

Affect Recognition from Physiological Signals using Autoregressive Hidden Markov Models

Fatma Patlar Akbulut^{1,2} *Member, IEEE*, Harry G. Perros¹ *Fellow, IEEE*,
Muhammad Shahzad¹ *Member, IEEE*

Abstract—Affect provides contextual information about the emotional state of a person as he/she communicates in both verbal and/or non-verbal forms. While human's are great at determining the emotional state of people while they communicate in person, it is challenging and still largely an unsolved problem to computationally determine the emotional state of a person. Emotional states of a person manifest in the physiological biosignals such as electrocardiogram (ECG) and electrodermal activity (EDA) because these signals are impacted by the peripheral nervous system of the body, and the peripheral nervous system is strongly coupled with the mental state of the person. In this paper, we present a method to accurately recognize six emotions using ECG and EDA signals and applying autoregressive hidden markov models and heart rate variability analysis on these signals. The six emotions include happiness, sadness, surprise, fear, anger and disgust. We evaluated our method on a comprehensive new data set collected from 30 participants. Our results show that our proposed method achieves an average accuracy of 94.6% in distinguishing across the 6 emotions.

Index Terms—Affect Recognition, Auto regressive hidden markov models, Machine learning, Linear discriminant analysis, Heart rate variability.

1 INTRODUCTION

AFFECT plays an important role in human life as it provides contextual information about the emotional state of a person as he/she communicates in both verbal and/or non-verbal forms. A person's mood greatly influences the way he/she communicates and behaves. It guides the decision-making processes in response to social cues and improves one's capacity to develop and nourish social connections. While human's have evolved to determine the emotional state of people while they communicate in person, it is challenging and still largely an unsolved problem to computationally determine the emotional state of a person using appropriate sensors. Automatic and computational determination of emotional state of a person finds many applications such as more natural interaction with computing devices, treatment of conditions such as schizophrenia, autism, and psychopathy, and many more.

Physiological biosignals such as electrocardiogram (ECG), electrodermal activity (EDA), and electroencephalogram (EEG) are impacted by the emotional state of a person because these signals are impacted by the peripheral nervous system of the body, and the peripheral nervous system is strongly coupled with the mental state of the person [1], [2]. The peripheral nervous system is composed of nerves that convey information from receptors to the brain. Autonomic nervous system (ANS) which is a sub part of this system controls involuntary movements in organism, especially in the functioning of internal organs [3]. It is composed of two parts: a sympathetic system and

a parasympathetic system. Each organ in human body is connected to nerves belonging to both these systems. When a person's emotional state changes, the sympathetic and/or parasympathetic systems activate, which lead to changes in the operational state of various internal organs, and these states can be measured using the biosignals we mentioned earlier. For example, when a person feels fear, his/her sympathetic system activates and the nerves of this system release epinephrine in the body, which leads to increase in heart rate, and this increase in heart rate can be measured in the ECG signals. Although the events that trigger human emotions often vary, the physiological signals and responses produced by the body are the same for everyone. This way, we can use the variations in such biosignals to identify the affective/emotional state of the person.

It is widely accepted in the literature that emotions can be classified into six main categories: happiness, sadness, surprise, fear, anger and disgust [4]. Different theoreticians have proposed diverse categorizations of emotions, shown in Table 1. Among these six emotions, happiness is the most important representative of positive feelings, and it emerges in situations where individuals are most comfortable and life experiences are most intense. It has been observed that individuals take actions to maximize the possibility of feeling positive emotions and minimize feeling negative emotions. For example, people strive to avoid situations that can lead to sadness or anxiety and search for social activities and/or hobbies that give a sense of happiness, pleasure, and excitement. Negative feelings such as, anger, fear, disgust and sadness give rise to the stress of various levels depending on the individual. The definition of stress varies depending on the discipline (psychology, biomedicine, or sociology), but in general it is defined as a collection of emotional, mental, physical, and behavioral responses when

- ¹ F. P. Akbulut, H.G. Perros, and M. Shahzad are with the Department of Computer Science, North Carolina State University, Raleigh, NC 27606, USA. e-mail: fpatlar@ncsu.edu, hp@ncsu.edu, mshahza@ncsu.edu
- ² F. P. Akbulut is also with the Department of Computer Engineering, Istanbul Kultur University, 34156 Istanbul, Turkey. e-mail: f.patlar@iku.edu.tr

an individual feels anxious.

Many studies have shown that heart rate variability (HRV) extracted from ECG is a significant metric in determining physiological and psychosocial conditions, such as depression, anxiety, stress, and panic [6], [7], [8]. It has been shown that HRV changes occur according to psychosocial factors. Hon and Lee reported the manifestation of emotional factors in the form of changes in HRV in [9]. Ewing et al. [10] used RR intervals to detect autonomic neuropathy in diabetic patients. Wolf [11], Kannel et al. [12], and Antoni et al. [13] studied the relationship between low HRV and high mortality [14]. These studies showed that there is a link between negative emotions and low HRV. In anxious and depressed people, HRV has been shown to be low [15]. The reason is that both the sympathetic and parasympathetic branches of the autonomic nervous system take part in regulating the heart rate (HR). Sympathetic nervous system activity increases HR and decreases HRV, whereas parasympathetic nervous system activity decreases HR and increases HRV [16].

Our survey of the research performed on the topic of affect recognition revealed many different methods including Hidden Markov Models (HMMs) [17], support vector machines (SVM) [18] [19] [20] [21] [22], k nearest neighbor algorithm (k-NN) [23] [24], and Fishers linear discriminant analysis (FLDA) [25]. Unfortunately, prior schemes suffer from the problems of low accuracy and limited dataset. In fact, most studies used the same datasets to evaluate their methods, which were collected in isolated and noise-free hospital and laboratory environments. The only study that is close to ours was proposed by Garcia et al. [26] that used Fisher kernels and autoregressive (AR) HMMs with DEAP (Dataset for Emotion Analysis using Physiological signals) [27]. DEAP contains EEG (electroencephalogram), EMG (electromyography), EOG (electrooculography), GSR (galvanic skin response), respiration, plethysmograph and temperature data from 32 participants. Unfortunately, their method provided accuracies of only 70%, 63% and 64% on Positive-Negative emotions, Pleasant-Unpleasant Valence, and Active-Passive Arousal, respectively. As we can see, in addition to relatively low accuracy, this work did not actually distinguish between the six emotions that we mentioned earlier.

In this paper, we present a new method that applies AR-HMMs on biosignals and very accurately recognizes the six affects mentioned earlier. The biosignals that we use include the EDA signal and HRV extracted from the ECG signal. We evaluate our method on a new data set collected from 30 people. Our results show that our proposed method achieves an average accuracy of 94.6% in distinguishing across the 6 emotions. AR-HMM based techniques are popular and they have been applied in many domains including the medical field. For example, [28] used AR-HMM on speech signals to detect words and syllable boundaries. Stanculescu et. al. [29] modelled the presence of blood-borne infections with an AR-HMM model with the objective of reducing the waiting period for sepsis analysis of premature babies. Other studies include estimating missing data for modeling functional MRI imaging data [30] and finding differentially expressed genes in tumors [31].

To summarize, in this paper, we make the following

three contributions.

- We have developed a user-independent affect recognition method that uses AR-HMM for electrodermal activity (EDA) features and HRV analysis for ECG features with a machine learning classifier.
- We show that using ECG feature in addition to EDA for affect recognition increases the accuracy significantly.
- We show that instead of analyzing emotions directly, first performing mood analysis and then performing emotion analysis increases affect recognition accuracy significantly.

