWARD NEURAL NETWORK VS. CLASSICAL
MACHINE LEARNING ALGORITHMS WITH FUSED SIGNALS FROM BP4D+.

Classifier	Accuracy
Feedforward Neural Network	**91.59%**
Support Vector Machine	88.69%
Naïve Bayes	86.67%
Random Forest	86.17%

trained 8 different networks, one on each of the signal types (Exp 2). The average accuracy across each network was taken as the final report. Both of these experiments were conducted to compare the results of our fusion approach to non-fused physiological signals. For the fused, Exp 1, and Exp 2 experiments, using our feedforward neural network, we achieved an average accuracy (across all 10 emotions) of 91.59%, 63.93%, and 81.16%, respectively.

For all emotions, Exp 2 outperformed Exp 1 for emotion recognition accuracy. These results can be explained, in part, due to the large differences in signals (Fig. 6). In Exp 1, we trained one network on all 8 signal types. The network may have had difficulty in learning the correct features due to these differences. Although Exp 2 outperformed Exp 1, the fused signals outperformed both of them for all emotions (Table I). The lowest accuracy of fused signals is 85.14%, from disgust, which has a higher accuracy compared to all single signal experiments, except for sad from Exp 2, which had an accuracy of 86.55%. These results show the expressive power of the proposed fusion method for emotion recognition

For the fused signals, for many of the emotions, there are few misclassifications of the other emotions (Table II). For example, anger was misclassified as surprise and pain 0.8% of the time (1 misclassified signal each), and the rest of the signals were correctly recognized. Disgust was the lowest accuracy at 85.14% and was incorrectly recognized as all other emotions, at least once, except for embarrassed. It was misclassified as anger the most often at 4% (6 signals). This can partially be explained as disgust and anger are considered

similar expressions of condemnation [25].

Along with testing the validity of the fusion method, we also wanted to investigate if a feedforward neural network is the best approach to emotion recognition with fused signals. To investigate this, we conducted the same 10-fold cross validation experiment on a Random Forest [3], Support Vector Machine [6], and Naïve Bayes classifier [15]. Each classifier was trained with 90% of the signals, and the other 10% were used for testing. In this experiment, the neural network outperforms all three of the classical machine learning methods (Table III). However, its important to note that each of other methods still performed reasonably well, with a minimum of 86.71% accuracy achieved by a random forest. Along with validating the use of the network, this also shows the expressive power of the proposed method, as it can be used with a range of different classifier types.

Pain Recognition Application. An important and growing concern is the dependency of U.S. military on opioids, including the increase in opioid-related overdoses [2]. The overall quality of care can be improved by assessing pain. If it is left unmanaged it can lead to adverse outcomes, both physically and psychologically [35]. Motivated by this and the use of wearable devices to collect physiological data [7], we investigated using the proposed fusion method to accurately recognize pain (as found in BP4D+). Again, using 10-fold cross validation, we investigated our feedforward neural network, and the 3 previously investigated classical machine learning methods (Random Forest, Support Vector Machine, Naïve Bayes). We conducted experiments with 2 emotions (i.e. pain and no pain), where all fused physiological signals that were not labeled as Pain, were given the No Pain class.

Based on this experimental design, the proposed fusion method can distinguish between Pain and No Pain with a high degree of accuracy with the neural network (Table V). Less

Classifier	Accuracy
feedforward Neural Network	**98.48%**
Support Vector Machine	92.64%%
Random Forest	90.27%
Naïve Bayes	89.77%

TABLE V
CONFUSION MATRIX OF PAIN VS. NO PAIN.

	Pain	No Pain
Pain	0.984	0.016
No Pain	0.015	0.985

TABLE VI
COMPARISONS TO CURRENT STATE OF THE ART ON DEAP.

	Arousal	Valence	Liking	Dominance
Proposed Method	**95.27%**	**95.50%**	**96.03%**	**96.47%**
Liu et al [20]	80.50%	85.20%	82.40%	84.90%
Rozgic et al [28]	69.10%	76.90%	75.30%	73.90%
Li et al [18]	64.30%	58.40%	66.90%	65.80%
Koelstra et al [14]	63.10%	65.20%	64.20%	N/A

than 2% of the signals were misclassified as Pain, and the same results can be seen for No Pain. Similarly, the proposed fusion method can also be used with the classical methods for pain recognition (Table IV). This is encouraging for future real-time applications for recognizing pain in soldiers, and potential use in home-based pain management systems [25].

