Expert Systems With Applications 112 (2018) 77–86

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A wavelet-based approach to emotion classification using EDA signals

Huanghao Feng, Hosein M. Golshan*, Mohammad H. Mahoor

Computer Vision Lab, ECE Dept., University of Denver, Denver, CO, USA

ARTICLE INFO

Article history:
Received 18 August 2017
Revised 15 May 2018
Accepted 7 June 2018
Available online 18 June 2018

Keywords:
Emotion classification
Feature extraction
Time-frequency analysis
Wearable device
Eletrodermal activity

ABSTRACT

Emotion is an intense mental experience often manifested by rapid heartbeat, breathing, sweating, and facial expressions. Emotion recognition from these physiological signals is a challenging problem with interesting applications such as developing wearable assistive devices and smart human-computer interfaces. This paper presents an automated method for emotion classification in children using electrodermal activity (EDA) signals. The time-frequency analysis of the acquired raw EDAs provides a feature space based on which different emotions can be recognized. To this end, the complex Morlet (C-Morlet) wavelet function is applied on the recorded EDA signals. The dataset used in this paper includes a set of multimodal recordings of social and communicative behavior as well as EDA recordings of 100 children younger than 30 months old. The dataset is annotated by two experts to extract the time sequence corresponding to three main emotions including "Joy", "Boredom", and "Acceptance". The annotation process is performed considering the synchronicity between the children's facial expressions and the EDA time sequences. Various experiments are conducted on the annotated EDA signals to classify emotions using a support vector machine (SVM) classifier. The quantitative results show that the emotion classification performance remarkably improves compared to other methods when the proposed wavelet-based features are used.

1. Introduction

Emotion understanding plays an important role in effective social communication among human beings. It can affect different aspects of human life such as learning, innovation, creativity, motivation, decision making, perception, and social interaction (Canento, Fred, Silva, Gamboa, & Lourenço, 2011). Hence, emotion recognition has remained one of the most important research topics in areas ranging from psychology (Kim, Bang, & Kim, 2004; Rigas, Katsis, Ganiatsas, & Fotiadis, 2007) to computer science (Blain, Mihailidis, & Chau, 2008). Researchers have been working on developing intelligent systems with the ability of recognizing emotions (e.g. facial expressions) for a variety of applications such as assistive devices and smart human-computer interfaces (Blain et al., 2008; Oberman, Winkielman, & Ramachandran, 2009). A possible application of automated emotion classification methods would be monitoring the emotional states of individuals suffering from socio-emotional disorders and eventually give them feedback about their emotional status.

In recent years, there have been efforts to develop wearable assistive devices that can recognize individuals' emotions from either facial expressions or biological signals (e.g. electrodermal activity) and give feedback to the user (Poh, Swenson, & Picard, 2010). The success of these assistive systems depends on how accurately machine learning algorithms can recognize emotions from biological signals. If successful, the systems could be a tremendous benefit to individuals, such as children with autism, who may use wearable devices to understand their own emotions and feelings, and eventually adjust their social behavior accordingly (Goodwin et al., 2016; Lord, Ruter, & Le Couteur, 1994; Prince et al., 2017). Hence, it is important to investigate and evaluate emotions/expressions caused by different social activities and stimuli and their effects on biological signals.

The human body responds to external emotion stimuli in various ways such as facial expressions, increased heart rate (HR), opening of the eye aperture, as well as sweating of the skin. Emotion recognition using biological signals has been studied in various multidisciplinary literature ranging from psychology to computer science (Atkinson & Campos, 2016; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000; Calvo & D'Mello, 2010; Gamboa, 2008; Haag, Goronzy, Schaich, & Williams, 2004; Perez-Gaspar, Caballero-Morales, & Trujillo-Romero, 2016; Whang & Lim, 2008). To gain access to these bio-signals, different electrical devices and sensors are needed, which are directly attached to specific

* Corresponding Author.
E-mail addresses: huanghao.feng@du.edu (H. Feng), hosein.golshanmojdehi@du.edu (H.M. Golshan), mohammad.mahoor@du.edu (M.H. Mahoor).

https://doi.org/10.1016/j.eswa.2018.06.014
0957-4174/

parts of the body. For instance, EDA, one of the most commonly used bio-signals to analyze body response and emotion features, can be measured by simple wearable devices such as a Q-sensor (Mardaga, Laloyaux, & Hansenne, 2006; Ooi et al., 2016). Even though there are some difficulties in using EDA-sensitive sensors, such as the skin status and skin contact (SC) level, they are easily accessible devices for emotion recognition purposes.

The objective of this paper is to develop an automatic emotion classification method using EDA signal. The proposed approach in this paper can be used in wearable assistive devices. Although, the dataset used in this research is collected from children, the proposed signal processing approach is not limited to any demographics or subjects' age range or disorders. However, the recognition accuracy is solely based on the data used in this research and that may or may not be generalizable to other datasets.