The rest of the paper is organized as follows. In the next section, we discuss the related work on affect recognition. In Section III, we describe our proposed affect recognition method in detail. In Section IV, we extensively evaluate our proposed method on a real data set and present the results. Finally, in Section V, we conclude the paper.

2 RELATED WORK

Several prior studies exist on human emotion analysis and recognition from physiological and non physiological signals. For example, [17] targeted the recognition of four emotions (happiness, anger, sadness, and neutral state) using facial expressions and speech. To recognize emotions, the proposed method uses markers on the face in a video in conjunction with acoustic information extracted from speech recorded simultaneously with the video. The results showed that the two modalities used together gave higher accuracy. Wang and Guan also conducted a study on recognition of individuals emotional states from audiovisual signals [25]. They represented visual characteristics with Gabor wavelet features and the audio information was constituted with the extracted prosodic, formant frequency, and Mel-frequency Cepstral Coefficient (MFCC) features. The multinomial classifier of the proposed model achieved 82.14% accuracy on a dataset consisting of 400 videos of 6 different emotions.

In [18], a headband is developed to collect EEG signals, which are then utilized for emotion analysis. Five male subjects were assessed for pleasant, neutral, and unpleasant emotional states. This method achieved an accuracy of 66.7% using support vector machines. Soleymani et. al. [19] described a system to recognize arousal/valence states of individuals using features extracted from EEG data and pupillary reflexes. The 24 participants' EEG responses and eye-gaze data were recorded during the emotional videos watched by them. At the end of the follow-up, a survey and evaluation were requested. This study showed that gaze distance and pupillary responses provide important affective feedbacks and the proposed approach achieved 68.5% and 76.4% accuracy for classifying three labels of valence and arousal, respectively. In another study [26] investigating valence/arousal levels, emotional regression was computed using machine learning methods. Physiological signals (EEG, EMG, EOG, GSR, respiration rate, plethysmograph, and temperature) were mapped to a Fisher score-space-based five state Hidden Markov Model to train for every subject and discriminative regression was performed via support vector machines. The performance of proposed

TABLE 1
Emotion clusters created by different theorists [5].

Theorists	Year	Emotion Clusters	Basis for Inclusion
James	1884	Rage, Fear, Grief, Love	Bodily involvement
Ekman et. al.	1982	Anger, Fear, Sadness, Joy, Disgust, Surprise	Universal facial expression
Clynes	1977	Joy, Grief, Anger, Hate, Reverence, Love, Sex, Depicted	Emotional expressive behavior
Panskepp	1982	Rage, Fear, Panic, Expectancy	Hardwired
Plutchik	1980	Anger, Fear, Acceptance, Anticipation, Disgust, Joy, Surprise, Sadness	Relation to adaptive biological process
Izard	1971	Anger, Fear, Contempt, Disgust, Distress, Guilt, Interest, Joy, Shame, Surprise	Hardwired
Frijda	1986	Communication, Desire, Happiness, Interest, Wonder, Sorrow, Surprise	Forms of action readiness

multimodal techniques resulted between 65.39% and 71.08% accuracy.

The system proposed by Verma and Tiwary [20] estimates the valence, arousal, and dominance values of several emotions using EEG, GSR, EMG, and EOG data in the DEAP database. Following the implementation of Discrete Wavelet Transform for signal analysis, tests were performed for 4 different classifiers. The average accuracies turned out to be 81.45%, 74.37%, 57.74%, and 75.94% for support vector machine, multilayer perceptron, k-nearest neighbor, and meta-multiclass classifiers, respectively. In another study [23] that also used the DEAP database, the Frequency Cepstral Coefficient (FCC) method was applied to determine the features. This technique used the Kernel Density Estimation (KDE) along with the K-Nearest Neighbor classifier to classify Happy and Sad emotions. The FCC method achieved 90% accuracy, which was 10% higher compared to the accuracy achieved by KDE. Yin et. al. [32] applied deep learning to DEAP database. Their proposed method, multiple-fusion-layer based ensemble classifier of stacked autoencoder, recognizes arousal/valence plane characteristics with average accuracy of 83.61%. Atkinson and Campos [33] applied minimum-Redundancy-Maximum Relevance (mRMR) technique to DEAP database to increase the impact of feature selection as a signal preprocessing step. Genetic Algorithm metaheuristic reinforced support vector machines kernel classifier achieved 60.7% and 62.33% accuracies for classifying arousal and valence, respectively. Chen et. al. [34] proposed to use ontological models for representing EEG records of DEAP database. Proposed model achieved a 69.09% accuracy for arousal and 67.89% accuracy for valence. C4.5 binary classifier reached the highest performance in emotion recognition among other techniques such as SVM, Multilayer Perception, and k-Nearest Neighbor. Cheng et. al. [21] used feature fusion method on one-channel ECG data to detect negative emotions, and achieved an overall accuracy of 79.51%.

Zhong et. al. [22] proposed a framework for multi-modal emotion recognition using physiological and facial expression data streams, called Temporal Information Preserving Framework (TIPF). GSR, ECG, respiratory amplitude, and skin temperature are used to distinguish valence and arousal distribution of streams. The primary observation in this study is that the use of temporal information of physiological signals increases the accuracy of the recognition of affect. Chen et. al. [34] proposed a three-stage decision method to recognize four emotions based on physiological signals. The first stage removes the influence of individual differences by transforming mixed training into separate groups. In the second stage, four emotions are divided

into two emotion pools to reduce recognition complexity. In the third and last stage, a classifier is trained based on affect in each emotion pool. This three-staged decision approach gave the highest accuracy of 77.57% using k-NN, SVM, C4.5, Random forest, One-against-Rest, and One-against-One classification schemes. To make the comparison between prior work easy, we have summarized the prior work in Table 2.

In comparison with all prior work, the method that we propose in this paper uses EDA as the information source in addition to ECG and achieves a high accuracy of 94.6%, which is greater than the accuracies achieved by any prior scheme. Furthermore, we have not used the DEAP database because it is suitable for working only on the valence-arousal plane and does not contain data that one can use to evaluate the accuracy of one's scheme on detecting emotions. We generated our own database that contains data which can be used to evaluate emotion recognition schemes.

3 METHODOLOGY

This section we provide a detailed description of our method to accurately identify emotions. Our method consists of four stages. In the first stage, we preprocess the data where we remove the noise from the signals. In the second stage, we extract appropriate features that can be used to distinguish between different emotions. In the third stage, we apply linear discriminant analysis to reduce the number of features. Finally, in the last stage, we generate neural networks based classification models. Next, we describe these four stages in detail, but before that, we first describe how we collected physiological data from volunteers for the six emotions, namely calmness, fear, sadness, anger, disgust, and happiness.