Comparison to State of the Art. Zhang et al [44], presented the first baseline using physiological signals from BP4D+, however, this was for 45 randomly selected subjects, not the entire dataset. Extracting hand-crafted features, they report the results from two experiments. First, testing on 5 emotions (happiness, sadness, startle, fear, and disgust) they report an accuracy of 59.5% using an RBF kernel SVM. Their second experiment was on all 10 emotions, where they perform a binary classification problem of low/high arousal. On this experiment, they report an accuracy of 60.5%. Again, this was only on 45 subjects from BP4D+, where we achieved an accuracy of 91.59% across all subjects and emotions.

D. Results on DEAP

To conduct our experiments on the DEAP dataset, we conducted single trial classification [14] on EEG data. We chose this approach, as it will allow us to make fair comparisons to current state of the art on this dataset. For these experiments, we did not use the classical machine learning approaches, as our evaluation in Section IV-C has shown the neural network to outperform all of the tested classifiers on BP4D+ data. Considering this, we only used our feedforward neural network. We investigated four binary problems. Namely, low/high arousal, low/high valence, low/high liking, and low/high dominance. As each signal was given a scale between 1 and 9, the threshold for the low and high classes was placed in the middle of the scale. This approach is the same as detailed previously on the DEAP dataset [14]. We fused all 32 channels of EEG data, using weights of [0,1], and performed single trial classification, for each subject.

For the investigated binary problems, we obtained an average accuracy (across all subjects) of 95.27%, 95.5%, 96.03%, and 96.47% for arousal, valence, liking, and dominance respectively. These results can be explained in part, by the nature of single-trial classification experiments. As they consider the variance within subjects, these experiments have been found to go beyond the study of the average brain [26]. We also compared our results to 4 state-of-the-art approaches from this dataset (Table VI). We outperform all compared by at least 10% across arousal, valence, liking, and dominance. We outperform the initial baseline [14], by approximately 30% across arousal, valence, and liking (dominance was not analyzed). This can be attributed to the ability of the proposed method to give accurate representations of a range of emotion.

V. CONCLUSION

We have presented a new method for fusing physiological signals for emotion recognition. The proposed method increases the influence of high-variance signals and decreases the influence of low-variance signals on emotion recognition. We tested the utility of the proposed method by comparing to all non-fused signals from BP4D+, showing the improved performance of the fused signals. To the best of our knowledge, this is the first work to present such work using physiological signals on all subjects from BP4D+. We also detailed state-of-the-art performance on DEAP EEG signals. The proposed method outperformed previous works by at least 10% on arousal, valence, liking, and dominance.

The proposed method has the ability to generalize across different types of physiological signals (e.g. blood pressure, heart rate, EEG). As the only constraint on the signals is that they have the same length, the proposed method has potential to be useful for representing a variety of signal types not discussed here such as heart signals (e.g. electrocardiography (ECG) and arterial pressure). Considering this, it has broad applications in pain recognition, stress analysis, lie detection, prediction of ASD in children, and potential to predict heart attacks from ECG data. With the increase of ubiquitous computing the proposed method can be used in real-time settings such as home health care systems, and for increasing solider survivability on the battlefield.

Although the proposed approach shows encouraging results on physiological data, emotion is a subjective experience that can be seen in more modalities than just physiological (e.g. image, thermal). This is one limitation of the current work, as only physiological data is used. Considering this, we will extend the work to fusion of multiple modalities including using variance to guide the fusion of deep features across multiple modalities.

ACKNOWLEDGMENT

REFERENCES

[1] I. Abbasnejad, S. Sridharan, D. Nguyen, S. Denman, C. Fookes, and S. Luccey. Using synthetic data to improve facial expression analysis with 3D convolutional networks. Computer Vision and Pattern Recognition, 2017.