In this paper, we use a dataset provided by Georgia Institute of Technology (Presti, Sclaroff, & Rozga, 2013; Rajagopalan, Dhall, & Goecke, 2013; Rajagopalan, Murthy, Goecke, & Rozga, 2015; Rehg et al., 2013), which contains the EDA signals recorded form 100 subjects aged less than 30 months old collected during 3–5 minute social interactions with a caregiver (the details of the dataset are elaborated in Section 3). The collected socio-emotional data is then used for developing an automated supervised learning approach for recognition of emotions based on EDA signal. The Circumplex model of affect was adopted for current research (Posner, Russell, & Peterson, 2005; Russell, 1980), with happy being labeled as "Joy", bored as "Boredom" and "Acceptance" as a neutral statement. In our work, the dataset is first annotated to label perceived emotions (e.g., Acceptance, Joy, Boredom) expressed by each subject. Afterwards, we utilize the continuous wavelet transform to develop a new feature space for classification purposes. Using the complex Morlet function (Grossmann & Morlet, 1984), the wavelet coefficients of the EDA signal at different scales are calculated, providing a more detailed representation of the input signal. The performance of the proposed feature space on emotion classification task is evaluated using the canonical support vector machine (SVM) classifier with different types of kernel functions as well as the K-nearest neighborhood (KNN) classifier.

The rest of this paper is organized as follows: Section 2 presents some related works concerning emotion classification using different types of wearable devices and bio-signals. Section 3 elaborates on the dataset used in this research and the annotation procedure. The proposed classification scheme is presented in Section 4. Experimental results are given in Section 5. Finally, Section 6 concludes the paper with some remarks for future research.

2. Related work

Physiological responses have been identified as reliable indicators of human emotional and cognitive states. This section is dedicated to review some existing methods used for human emotion recognition based on various physiological responses, such as facial expression and other types of bio-signals.

A wearable glass device was designed by Kwon, Kim, Park, and Kim (2016) to measure both electrodermal activity (EDA) and photoplethysmogram data for emotion recognition purposes. A built-in camera was also used in this device for capturing partial facial expression from the eye and nose area. This approach obtains remarkable performance in facial expression recognition in the subject-dependent cases. However, for subject-independent cases, it results in different accuracies across different types of emotions, which is an undesirable feature.

EDA has been used as an effective and reproducible electrophysiological method for investigating sympathetic nervous system function (Kwon et al., 2016; Shahani, Halperin, Boulu, & Cohen, 1984; Stagg, Davis, & Heaton, 2013; Tarvainen, Karjalainen, Koistinen, & Valkonen-Korhonen, 2000). Note that the sympathetic nervous burst changes the skin conductance, which can be traced by analyzing the EDA signals (Kylliäinen & Hietanen, 2006; Lidberg & Wallin, 1981; Rehg et al., 2013). The Q-sensor is a convenient wireless-based EDA device with no need for cables, boxes, or skin preparation. This device can track three types of data including EDA, temperature, and acceleration at the same time (Kappas, Küster, Basedow, & Dente, 2013). It is worth mentioning that as of today, there has been no published work on emotion classification using the EDA signals collected by this dataset collected at the Georgia Institute of Technology (Rehg et al., 2013).

Several emotion classification methods have been presented in the literature using different bio-signals (Legiša, Messinger, Kermol, & Marlier, 2013; Liu, Liu, & Lai, 2014; Schmidt & Walach, 2000; Xu, Fu, Jiang, Li, & Sigal, 2016). Due to the variety of the signals used in these methods, different approaches have been designed to comply with their specific characteristics. Analysis of variance (ANOVA) and linear regression (Schmidt & Walach, 2000) are the commonly used methods to extract features from bio-signals and to recognize different emotional states. These methods are based on the assumption of a linear relationship between the recorded signals and emotional states. A fuzzy-based classification method (Liu et al., 2014) has been used in to transform EDA and facial electromyography (EMG) to valence and arousal states. These states were then used to classify different emotions.

Artificial neural networks (ANN) have also been applied for emotion classification tasks based on physiological responses. Lin, Wang, Wu, Jeng, and Chen (2007) developed a multilayer perceptron network (MLP) architecture capable of recognizing five emotions using various features from Electrocardiography (ECG) and EDA signals, and obtained very accurate classification performance. Nasoz, Alvarez, Lisetti, and Finkelstein (2004) employed K-nearest neighborhood and discriminant function analysis to perform the emotion classification task using different features extracted from the EDA signals, body temperature and heart rate.