3.1 Data Collection

We collected physiological signal data from 30 participants comprising 11 males and 19 females between the ages of 19 to 81 years. Before starting data collection session with any participant, we first asked the participant to sit down and relax for 10 minutes. After that, the participant put on headphones and watched six 1-minute videos in random order corresponding to the six emotions in our study. While transition from video for one emotion to the video of another, we kept a pause of 5 seconds for a quick rest. While the participant watched the videos, we recorded participant's psychophysiological response (more specifically, ECG and EDA signals) at a sampling rate of 290Hz using the Cardiovascular Disease Monitoring (CVDiMo) wearable system. Most participants sat through two sessions, and

TABLE 2
PERFORMANCE COMPARISON WITH EXISTING WORK ON AFFECT RECOGNITION

Study	Methods	Source	Dataset	Participant	Recognition	Accuracy Rate(%)
Busso et al. (2004) [17]	HMM, MFCC, Bimodal Classifier	Speech, Facial Expression	N/A	N/A	Anger, Sadness, Happiness, Neutral	89
Wang and Guan (2008) [25]	MFCC, Gabor Wavelet, PCA, Multiclassifier	Audiovisual Signals	N/A	N/A	Anger, Disgust, Fear, Happiness, Sadness, Surprise	82.14
Schaaff and Schultz (2009) [18]	Fourier Transform, SVM	EEG	N/A	5	Valence, Arousal	66.70
Soleymani et al. (2012) [19]	Modality fusion strategy, SVM	EEG, Pupillary Response, Gaze distance	N/A	24	Valence, Arousal	68.50 - 76.4
Garcia et al. (2013) [26]	AR, HMM, Fisher kernels	EEG, EMG, EOG, GSR, respiration, PPG, temperature	DEAP	32	Positive-Negative, Pleasant-Unpleasant Valence, Active-Passive Arousal	70 - 63 - 64
Verma and Tiwary (2014) [20]	Discrete Wavelet Transform, SVM	EEG, GSR, EMG, EOG,	DEAP	32	Valence, Arousal, Dominance	81.45
Cheng et al. (2017) [21]	HRV, SVM	ECG	Bio Vid Emo DB dataset	N/A	Negative Emotions	79.51
Lahane and Thirugnanam (2017) [23]	Kernel Density Estimation (KDE), k-NN	EEG	DEAP	32	Happy, Sad	90
Zhong et al. (2017) [22]	HRV, SCR detection, SVM	Facial expression, GSR, ECG, Respiration, Temperature	MAHNOB-HCI database	27	Arousal, Valance	70 - 73
Chen et al. (2017) [24]	FFT, k-NN, Random Forest	EEG, EOG, EMG, Temperature, BVP, Respiration, GSR	DEAP	32	Arousal, Valance	77.57 - 43.57
Yin et al. (2016) [32]	Deep Learning	EEG, EOG, EMG, Temperature, BVP, Respiration, GSR	DEAP	32	Arousal, Valance	84.18 - 83,04
Atkinson and Campos (2016) [33]	Genetic Algorithm supported SVM	EEG, EOG, EMG, Temperature, BVP, Respiration, GSR	DEAP	32	Arousal, Valance	60,7 - 62,33
Chen et al. (2015) [34]	C4.5 classifier	EEG	DEAP	32	Arousal, Valance	69,09 - 67,89
Our Model (2018)	HRV, AR-HMM, LDA, NN	ECG, EDA	CVDiMo (own)	30	Calm, Fear, Sadness, Disgust, Anger, Happiness	94.6

DEAP: Dataset for Emotion Analysis using Physiological signals

thus watched each of the twelve videos (we prepared two unique videos per emotion). Some of the participants only participated in the first session. On the completion of the data collection, we had a total of 312 records. The data collection was approved by the Istanbul Cerrahpasa Medical Faculty Hospital's (in Istanbul, Turkey) ethical committee on the use of humans as experimental subjects.

In the video intended to induce the calmness emotion, we showed the participants nature images and underwater shots. The effect was enhanced by the use of background music. In the video intended to induce the fear emotion, we showed the participants horror scenes from the the horror films *Ring* and *The Grudge*. In the video intended to induce the sadness emotion, we showed the participants emotional

parts of the famous Turkish drama films *Ekya* and *Babam ve Olum*. In the video intended to induce the anger emotion, we showed the participants a baby being beaten by his caregiver and people being beaten unfairly by the police. In the video intended to induce the disgust emotion, we showed the participants scenes of eating insects and dead animals by the actor Bear Grylls from the documentary *Man vs. Wild*. Finally, in the video intended to induce the happiness emotion, we showed the participants cute animal images.

3.2 Preprocessing

The signals obtained from the sensors may be distorted due to the movement of the body and other issues that occur

during the data collection. To eliminate such artifacts, a preprocessing step is needed.

The ECG signals were processed using the following three steps: noise reduction, RR interval detection, and interbeat interval outlier removal. In the noise reduction step, low-pass filters, median filters and discrete wavelet transforms (DWT) were applied. These reduce the high-frequency components so that we can obtain the low-frequency features, trends, and a constant value of the of the raw ECG in each window.

We applied a derivative operator on the RR intervals to suppress the emergence of the QRS complex, which reflects rapid changes in the signal. This allowed us to capture the slowly varying low frequency T, P, and U waves in the cardiac cycle. Next, the negative parts coming from the derivation process were eliminated and the signal was squared to increase nonlinearly the dominant peaks. In other words, the smaller values were further reduced and the larger values were increased to emphasize the R peak slope. Finally, we performed motion window integration to obtain wave form features in addition to the slope of the R peaks. A vital component in determining the cardiac cycle from the noise-reduced ECG signal is to find the RR intervals. This is often done by determining the location of each pulse using techniques, such as, linear filtering and nonlinear transformations, and rule-based approaches. The R peaks, which were determined as characteristic points, were determined physiologically on the basis of a rule that a new wave could not occur with $200ms$ of the most recent wave. The highest peak in that cycle was determined as R, after one cycle was captured. The peak point of each participant is automatically adapted using the adaptive threshold value determined according to each signal.

EDA signals are one of the most robust physiological signals in affect recognition because of the fact that responses are sympathetic-centered. However, the signals also contain a large number of artifacts like ECG, such as electrical line noise. The spectral distributions of these signals are in the range of $0.08 - 0.2Hz$. For this reason, in the first step, we used a third-order low-pass filter with a cut-off frequency of $0.3Hz$ to reduce noise and then applied normalization to calculate the change in the fundamental fluctuations typically found in skin conductivity measurements. Normalization also limits the variance due to differences in physiology between participants, in addition to the long term changes in the physiological signal over time such as ambient temperature, time of the day etc.

3.3 Feature Extraction

All signals have been calculated over equal length time intervals in order to avoid bias in duration-dependent metrics (such as the standard deviation). To this end, the first minute (which is the total duration of watching activity) of the measurements was used to derive the value of the features.

The EDA signal is comprised of two components, skin conductivity responses (SCR) and skin conductance level (SCL). SCL is the slower basal component of EDA. Therefore, we only used SCR values from the EDA signal and modeled the entire SCR signal using an AR-HMM and used the model parameters as features. From the EDA signal, we

obtained a total of 19 features. From the ECG signal, we performed HRV analysis and obtained a total of 17 features. Next, we first describe how we modeled the EDA signal using AR-HMM and after that explain the HRV analysis on the ECG signal. Table 3 summarizes the features that we have used.

3.3.1 AR-HMM Modeling of Emotions using EDA Signals

An HMM is a discrete-time Markov Chain, whose states are hidden. At each state, an observable output is produced according to a state-dependent distribution. In an autoregressive HMM (AR-HMM) the observable outputs are generated by an autoregressive model whose coefficients depend on the current state of the HMM. HMMs have been extensively used for the purpose of classification using three different approaches: a) given a sequence of observations and a number of different HMMs, determine the most likely HMM from which it came from; b) given a sequence of observations that came from a given HMM, determine the most likely sequence of states which gave rise to these observations; and c) given a sequence of observations, determine the most likely parameters of the HMM that gave rise to the sequence. In this paper, we will use the third approach to classify emotions.