[2] A. Bennett, L. Elliott, A. Golub, B. Wolfson-Stofko, and H. Guarino. Opioid-involved overdose among male Afghanistan/Iraq-era US military veterans: A multidimensional perspective. Substance Use and Misuse, 52(13): 1701-1711, 2017.

[3] L. Brieman. Random forests. Machine Learning, 45(1): 5-32, 2001.

[4] E. Cambria. Affective computing and sentiment analysis. IEEE Intelligent Systems, 31(2): 102-107, 2016.

[5] M. Chen, Y. Zhang, Y. Li, M. Hassan, and A. Alamri. AIWAC: Affective interaction through wearable computing and cloud technology, 22(1): 20-27, 2015.

[6] C. Cortes and V. Vapnik. Support-vector networks. Machine Learning, (20)3: 273-297, 1995.

[7] K. Divis, C. Anderson-Bergman, R. Abbott, V. Netwon, and G. Emannuel-Avina. Physiological and cognitive factors related to human performance during the Grand Canyon rim-to-rim hike. Journal of Human Performance in Extreme Environments, 14(1): 5, 2018.

[8] O. Faust, Y. Hagiwara, T. Hong, O. Lih, and U. Acharya. Deep learning for healthcare applications based on physiological signals: A review. Computer Methods and Programs in Biomedicine, 2018.

[9] W. Hariri, H. Tabia, N. Farah, A. Benouareth, and D. Declercq. 3D facial expression recognition using kernel methods on Riemannian manifold. Engineering Applications of Artificial Intelligence, 64, pp. 25-32, 2017.

[10] K. Han, Y. Dong, and I. Tashev. Speech emotion recognition using deep neural network and extreme learning machine. Fifth Annual Conference of the International Speech Communication Association, 2014.

[11] H. Jasper. The ten-twenty electrode system of the international federation. Electroencephalogram Clin. Neurophysiological, 10, pp. 370-375, 1958.

[12] D. Kingma and J. Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv: 1412.6980, 2014.

[13] J. Kim and E. Andre. Emotion recognition based on physiological changes in music listening. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(12): 2067-2083, 2008.

[14] S. Koelstra, C. Muhl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, and A. Patras. Deap: A database for emotion analysis using physiological signals. IEEE Transactions on Affective Computing, (3), pp. 18-31, 2012.

[15] I. Kononenko. Comparison of inductive and nave Bayesian learning approaches to automatic knowledge acquisition. Current Trends in Knowledge Acquisition, 8, p. 190, 1990.

[16] C. Lara, J. Flores, H. Mitre-Hernandez, and H. Perez. Induction of emotional states in educational video games through a fuzzy control system. IEEE Transactions on Affective Computing, 2018.

[17] H. Li, S. Jian, X. Zongben, and C. Liming. Multimodal 2D+3D facial expression recognition with deep fusion convolutional neural network. IEEE Transactions on Multimedia, 19(12):2816-2831, 2017.

[18] X. Li, P. Zhang, D. Song, G. Yu, Y. Hou, and B. Hu. EEG based emotion identification using unsupervised deep feature learning. 2015.

[19] C. Lisetti and F. Nasoz. Using noninvasive wearable computers to recognize human emotions from physiological signals. EURASIP jounral on Applied Signal Processing, pp. , 2004.

[20] W. Liu, W. Zheng, and B. Lu. Multimodal emotion recognition using multimodal deep learning. arXiv preprint arXiv 1602.08225, 2016.

[21] R. Picard. Affective Computing. MIT Press, 1995.

[22] R. Picard, S. Papert, W. Bender, B. Blumberg, C. Breazeal, D. Cavallo, T. Machover, M. Resnick, D. Roy, and C. Strohecker. Affective learning a manifesto. BT Technology Journal, 22(4): 253-269, 2004.

[23] C. Puri, L. Olson, I. Pavlidis, J. Levine, and J. Starren. StressCam: non-contact measurement of users emotional states through thermal imaging. CHI Extended Abstracts on Human Factors in Computing Systems, pp. 1725-1728, 2005.

[24] R. W. Picard, E. Vyzas, and J. Healey. Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(10):1175-1191, 2001.