Support Vector Machine (SVM) is a well-known supervised learning algorithm that has extensively been used for pattern classification and regression (Cortes & Vapnik, 1995). The SVM classifier tends to separate dataset by drawing an optimal hyperplane between classes such that the margin between them becomes maximum. The samples of each class that are located within the margin are called support vectors and play the main role in calculating the parameters of the hyperplanes between the corresponding classes. Machine learning algorithms such as SVM, linear discriminant analysis (LDA), and classification and regression tree (CART) have been employed for emotion classification purposes. For instance, in several works including Jang et al. (2014) and Sano et al. (2011), the authors combined various types of bio-signals such as ECG, skin temperature (SKT), HR, and Photoplethysmogram (PPG) for emotion classification purposes. Amershi, Conati, and McLaren (2006) proposed unsupervised clustering methods for emotion recognition. Their method benefited from several features obtained from different body responses such as SC, HR, and EMG. They showed that only a few statistical features such as the mean and standard deviation of the data can be relevant identifiers for defining different clusters.

EDA signals are nonstationary and noisy; hence, wavelet-based analysis of EDA signals has been considered in the literature (Laparra-Hernández, Belda-Lois, Medina, Campos, & Poveda, 2009; Swangnetr et al., 2013) either as a pre-processing step or a feature extraction approach for emotion classification. Swangnetr et al. (2013) used a set of wavelet coefficients representing EDA features together with heart rate signal to increase the percentage of correct classifications of emotional states and provide clearer relationships among the physiological response and arousal and valence. Sharma, Prakash, and Kalra (2016) used a

feature space based on the discrete wavelet transform (DWT) of the EDA signal to distinguish subjects suffering social anxiety disorder (SAD) and a control group. Using MLP and DWT features, they achieved a classification accuracy of ∼85%.

To the best of our knowledge there are a few works (Legiša et al., 2013; Oberman et al., 2009) that have studied and compared different automated classification techniques for emotion recognition of children using EDA signals. This motivated us to conduct this study using an existing dataset, which concentrates on emotion classification of children based on the relationship between their facial expressions and the collected EDA signals.

3. Data annotation and specification

The dataset used in this paper constitutes a set of multimodal recordings of social and communicative behavior of 100 children younger than 30 months old provided by the Georgia Institute of Technology (Rehg et al., 2013). All data was collected in the Child Study Lab (CSL) at Georgia Tech, under a university-approved IRB protocol. The laboratory was 300-square feet space, and the temperature/humidity of the room for all sessions was kept the same.

Note that, in general, the operating temperature of the Q-sensor used for data acquisition sessions is in the range of 5 °C to 40 °C (40°F to 104°F), and the humidity is in the range of 5% to 95% relative humidity non-condensing (45–65% is recommended). For experimental purposes in the laboratory, the ambient temperature and humidity were kept constant to guarantee the proper functionality of the Q-sensors as well as the comfort of subjects in light clothing. To obtain appropriate results and ensure the best possible function of the device, the sensor was cleaned before each usage. Q-sensors are equipped with Bluetooth which allows the user to monitor the acquired signals in real time; if there is any abnormal behavior with the receiving signals, the position of the Q-sensor is adjusted, or the experiment is repeated.

According to the dataset description, each session lasted 3–5 minutes during which the EDA signals (the sampling rate is 32 Hz) were collected using two Q-sensors attached to left and right wrists, and the entire experiment was video-recorded. A set of semi-structured play interactions with adults, known as Multimodal Dyadic Behavior (MMDB), was designed for the experimental sessions to stimulate different emotions; event 1: “greeting”, event 2: “playing with a ball”, event 3: “looking at a book and turning its pages”, event 4: “using the book as a hat”, and event 5: “tickling” . These experiments are aimed at analyzing and decoding the children’s social communicative behavior at early ages and are consistent with the Rapid-ABC play protocol (Ousley, Arriaga, Abowd, & Morrier, 2012).

Note that key socio-communicative milestones at early ages, especially younger than two years old, can be reflected from the behavior in the aforementioned five events. In addition, their diminished occurrences and qualitative differences in expression represent early clues of autism (Wetherby et al., 2004).

The dataset provided by the publishers has a specification sheet containing the annotation information of some subjects. Note that, for some of the subjects (36 subjects), the available EDA files were incomplete, so we excluded those subjects from our annotation process. To take advantage of the maximum capacity of the data in our experiments, two experts were employed to annotate the data based on the facial expression of subjects using their corresponding video files. As a result, from 100 subjects (one recording session per subject) participated in the experiments, 64 subjects (including fully and partially engaged subjects, see Table 1) were annotated.

In terms of the annotation process, both annotators needed to go through all the videos and record the frames with target emotions. Afterwards, the results were compared to ensure high interrater reliability between the two annotators.

Table 1
The number of participants in different engagement level among all events.