Let us consider an AR-HMM with N states, a one-step transition matrix $A = [a_{ij}]$, and an initial probability vector $B = [b_i]$. Let $Y = y_1, y_2, \dots, y_T$ be a sequence of affect observations. In the AR-HMM model the observations follow an AR(p) model, whose parameters depend on the current state of the Markov Chain, i.e.,

$$Y_t = c_{i,0} + c_{i,1} + \dots + c_{i,p}Y_{t-p} + \varepsilon_t \quad (1)$$

where $i = 1, 2, \dots, N$ indicates the state of the Markov Chain and $\varepsilon_t \sim N(0, \Sigma)$.

We set the number of states of the Markov Chain to three in order to reflect the intensity of emotion experienced by the participant at three different levels, i.e., low, medium, and high. Using the partial autocorrelation function (PACF) of the time series, we fixed the order of the AR model to three. Thus, our model is an AR(3)-HMM(3). 1.

Figure 1 consist of three graphs. The first one gives the filtered EDA signal to be modelled. The second one depicts the conditional standard deviation of the model, and the third graph shows the smoothed states probabilities of the model. We note that in the first few seconds of the experiment, state 1 is floating around zero, and states 2 and 3 around 0.4 - 0.6. After that, some states increase to 0.9 - 1 and others decrease. Based on these observations we conclude that every emotion carries multiple stimulations within itself. For the most of the part of the smoothed state probabilities generated by the Markov algorithm clearly distinguishes the most probable state.

The hidden Markov chain parameters can be estimated using the Expectations Maximization (EM)[35] algorithm. In the maximization step, we re-estimate the parameters associated with each regime, the transition probability matrix and the smooth probabilities. During the estimation, the predictor coefficient C was calculated with respect to EDA segments. Initial model of the unknown state, each set of sequences is segmented into the maximum likelihood state sequence. That denotes if an observation is cataloged into

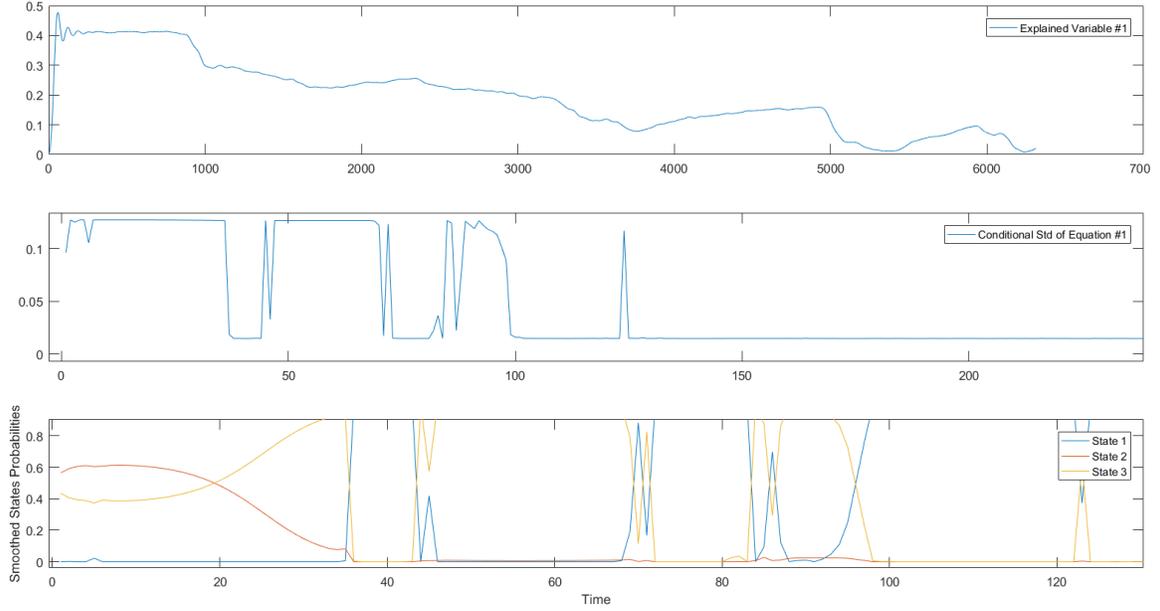


Fig. 1. Plot of the observed time series of EDA with conditional standard deviation and smoothed state probabilities modeled of the AR(3). (Subject 01, Emotion 02:Fear)

the i_t th state, we presume the corresponding state X_t be the i_t th state. We delegate each observation to the related state by corresponding its norm to a threshold.

This paragraph is also very unclear. I can't understand anything written here. Please elaborate it in more detail. To perform the gradual estimation, the Euclidean distance calculate between the AR coefficients of varied segment. The observation matrix $C = I, C_2, C_3, \dots, C_{n+1}$ where I is an identity vector and the column vectors C_2, C_3, \dots, C_{n+1} consist of coefficients $c(2), c(3), \dots, c(n+1)$ respectively of the skin response segments for estimation. The predictor C was computed at each stage of estimation based on the chosen sets. The procedure is iterated after replacing the old model with the new model. The steps assure an increase in likelihood after each iteration and will converge to a local optimum. The procedure for estimating the AR HMM model parameters, outlined in Algorithm 1.

Algorithm 1 Algorithm for Estimation of AR Model

Input: $X^* \leftarrow X_1^*, X_2^*, \dots, X_T^*$, {initial state}

$Lag \leftarrow Lag_{1:3}$, $State \leftarrow State_{1:3}$

Output: AR Model

Initialization : $\pi \leftarrow A$, $\lambda \leftarrow \pi$, A, B

1: **for** $i \rightarrow State$ **do**

2: **for** $j \rightarrow Lag$ **do**

3: $a_{ij} \leftarrow \frac{\#of\ transitions\ from\ state\ i\ to\ j}{\#of\ transitions\ from\ state\ i}$

4: $\mu_j^i \leftarrow \frac{\sum_{X_t=i} Y_{t,j}^{X_t}}{State_i}$ {mean vector in state i}

5: $C_j^i \leftarrow \frac{\sum_{X_t=i} (Y_{t,j} - \mu_j^{X_t})(Y_{t-1,j} - \mu_j^{X_{t-1}})}{\sum_{X_t=i} (Y_{t,j} - \mu_j^{X_t})^2}$
 {j_th element AR argument in state i}

6: $\Sigma \leftarrow \frac{\sum_{t=1}^T \epsilon_t}{T}$ {Covariance matrix of noise}

7: *EstimateModel* : $\hat{\lambda} \leftarrow \hat{\pi}, \hat{A}, \hat{B}$

8: *CalcSmoothedProb* : $P(X_t|Y), P(X_t, X_{t-1}|Y)$

9: *FindStateSeq* : X^*

10: $a_{ij} = \frac{\sum_{t=2}^T P(X_{t=j}, X_{t-1}=i|Y)}{\sum_{t=2}^T P(X_{t-1}=i|Y)}$ {Re-estimate}

11: **end for**

12: **end for**

13: **return** $ARModel$

After developing the AR-HMM model, we used the following values as features to classify the emotions: log-likelihood, standardized residuals, switching coefficients for the 3 lags (recall that we we used AR(3)), Akaike information criterion, Bayesian information criterion, transition probability, number of estimated parameters, estimated parameters in vector, standard errors of coefficients, smoothed probabilities of regimes, and covariance matrix.