[25] C. Molho, J. Tybur, E. Guler, D. Balliet, and W. Hofmann. Disgust and anger relate to different aggressive responses to moral violations. Psychological Science, 28(5): 609-619, 2017.

[26] C. Pernet, P. Sajda, and G. Rousselet. Single-trial analysis: why bother? Frontiers in Psychology, 2, pp. 322, 2011.

[27] M. Ragot, N. Martin, S. Em, N. Pallamin, and J. Diverrez. Emotion recognition using physiological signals: laboratory vs. wearable sensors. International Conference on Applied Human Factors and Ergonomics, 2017.

[28] G. Rousselet, C. Gasper, K. Wieczorek,a nd C. Pernet. Modelling single-trial ERP reveals modulation of bottom-up face visual processing by top-down task constraints (in some subjects). Frontiers in Psychology, 2 (137), 2011.

[29] V. Rozgic, S. Vitaladevuni, and R. Prasad. Robust EEG emotion classification using segment level decision fusion. IEEE Acoustics, Speech, and Signal Processing, 2013.

[30] P. Sullivan. Empty Nest Syndrome. Our Children, 23(9): 32-33, 1998.

[31] J. Tan, Y. Hagiwara, W. Pang, I. Lim, S. Oh, M. Adam, R. Tan, M. Chen, and U. Acharya. Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. Computers in Biology and Medicine, (94), pp, 19-26, 2018.

[32] L. Trujillo, G. Olague, R. Hammoud, and B. Hernandez. Automatic feature localization in thermal images for facial expression recognition. Computer Vision and Pattern Recognition Workshop, 2005.

[33] M. Valstar, E. S.-Lozano, et al., FERA 2017 Addressing head pose in the third facial expression recognition and analysis challenge, arXiv: 1702.04174, 2017.

[34] J. Wagner, K. Jonghwa, and E. Andre. From physiological signals to emotions: implementing and comparing selected method for feature extractions and classification. International Conference on Multimedia and Expo, 2005.

[35] N. Wells, C. Pasero, and M. McCaffery. Improving the quality of care through pain assessment and management, 2008.

[36] Y. Xu, I. Hubener, A. Seipp, S. Ohly, and K. David. From the lab to the real-world: An investigation on the influence of human movement on emotion recognition using physiological signals. Pervasive Computing and Communications Workshops, 2017.

[37] H. Yang, U. Ciftci, and L. Yin. Facial expression recognition by de-expression residue learning. Computer Vision and Pattern Recognition, 2018.

[38] H. Yang, Z. Zhang, and L. Yin. Identity-adaptive facial expression recognition through expression regreneration using conditional generative adversarial networks. Automatic Face and Gesture Recognition, 2018.

[39] T. Yang and H. Hsieh. Classification of acoustic physiological signals based on deep learning neural networks with augmented features. Computing in Cardiology Conference, 2016.

[40] Z. Yin, M. Zhao, Y. Wang, and J. Zhang. Recognition of emotions using multimodal physiological signals and an ensemble deep learning model. Computer methods and programs in biomedicine, 140, pp. 93-110, 2017.

[41] G. Zamzmi, C. Pai, D. Goldgof, R. Kasturi, T. Ashmeade, and Y. Sun. An approach for automated multimodal analysis of infants pain. International Conference on Pattern Recognition, 2016.

[42] S. Zhang, T. Huang, W. Gao, a Q. Tian. Learning affective features with a hybrid deep model for audio-visual emotion recognition. IEEE Transactions on Circuits and System for Video Technology, 2017.

[43] Y. Zhang, M. Qiu, C. Tsai, M. Hassan, and A. Alamri. Health-CPS: Healthcare cyber-physical system assisted by cloud and big data. IEEE Systems Journals, 11(1): 88-95, 2017.

[44] Z. Zhang, J. Girard, Y. Wu, X. Zhang, P. Liu, U. Ciftci, S. Canavan, M. Reale, A. Horowitz, H. Yang, J. Cohn, and L. Yin. Multimodal spontaneous emotion corpus for human behavior analysis. Computer Vision and Pattern Recognition, 2016.