	Engaged	Partially-engaged	Disengaged
Event 1	53	9	2
Event 2	57	6	1
Event 3	53	5	6
Event 5	55	7	2

The annotation was carried out based on the synchronicity between the video frames and the recorded EDA sequences of each subject. In other words, the annotators went through the entire video file of each event frame by frame, and selected the frames related to the initiation and end of an emotion. At the same time, the corresponding sequences of the EDA signals were stored to generate the dataset for each perceived emotion. During the annotation process, two dominant emotions were recognizable; events 2 (with average duration of 45 seconds) and 5 (average duration of 35 seconds) stimulate the “Joy” emotion and event 3 (with an average duration of 60 seconds) stimulates “boredom” . With respect to event 1, “greeting”, it was difficult to assign a specific emotion to it; however, the annotators most often used “Acceptance” for this event. In addition, we excluded event 4 from our experiments since the length of this event (on average 9 seconds) was very short in comparison with other events (on average 50 seconds for other events), and the annotators were not able to identify any specific emotion triggered by this event. Moreover, we did not find meaningful information inside the EDA files associated with this event. Fig. 1 shows the above-described procedure graphically. Besides, the distribution of different emotions across all subjects and events is given in Fig. 2.

4. Proposed classification method

Since we developed our emotion classification method based on the time-frequency analysis of the EDA signals, the main properties of the continuous wavelet transform assuming complex Morlet wavelet is first presented here. Then, the pre-processing steps, as well as the wavelet-based feature extraction scheme, are discussed. Finally, we briefly review the characteristics of the support vector machine as the classifier used with our approach.

4.1. Continuous wavelet transform

The EDA data recorded using the SC sensors are categorized as non-stationery signals (Najafi et al., 2003; Swangnetr et al., 2013). Hence, multiresolution analysis techniques are essentially suitable to study the qualitative components of these kinds of bio-signals (Najafi et al., 2003). Note that continuous wavelet transform (CWT) is one of the strongest and most widely used analytical tools for multiresolution analysis. CWT has received considerable attention in processing signals with non-stationary spectra (Mallat, 1989; Vetterli & Herley, 1992); therefore, it is utilized here to perform the time-frequency analysis of the EDA signals. In contrast to many existing methods that utilize the wavelet coefficients of the raw signal to extract features, our proposed method is essentially based on the spectrogram of the original data in a specific range of frequency (0.5, 50)Hz, which provides more information for other post-processing (i.e., feature extraction and classification) steps. We apply the wavelet transform at various scales corresponding to the aforementioned frequency range to calculate the spectrogram of the raw signal (i.e., Short Time Fourier Transform (STFT) can also be used to calculate the spectrogram of the

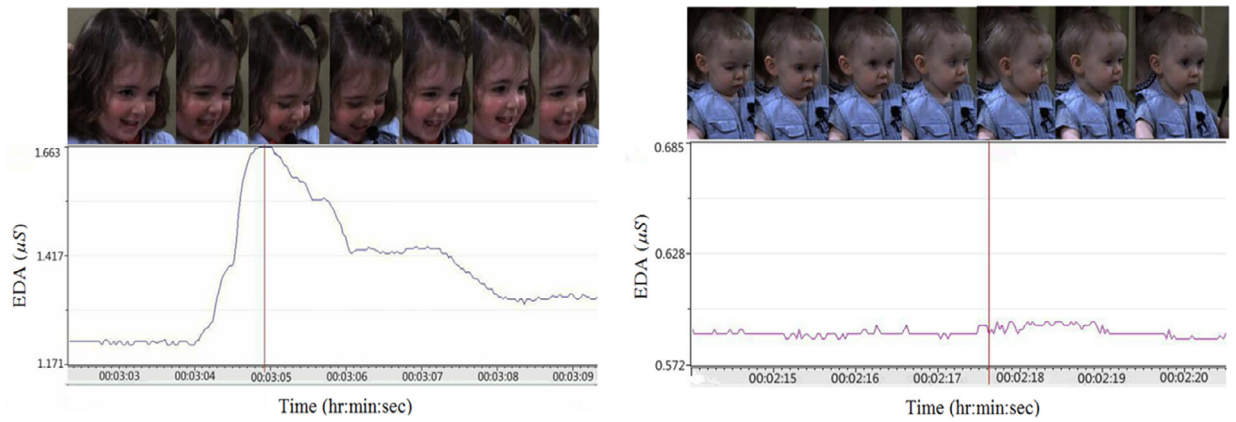

Fig. 1. Two samples of the annotation process. Left shows some video frames associated with event 5 "tickling" and right shows the video frames of event4 "using the book as a hat". The corresponding EDA signals are shown under each case. While for the event 5 the EDA signal contains meaningful information, the EDA signal of event 4 does not contain useful information, likely due to the disengagement of the subject.