3.3.2 Heart Rate Variability Analysis

Heart Rate Variability (HRV) is a physiological sign of cardiac autonomic activity and refers to the time series

magnitude of the standard deviation of cardiac periods in the pulse cycle [16]. HRV refers to the change in heart rate intervals which is also known as RR intervals. It has been used in many studies involving cardiovascular research and human health as an indirect tool to assess the functioning and balance of the autonomic nervous system [14]. HRV is also often used in emotion recognition research and shows different values when a person experiences different emotions.

The time between sequential heartbeats (RR interval) is the sum of the two sequences ($QRS_k = RR_i, RR_{i+1}, \dots, RR_n$ and $QRS_{k-1} = RR_{i+1}, RR_{i+2}, \dots, RR_{n+1}$) [14]. We calculated the average RR interval and standard deviations of one minute segments, the shortest, longest, and average RR interval, the segment number (pNN50) in which the successive difference of two RR intervals is more than 50 ms, the square root mean square difference of consecutive RR intervals, the HRV triangle index (HRVI) measurement, and the integrals of the RR interval frequency density distribution, and used all these values as features. To calculate the HRVI, the total range of all RR intervals was divided by the height of the density histogram of RR intervals measured on a separate scale with bins of 7.8125 ms (1/128 seconds). In addition to time frequency analysis, we performed a frequency spectrum analysis of the HRV with the power spectrum obtained by fast Fourier Transform. The ratios (LF/HF) of low (LF) and high frequency (HF) bands were obtained. We used the Poincar plot (SD1, SD2) from the analysis methods to evaluate the dynamics of HRV because it contains non-linear features of HRV. SD1 is an index that represents HRV in long-term recordings and reflects global variability, and SD2 is an index that shows the variability of pulse rate instantly and shows parasympathetic activity. The ratio (SD1/SD2) shows the ratio between the short and long variations of the RR intervals.

3.4 Feature Selection with Linear Discriminant Analysis

We as can see from the Table 3, the number of features that we extract per video record (recall that we have 312 records in total), is 36. For 36 features, to achieve good accuracy, we should have atleast 10 times as many records as the number of features times number of classes. Otherwise, due to the curse of dimensionality, the accuracy deteriorates. Unfortuntaely, obtain such a large number of records is prohibitive.

To overcome this problem, we employed linear discriminant analysis to reduces the high dimensionality due to the large number of features by merging features while preserving as much of the class discriminatory information as possible. The decrease in the number of features and thus the dimensionality further increases computational efficiency and reduces possibility of overfitting.

Next, we describe how LDA reduces the dimensionality. Given n features x_1, x_2, \dots, x_n that we want to reduce to m features, where $m < n$, as a first step we calculate an m -dimensional mean vector for each class in the dataset,

$$\mu_i = \frac{1}{N_i} \sum x \quad (2)$$

TABLE 3
List of Features

Source	Number	Feature
EDA	1	Log-likelihood
	2	Standardized Residuals from the model
	3	P value
	4	Akaike information criterion
	5	Bayesian information criterion
	6	Transition Probability for lag 1
	7	Transition Probability for lag 2
	8	Transition Probability for lag 3
	9	Mean
	10	Number of Observation
	11	Switching Betas(Coefficients) for lag 1
	12	Switching Betas(Coefficients) for lag 2
	13	Switching Betas(Coefficients) for lag 3
	14	Number of estimated parameters
	15	Estimated parameters in vector
	16	Standard errors of coefficients
	17	Smoothed probabilities of regimes
	18	Conditional Std
	19	Covariance matrix
ECG	20	Mean Heart Rate
	21	Minimum Heart Rate
	22	Maximum Heart Rate
	23	Global variability Index SD1
	24	Instant variability Index SD2
	25	SD Ratio
	26	High frequency (HF) bands
	27	Low frequency (LF) bands
	28	HF/LF Ratio
	29	Square root mean square difference (RMSSD)
	30	Mean RR
	31	Minimum RR
	32	Maximum RR
	33	Segment number successive difference >50 NN50
	34	Percentage of NN50 (pNN50)
	35	HRV triangle index (HRVI)
	36	RR interval standard deviations

Next, we compute the scatter matrices within and between-classes as follows:

$$S_w = \sum_{i=1}^c \sum_{x_k \in c_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad (3)$$

$$S_b = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (4)$$

After this, we calculate the eigenvectors and eigenvalues of the scatter matrices, and select n eigenvectors with the highest eigenvalues to shape a $m \times n$ dimensional matrix whose columns represent the eigenvectors.

$$S_w^{-1} S_b w = \lambda w \quad (5)$$

Next, we use the $m \times n$ eigenvector matrix to convert the samples to the new subspace and get the new feature vectors Y ,

$$Y = w^T x_k \quad (6)$$

where $k = 1, 2, \dots, N$ and w is a matrix with orthonormal columns. In our implementation, we used $m = 17$, i.e, we projected the 36 dimensional feature space onto 17 dimensional feature space, while maximizing distance between classes in the m dimensional space.

3.5 Neural Network Models

After reducing the number of dimension to seventeen (17), we then used the values of these parameters from our 312 records to train neural network classifiers and classify the emotions. Before training the neural network, we scaled the feature set to bring the values of all features in the range $[1, 1]$. The motivation behind doing this was to prevent features with larger values suppressing those with smaller values. We calculated the accuracy achieved by the neural networks using 10-fold cross-validation.

We observed from our experiments that when we tried to recognize the six emotions together, the accuracy was not as high as we desired. To overcome this problem, we divided the classification problem into two levels of hierarchy. The first level of classification only determine whether an emotion is positive or negative. Among our six emotions, calmness and happiness constitute positive emotions while fear, sadness, anger, and disgust constitute negative emotions. After identifying which category does the emotion belong to, we perform the second level of classification, where we determine exactly which emotion is represented by the unknown sample. Algorithm 2 described provides the pseudo code of the training process.

Algorithm 2 Algorithm for Hierarchical Training

Input: Feature $\leftarrow f_1, f_2, \dots, f_n$

Output: Model

Initialization : Epochs, Model, Kfold

Finding Long-Term State

```

1: for  $kl \rightarrow Kfold$  do
2:   for  $epoch \rightarrow Epochs$  do
3:     for  $input \in Feature$  do
4:        $m \leftarrow trainmodel(input, Model)$ 
5:     end for
6:   end for
7: end for
8: if  $(m \neq 0)$  then
9:   Finding Positive State
10:   $m_p \leftarrow trainmodel(input, m)$ 
11: end if
12: if  $(m \neq 1)$  then
13:   Finding Negative State
14:   $m_n \leftarrow trainmodel(input, m)$ 
15: end if
    {Evaluate model accuracy}
16: return Model

```

4 EXPERIMENTAL RESULTS

In this section, we first present an exploratory study on our data set that helps us understand the separability across different emotions using the features that we discussed in the previous section. After that, we present the results on the accuracy of our proposed scheme in identifying emotions. The experiments on evaluating the accuracy of our proposed method are conducted on three levels. At the first level, we simply classify whether an emotion was positive or negative. Recall that happiness and calmness belong to the

positive class while the remaining four emotions belong to the negative class. We call this positive or negative categorization of emotions the long term affective state. At the second level, find out exactly which emotion is experienced by the participant without using the results from the first level of classification. Finally, at the third level, we perform hierarchical classification, where we first classify a given sample into either positive or negative effective state and then further identify the exact emotion within that affective state. Figure 2 summarizes this hierarchical approach to emotion classification. All the algorithms used were implemented in