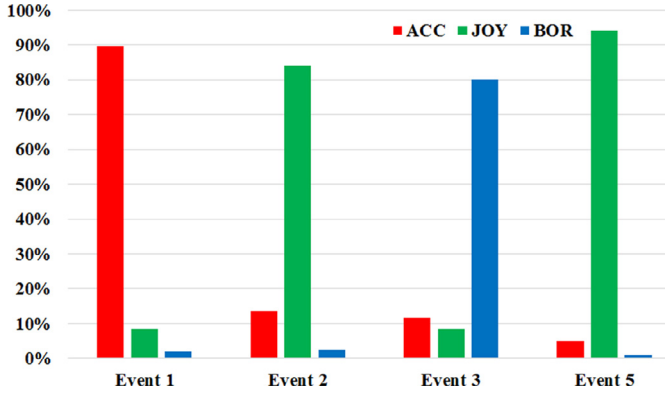

Fig. 2. The distribution of the emotions across all subjects and events. The abbreviations "ACC", "BOR", and "JOY" respectively correspond to the emotions "Acceptance", "Boredom", and "Joy".

raw signal). In addition, as opposed to many related studies that utilize real-valued wavelet functions for feature extraction purposes, we have employed the complex Morlet (C-Morlet) function with the proposed approach, as it takes into account both the real and imaginary components of the raw signal, leading to a more detailed feature extraction.

The wavelet transform of a 1-D signal provides a decomposition of the time-domain sequence at different scales, which are inversely related to their frequency contents (Godfrey, Conway, Leonard, Meagher, & Ólaighin, 2009; Mallat, 1989). This requires the time-domain signal under investigation to be convolved with a time-domain function known as "mother wavelet". The CWT applies the wavelet function at different scales with continuous time-shift of the mother wavelet over the input signal. As a consequence, it helps represent the EDA signals at different levels of resolution. For instance, it results in large coefficients in the transform domain when the wavelet function matches the input signal, providing a multiscale representation of the EDA signal.

Using a finite energy function $\psi(t)$ concentrated in the time domain, the CWT of a signal $x(t)$ is given by $X(\alpha,b)$ as follows (Vetterli & Herley, 1992):

$$X(a,b) = \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{a}}\psi^*\left(\frac{t-b}{a}\right)dt. \quad (1)$$

where, α, is the scale factor and represents dilation or contraction of the wavelet function and b is the translation parameter that slides this function on the time-domain sequence under analysis. Therefore, $\psi(\alpha,b)$ is the scaled and translated version of the corresponding mother wavelet. "*" is the conjugation operator.

Note that the wavelet coefficients obtained from Eq. (1) essentially evaluate the correlation between the signal $x(t)$ and the wavelet function used at different translations and scales. This implies that the wavelet coefficients calculated over a range of scales and translations can be combined to reconstruct the original signal as follows:

$$x(t) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} X(a,b)\psi\left(\frac{t-b}{a}\right)dadb. \quad (2)$$

4.2. Wavelet-based feature extraction

The time-frequency analysis of various bio-signals has been addressed in many related literature (Golshan, Hebb, Hanrahan, Nedrud, & Mahoor, 2016; Golshan, Hebb, Hanrahan, Nedrud, & Mahoor, 2017; Golshan, Hebb, Hanrahan, Nedrud, & Mahoor, 2018; Li, Zhang, Tao, Sun, & Zhao, 2009). It has been shown that the wavelet-domain feature space can improve the recognition performance of different human activities using the signals emanated from the body responses. Therefore, it essentially enhances the classification performance due to the more distinctive feature space provided.

In this paper, we focus on the time-frequency analysis of the EDA signal to provide a new feature space based on which emotion classification task can be done. As opposed to some related studies that employ the raw time-domain signals for classification purposes (Greco, Valenza, Citi, & Scilingo, 2017; Jang, Park, Kim, & Sohn, 2012), we use the amplitude of the CWT of the EDA signals to generate the features and drive the classifier. Working in the wavelet-domain is essentially advantageous since the wavelet transform probes the given signal at different scales, extracting more information for other post-processing steps. In addition, the localized support of the wavelet functions enables CWT-based analysis to match to the local variations of the input time sequence (Vetterli & Herley, 1992). As a result, a more detailed representation of the signal is provided in comparison with the raw time-domain signal.

Fig. 3 shows the amplitude of the CWT of a sample EDA signal at different scales using complex Morlet (C-Morlet) wavelet function. As can be seen, due to the localization property of the CWT, different structures of the input signal are extracted at each level of decomposition, providing useful information for analyzing the recorded EDA signals.