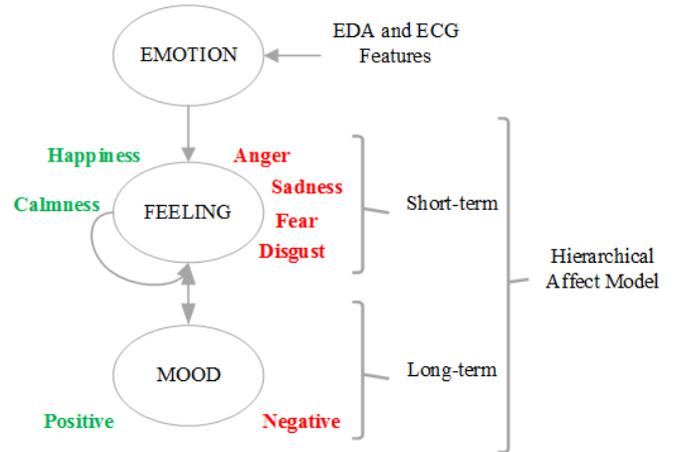


Fig. 2. Hierarchical model of affect

MatLab and the experiments were conducted using a high-performance computing (HPC) facility in North Carolina State University and also on a personal computer equipped with a 64-bit Intel (i7-7700 HQ) 2.8 GHz processor and 16 GB DDR4 RAM.

4.1 Exploratory Study

In this section, we try to understand the separability of the emotions through unsupervised clustering. For clustering, we use k-means algorithm. If we observe a good separation across clusters obtained using this unsupervised method, then our proposed method in the previous section should be able to distinguish across the six emotions. Recall that we have 312 records and each record has 36 features (before applying the LDA). As we have 6 emotions, we used $k = 6$ when applying k means clustering.

To display the results of clustering, we choose silhouette plots. A silhouette plot is a representation of a clusters that highlights the tightness of the clusters as well as their separation. A silhouette plot obtained from k means clustering contains k solhouettes. **Fatma, please elaborate the rest of this paragraph in blue color more detail. It is very unclear** The average silhouette width, which is an important ratio, maximizes the distance between the clusters and reduces the distance within the head. This average silhouette can be used to select an appropriate number of K sets, as well as provide an assessment of cluster validity. Silhouette points evaluation for cluster was realized as: ≤ 0.19 bad separation, $0.20 - 0.49$ poor separation, $0.50 - 0.69$, reasonable separation, and $0.70 - 1.00$ excellent separation. Silhouette coefficients close to 1 indicate that the sample is far from

neighboring clusters and the assignment is very accurate. However, a value close to 0 means to be very close to another cluster boundary. Sometimes these coefficients can be produced negatively, meaning that they are assigned to the wrong clusters.

Figure 3 shows a silhouette plot of the clusters obtained from the 19 EDA features and a scatter plot. (Fatma, which three features are plotted here) We observe from this figure that most points in all the clusters have low silhouette values (< 0.2) and some even have negative values. This shows that the clusters are not well separated when using only EDA based features. Similarly, the scatter plot does not show any visible clusters either. This implies that using AR-HMM features alone will not provide a high accuracy.

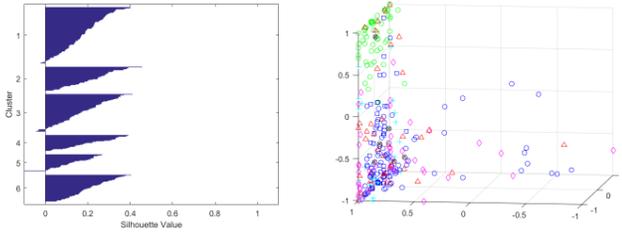


Fig. 3. Affect feature clusters a) silhouette plot of EDA features b) cluster plot of EDA features

Figure 4 shows a silhouette plot of the clusters obtained from both 19 EDA and 17 HRV features and a scatter plot. (Fatma, which three features are plotted here) We observe from this figure that most points in all the clusters have higher values (between 0.3 and 0.6) compared to what we saw in Figure 3, but are still not very high, indicating that while some clusters exist, they are not very well separated. The scatter plot in Figure 4 also visually shows the presence of some clusters, but the clusters are not very well separated.

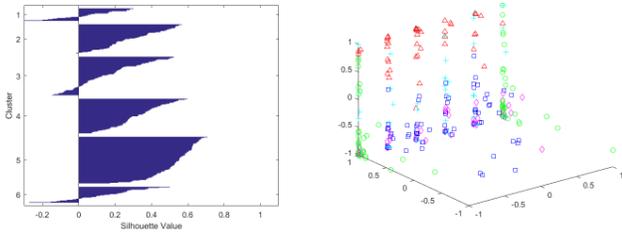


Fig. 4. Affect feature clusters a) silhouette plot of EDA and ECG features b) cluster plot of EDA and ECG features

Figure 4 shows a silhouette plot of the clusters obtained after applying LDA on both the EDA and HRV features with the total number of 17 features and a scatter plot. (Fatma, which three features are plotted here) We observe from this figure that most points in the third, fourth, and fifth clusters have large silhouette values (> 0.8), which means that these clusters are very well separated from the neighboring clusters. The first, second, and sixth clusters also contain many points with large silhouette values and the remaining having values between 0.4 and 0.5, which when looked in conjunction with the third, fourth, and fifth clusters turn out to be reasonably well separated as well. In the scatter plot in Figure 4, we observe quite distinct clusters

visually, which also implies that after applying LDA on both EDA and HRV features, we should be able to achieve high accuracy in classifying the emotions.

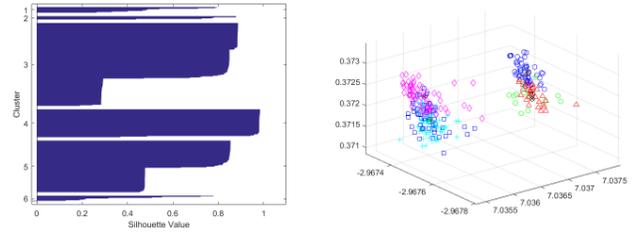


Fig. 5. Affect feature clusters with LDA a) silhouette plot of EDA and ECG features b) cluster plot of EDA and ECG features

4.2 Classification of Long Term Affect State

In this section, we present the accuracy of our proposed method in classifying whether an emotion was positive or negative. Recall that happiness and calmness belong to the positive class while the remaining four emotions belong to the negative class. Classification at this level can be useful in tele-health systems to identify if a patient has been experiencing negative emotions, so that a doctor can take appropriate steps to mitigate any negative impacts of the persistent negative emotions. We performed 10-fold cross validation on our data set, where the entire data set was divided into two classes, positive and negative. Our proposed method obtained an accuracy of 93.9% when using only the EDA based features and 94.2% when using only the ECG based features. We define accuracy for this set of evaluation as the number of correct assessments divided by number of all assessments. In other words, Accuracy = (TN + TP)/(TN+TP+FN+FP), where TN, TP, FN, and FP stand for True Negative, True Positive, False Negative, and False Positive. Figure 6 shows the confusion matrices obtained using only EDA features and using the 17 features obtained after applying LDA.