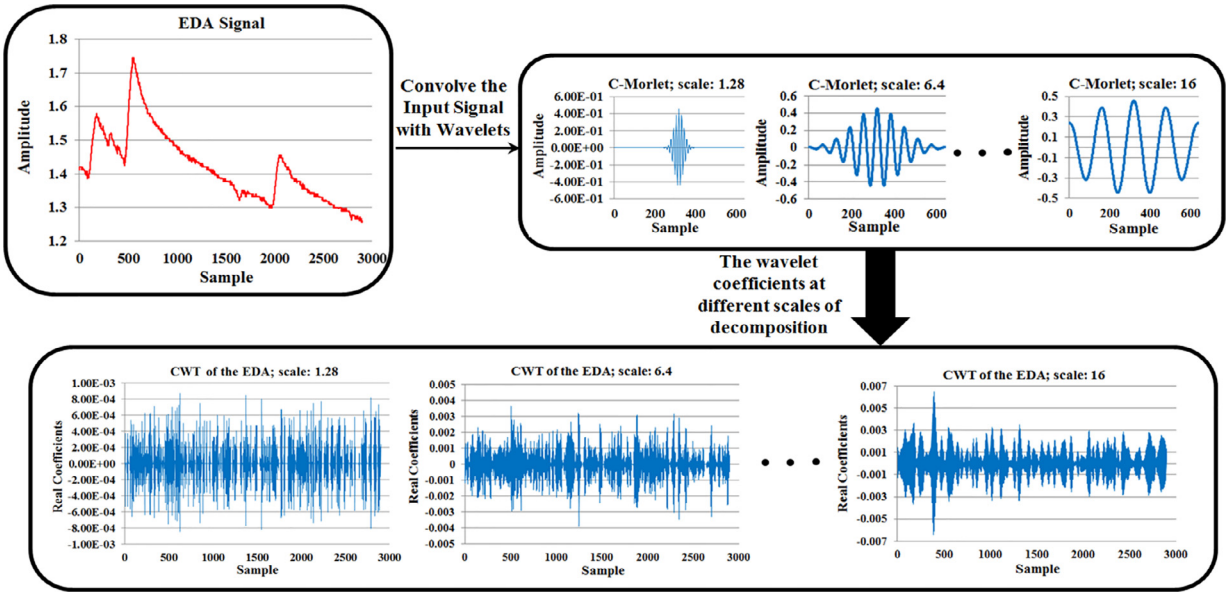

Fig. 3. The CWT of a typical EDA signal using the C-Morlet mother wavelet. Different scales of the wavelet functions are convolved with the original EDA signal to highlight different features of the raw data. As can be seen inside the bottom box, when the scaling parameter of the wavelet function increases, the larger features of the input signal are augmented. On the other hand, the detailed structures of the signal are better extracted when the scaling factor decreases.

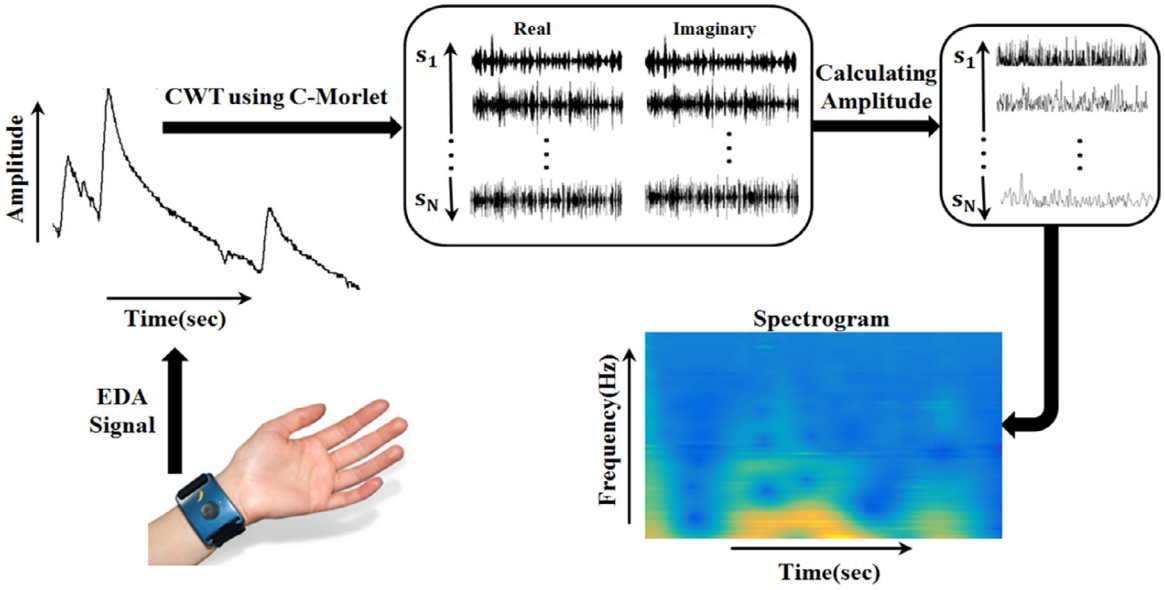

Fig. 4. The wavelet-based feature extraction. Using the C-Morlet mother wavelet, the real and imaginary wavelet coefficients are calculated at different scales. Then the amplitude of these coefficients is calculated to provide the corresponding spectrogram. This spectrogram is then used as the feature space.

In this work, we have employed the C-Morlet wavelet function to process the acquired EDA signals, as it has been well used for time-frequency analysis of different bio-signals and classification (Golshan et al., 2016). Note that the impact of different families of the wavelet functions (e.g., Symlets, Daubechies, Coiflets) on the emotion classification will be evaluated in Section 4. The equation of the C-Morlet mother wavelet with f_c as its central frequency and f_b as the bandwidth parameter is given as follows:

$$\psi(t) = \frac{\exp\left(-t^2/f_b\right)}{\sqrt{\pi f_b}} \exp\left(j2\pi f_c t\right). \tag{3}$$

Fig. 4 shows the above-described wavelet-based feature extraction approach graphically.

4.3. Support vector machine

The SVM classifier tends to separate data $D = \{\mathbf{x}_i, y_i\}_{i=1}^{N}, \mathbf{x}_i \in R^d, y_i \in \{-1, +1\}$ by drawing an optimal hyperplane $\mathbf{w}, \mathbf{x} + b = 0$ between classes such that the margin between them becomes maximum (Cortes & Vapnik, 1995). With reference to Fig. 5, H_1 and H_2 are the supporting planes and the optimal hyperplane (OH) splits this margin such that it stands at the same distance from each supporting hyperplane. This implies that the margin between H_1 and H_2 is equal to $2 / \|\mathbf{w}\|$.

In terms of linearly separable classes, the classifier is obtained by maximizing the margin $2/\|\mathbf{w}\|$, which is equivalent to minimizing $\|\mathbf{w}\|/2$ with a constraint in convex quadratic programming (QP) as follows:

$$\min \frac{1}{2}\|\mathbf{w}\|^2 s.t. y_i(\langle \mathbf{w}, \mathbf{x}_i \rangle + b) \geq 1. \tag{4}$$

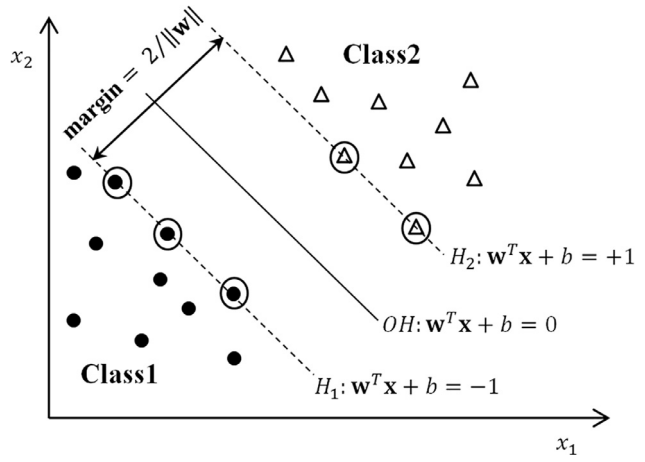

Fig. 5. Canonical SVM for classifying two linearly separable classes. The decision boundary is shown by OH. Two hyperplanes H_1 and H_2 pass the support vectors that are circled inside the figure.

where, **w** and b are the parameters of the hyperplane and ⟨.,.⟩ is the notation of the inner product.

However, different classes are seldom separable by a hyperplane since their samples are overlapped in the feature space. In such cases, a slack variable $\xi_i \geq 0$ and a penalty parameter $C \geq 0$ are used with the optimization step to obtain the best feasible decision boundary. It is given as:

$$\min \frac{1}{2}\|\mathbf{w}\|^2 + C\left(\sum_{i=1}^{N}\xi_i\right) s.t. y_i(\langle \mathbf{w}, \mathbf{x}_i\rangle + b) \geq 1 - \xi_i. \quad (5)$$