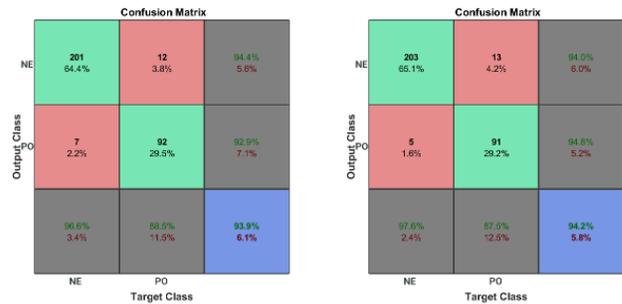


Fig. 6. Confusion Matrix of Long Term State Classification a) with AR model b) with AR model and HRV features

4.3 Classification of Short Term Affect State

4.3.1 Direct Classification

In this section, we present the accuracy of our proposed method in classifying the six emotions directly. Identifying

the emotional state of a person can be useful in many applications such as to identify offenders. We again performed 10-fold cross validation on our data set, where this time, the entire data set was divided into six classes, corresponding to the six emotions. Our proposed method achieved an accuracy of just 64.4% when using the EDA features and 89.1% using the features obtained after applying LDA. When using the features obtained after LDA, both positive emotions were recognized with an accuracy of 96.2%. The emotion with the lowest accuracy was sadness, that was recognized correctly 78.8% of the times. Figure 7 shows the confusion matrices obtained using only EDA features and using the 17 features obtained after applying LDA.

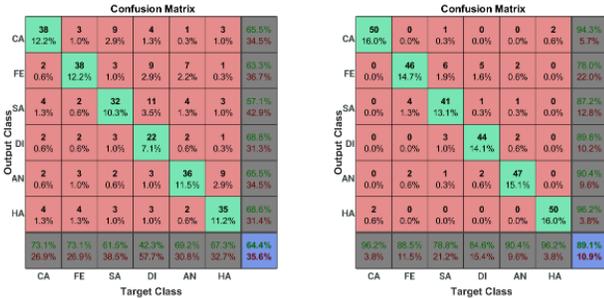


Fig. 7. Confusion Matrices of Short Term State Classification a) with AR model features b) with AR model and HRV features

4.3.2 Hierarchical Classification

Hierarchical Classification works in two stages. In the first stage, we identify whether an emotion is positive or negative. In the second stage, two different classifiers are used separately, one for positive emotions and the other for negative emotions. If an emotion is classified as negative by the first stage, the second stage uses the negative emotion classifier and classifies the emotion as one of the four negative emotions. Similarly, if an emotion is classified as positive by the first stage, the second stage uses the positive emotion classifier and classifies the emotion as one of the two positive emotions.

Figure 8 shows the confusion matrix for the detection of the two positive emotions. We observe from this figure that the positive emotion classifier achieved an accuracy of 98.9 in classifying the two positive emotions. As the first phase achieved 100% accuracy on identifying positive emotions as positive, our proposed scheme achieved an overall accuracy of 98.9% in classifying the two positive emotions, viz., happiness and calmness.

Figure 9 shows the confusion matrix for the detection of the four negative emotions. We observe from this figure that the negative emotion classifier achieved an accuracy of 96.1% in classifying the four negative emotions when using the 17 features obtained after applying LDA. As the first phase achieved 93.9% accuracy on identifying negative emotions as negative, our proposed scheme achieved an overall accuracy of 90.2% in classifying the four negative emotions, viz., fear, anger, surprise, and disgust.

The high accuracy of our proposed scheme results from the use of LDA in reducing the features as well as using the hierarchical approach of performing the classification

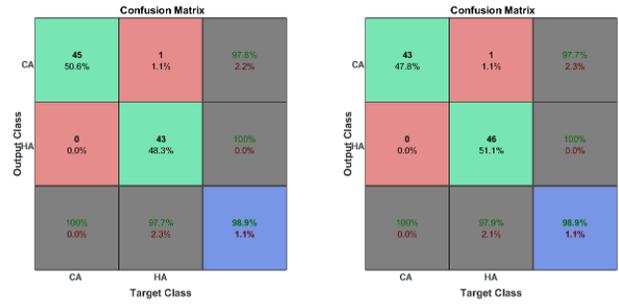


Fig. 8. Confusion Matrices for Hierarchical Classification of Positive Emotions a) with AR model features b) with AR model and HRV features Class labels are now included

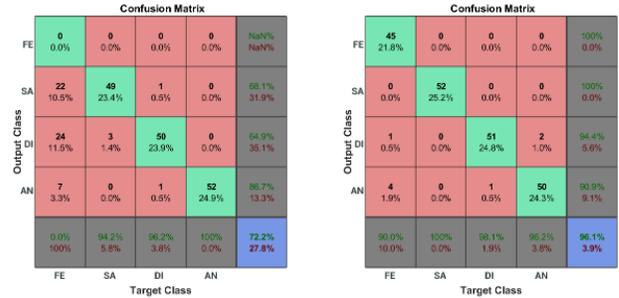


Fig. 9. Confusion Matrices Hierarchical Classification of Negative Emotions a) with AR model b) with AR model features and HRV features

of emotions. The accuracy values for the models with and without LDA are shown in Table 4.

TABLE 4
Effect of LDA on recognition accuracy

	Positive Emotions	Negative Emotions	Overall Accuracy
without LDA	94.2%	83.6%	88.9%
with LDA	98.9%	90.23%	94.6%

5 CONCLUSION

In this paper, we have presented a new and accurate method to recognize six different emotions using EDA and ECG signals. The key technical depth of the paper is in the use of the AR-HMMs to model the EDA signal and the use of LDA to enable accurate emotion recognition without requiring a large number of training samples. We have also presented an exploratory analysis of our data set that develops insights into the high classification accuracy of our approach in recognizing emotions. Unlike other studies, we have taken a hierarchical approach to classify emotions, where we first categorize the emotion as either positive or negative and then identify the exact emotion. Our experimental results indicate that with the use of this hierarchical method, our proposed system achieved an average accuracy of 94.6%

ACKNOWLEDGMENT

The authors would like to thank the members of the Department of Cardiology, Cerrahpasa Medical Faculty, Istanbul,

Turkey, for their help and suggestions in collect CVDiMo dataset.

Ethical approval: "All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards."

REFERENCES

- [1] M. Mumenthaler and H. Mattle, *Fundamentals of neurology: an illustrated guide*. Thieme, 2017.
- [2] S. Blain, T. Chau, and A. Mihailidis, "Peripheral autonomic signals as access pathways for individuals with severe disabilities: a literature appraisal," *The Open Rehabilitation Journal*, vol. 1, no. 11, pp. 27–37, 2008.
- [3] J. Haase, "Autonomic nervous system," in *Neurosurgery*. Springer, 2010, pp. 469–470.
- [4] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of personality and social psychology*, vol. 17, no. 2, p. 124, 1971.
- [5] O. Andrew, C. G. L., and C. Allan, "The cognitive structure of emotions," 1988.
- [6] E. Klein, E. Cnaani, T. Harel, S. Braun, and S. A. Ben-Haim, "Altered heart rate variability in panic disorder patients," *Biological Psychiatry*, vol. 37, no. 1, pp. 18–24, 1995.
- [7] I. Kawachi, D. Sparrow, P. S. Vokonas, and S. T. Weiss, "Decreased heart rate variability in men with phobic anxiety (data from the normative aging study)," *The American journal of cardiology*, vol. 75, no. 14, pp. 882–885, 1995.
- [8] S. Balogh, D. F. Fitzpatrick, S. E. Hendricks, and S. R. Paige, "Increases in heart rate variability with successful treatment in patients with major depressive disorder," *Psychopharmacology Bulletin*, 1993.
- [9] E. Horn and S. Lee, "Electronic evaluations of the fetal heart rate patterns preceding fetal death: further observation," *Am J Obstet Gynecol*, vol. 87, pp. 824–826, 1995.
- [10] D. J. Ewing, C. N. Martyn, R. J. Young, and B. F. Clarke, "The value of cardiovascular autonomic function tests: 10 years experience in diabetes," *Diabetes care*, vol. 8, no. 5, pp. 491–498, 1985.
- [11] M. Wolf, G. Varigos, D. Hunt, and J. Sloman, "Sinus arrhythmia in acute myocardial infarction," *The Medical journal of Australia*, vol. 2, no. 2, pp. 52–53, 1978.
- [12] W. B. Kannel, C. Kannel, R. S. Paffenbarger, and L. A. Cupples, "Heart rate and cardiovascular mortality: the framingham study," *American heart journal*, vol. 113, no. 6, pp. 1489–1494, 1987.
- [13] M. L. Antoni, H. Boden, V. Delgado, E. Boersma, K. Fox, M. J. Schlij, and J. J. Bax, "Relationship between discharge heart rate and mortality in patients after acute myocardial infarction treated with primary percutaneous coronary intervention," *European heart journal*, vol. 33, no. 1, pp. 96–102, 2011.
- [14] A. J. Camm, M. Malik, J. Bigger, G. Breithardt, S. Cerutti, R. J. Cohen, P. Coumel, E. L. Fallen, H. L. Kennedy, R. E. Kleiger *et al.*, "Heart rate variability: standards of measurement, physiological interpretation and clinical use. task force of the european society of cardiology and the north american society of pacing and electrophysiology," *Circulation*, vol. 93, no. 5, pp. 1043–1065, 1996.
- [15] J. M. Gorman and R. P. Sloan, "Heart rate variability in depressive and anxiety disorders," *American heart journal*, vol. 140, no. 4, pp. S77–S83, 2000.
- [16] G. G. Berntson, J. Thomas Bigger, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone *et al.*, "Heart rate variability: origins, methods, and interpretive caveats," *Psychophysiology*, vol. 34, no. 6, pp. 623–648, 1997.
- [17] C. Busso, Z. Deng, S. Yildirim, M. Bulut, C. M. Lee, A. Kazemzadeh, S. Lee, U. Neumann, and S. Narayanan, "Analysis of emotion recognition using facial expressions, speech and multimodal information," in *Proceedings of the 6th international conference on Multimodal interfaces*. ACM, 2004, pp. 205–211.
- [18] K. Schaaff and T. Schultz, "Towards emotion recognition from electroencephalographic signals," in *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*. IEEE, 2009, pp. 1–6.
- [19] M. Soleymani, M. Pantic, and T. Pun, "Multimodal emotion recognition in response to videos," *IEEE transactions on affective computing*, vol. 3, no. 2, pp. 211–223, 2012.
- [20] G. K. Verma and U. S. Tiwary, "Multimodal fusion framework: A multiresolution approach for emotion classification and recognition from physiological signals," *NeuroImage*, vol. 102, pp. 162–172, 2014.
- [21] Z. Cheng, L. Shu, J. Xie, and C. P. Chen, "A novel ecg-based real-time detection method of negative emotions in wearable applications," in *Security, Pattern Analysis, and Cybernetics (SPAC), 2017 International Conference on*. IEEE, 2017, pp. 296–301.
- [22] B. Zhong, Z. Qin, S. Yang, J. Chen, N. Mudrick, M. Taub, R. Azevedo, and E. Lobaton, "Emotion recognition with facial expressions and physiological signals," in *Computational Intelligence (SSCI), 2017 IEEE Symposium Series on*. IEEE, 2017, pp. 1–8.
- [23] P. Lahane and M. Thirugnanam, "A novel approach for analyzing human emotions based on electroencephalography (eeg)," in *Power and Advanced Computing Technologies (i-PACT), 2017 Innovations in*. IEEE, 2017, pp. 1–6.
- [24] J. Chen, B. Hu, Y. Wang, P. Moore, Y. Dai, L. Feng, and Z. Ding, "Subject-independent emotion recognition based on physiological signals: a three-stage decision method," *BMC medical informatics and decision making*, vol. 17, no. 3, p. 167, 2017.
- [25] Y. Wang and L. Guan, "Recognizing human emotional state from audiovisual signals," *IEEE transactions on multimedia*, vol. 10, no. 5, pp. 936–946, 2008.
- [26] H. F. Garcia, A. A. Orozco, and M. A. Alvarez, "Dynamic physiological signal analysis based on fisher kernels for emotion recognition," in *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*. IEEE, 2013, pp. 4322–4325.
- [27] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18–31, 2012.
- [28] J. D. Bryan and S. E. Levinson, "Autoregressive hidden markov model and the speech signal," *Procedia Computer Science*, vol. 61, pp. 328–333, 2015.
- [29] I. Stanculescu, C. K. Williams, and Y. Freer, "Autoregressive hidden markov models for the early detection of neonatal sepsis," *IEEE journal of biomedical and health informatics*, vol. 18, no. 5, pp. 1560–1570, 2014.
- [30] S. Dang, S. Chaudhury, B. Lall, and P. K. Roy, "Autoregressive hidden markov model with missing data for modelling functional mr imaging data," in *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*. ACM, 2016, p. 93.
- [31] M. Seifert, K. Abou-El-Ardat, B. Friedrich, B. Klink, and A. Deutsch, "Autoregressive higher-order hidden markov models: exploiting local chromosomal dependencies in the analysis of tumor expression profiles," *PloS one*, vol. 9, no. 6, p. e100295, 2014.
- [32] Z. Yin, M. Zhao, Y. Wang, J. Yang, and J. Zhang, "Recognition of emotions using multimodal physiological signals and an ensemble deep learning model," *Computer methods and programs in biomedicine*, vol. 140, pp. 93–110, 2017.
- [33] J. Atkinson and D. Campos, "Improving bci-based emotion recognition by combining eeg feature selection and kernel classifiers," *Expert Systems with Applications*, vol. 47, pp. 35–41, 2016.
- [34] J. Chen, B. Hu, P. Moore, X. Zhang, and X. Ma, "Electroencephalogram-based emotion assessment system using ontology and data mining techniques," *Applied Soft Computing*, vol. 30, pp. 663–674, 2015.
- [35] J. D. Hamilton, "Analysis of time series subject to changes in regime," *Journal of econometrics*, vol. 45, no. 1-2, pp. 39–70, 1990.



Fatma Patlar Akbulut received the B.S. and M.S. degree in computer engineering from Istanbul Kultur University, and Ph.D. degrees in biomedical engineering from Istanbul University in 2017. Her doctoral work focused on developing wearable systems for long-term cardiovascular disease monitoring. From 2017 to now, she is a Post-Doctoral Researcher at Computer Science Department of North Carolina State University. Her research interests include the signal processing, affective computing, machine learning, and wearable systems for healthcare.



Harry G. Perros is a Professor of Computer Science, an Alumni Distinguished Graduate Professor, and the co-founder and program coordinator of the Master of Science degree in Computer Networks at NC State University. He is an IEEE Fellow and he has published extensively in the area of performance modelling of computer and communication systems, and he has organized several national and international conferences. He has also published six print books and an e-book. His current research interests are in the

areas of resource allocation under QoS, IoT analytics, and queueing theory.



Muhammad Shahzad received his Ph.D. degree in computer science from Michigan State University in 2015. He is currently an Assistant Professor at the Department of Computer Science, North Carolina State University, USA. His research interests include design, analysis, measurement, and modeling of networking and security systems. He received the 2015 Outstanding Graduate Student Award, 2015 Fitch Beach Award, and 2012 Outstanding Student Leader Award at Michigan State University.