Usually, various kernel functions are used to deal with the nonlinearly separable data. As a result, the original data $\mathbf{x}_i$ is mapped onto another feature space through a projection function $\varphi(\bullet)$. It is not necessary to exactly know the equation of the projection $\varphi(\bullet)$, but one can use a kernel function $k(\mathbf{x}_i,\mathbf{x}_j) = \langle\varphi(\mathbf{x}_i),\varphi(\mathbf{x}_j)\rangle$. This function is symmetric and satisfies the Mercer's conditions. The Mercer's conditions determine if a candidate kernel is actually an inner-product kernel. Let $k(\mathbf{x}_i,\mathbf{x}_j)$ be a continuous symmetric kernel defined in the closed interval $t_1 \leq t \leq t_2$, the kernel can be expanded into series $\sum_{n=1}^{\infty}\lambda_n\varphi_n(\mathbf{x}_i)\varphi_n(\mathbf{x}_j)$, where $\lambda_n > 0$ are called eigenvalues and functions φ_n are called eigenvectors in the expansion. The fact that all the eigenvalues are nonnegative means that the kernel is positive semidefinite (Cortes & Vapnik, 1995). To maximize the margin, H_1 and H_2 are pushed apart until they reach the support vectors on which the solution depends. To solve this optimization problem, the Lagrangian dual of Eq. (5) is used as follows:

$$\begin{aligned} &\max_{\alpha} \sum_{i=1}^{N}\alpha_i - \tfrac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} y_i y_j \alpha_i \alpha_j k\left(\mathbf{x}_i, \mathbf{x}_j\right). \\ &s.t.\ 0 \leq \alpha_i \leq C,\ \sum_{i=1}^{N}\alpha_i y_i = 0,\ i = 1,\ldots,N \end{aligned} \quad (6)$$

where, $\{\alpha_i\}$s are the Lagrangian multipliers in which just a few number of them are non-zero. These non-zero values are corresponding to the support vectors determining the parameters of the hyperplane $\mathbf{w} = \sum_{i=1}^{N}\alpha_i y_i \mathbf{x}_i$. Therefore, the label of the test sample (y_z) is given by:

$$y_z = \mathrm{sgn}\left(\sum_{i=1}^{N}\alpha_i y_i k(\mathbf{x}_i, z) + b\right). \quad (7)$$

Table 2
Comparison of different wavelet functions on the feature extraction and emotion classification performance (%) of 2 and 3 classes using SVM classifier with different kernels. The abbreviations "Acc", "Bor", and "Joy" respectively stand for the emotions "Acceptance", "Boredom", and "Joy". The best value is highlighted in each case.

	Kernels	db1	coif1	sym2	C-Morlet
Acc-Bor	Linear	61	56	61	**75**
Acc-Joy		50	46	50	**69**
Bor-Joy		51	69	57	**90**
Bor-Joy-Acc		51	35	39	**66**
Acc-Bor	Polynomial	51	56	58	**64**
Acc-Joy		54	54	57	**81**
Bor-Joy		55	64	69	**86**
Bor-Joy-Acc		43	46	50	**61**
Acc-Bor	RBF	59	60	58	**74**
Acc-Joy		55	44	57	**84**
Bor-Joy		68	51	69	**89**
Bor-Joy-Acc		45	35	50	**69**

5. Experiments and results

To evaluate the accuracy of the proposed feature extraction method on the emotion classification performance, we employ the EDA signals of 64 subjects annotated based on their facial expression, as elaborated in Section 3. The EDA dataset is classified based on different emotions perceived in the annotation step, which includes Joy, Boredom, and Acceptance emotions. The SVM classifier is applied on the dataset using three different kernel functions including the Linear function $k(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T\mathbf{y} + c$, Polynomial function $k(\mathbf{x}, \mathbf{y}) = (\mathbf{x}^T\mathbf{y} + c)^d$, and Radial Basis Function (RBF) $k(\mathbf{x}, \mathbf{y}) = exp(\gamma\|\mathbf{x}-\mathbf{y}\|^2)$, where **x** and **y** are two feature vectors, and γ, c, and d are constant values.

Before we proceed with the quantitative performance of the proposed emotion classification approach, it is worthwhile to first examine the impact of different families of wavelet functions on the feature extraction step as well as emotion classification ability.

5.1. Determination of mother wavelet

Since the proposed feature extraction method used in this paper is based on the time-frequency representation of the input EDA signal, the selection of a suitable mother wavelet that can better differentiate various emotions in the feature space is of great importance. To this end, we have applied different types of the wavelet functions (Daubechies, Coiflets, Symlets, C-Morlet) (Godfrey et al., 2009) to evaluate their effect on the classification performance of the EDA dataset.

Table 2 shows the classification results given by different wavelet functions. For the sake of brevity, only the results of the "db1", "coif1", "sym2", and "C-Morlet" wavelets and all three kernels with the SVM classifier are shown. As can be seen, the time-frequency features calculated by the C-Morlet leads to a higher classification performance, likely due to the more distinctive feature space provided by this wavelet function. Fig. 6 shows the difference between the aforementioned wavelet functions. Note that the C-Morlet wavelet has successfully been applied on different types of bio-signals (e.g., EEG, LFP brain signals) and led to promising results (Golshan et al., 2016; Golshan et al., 2017), specifically for feature extraction purposes. One of the main characteristics of this wavelet function is its complex nature that essentially tends to extract more features from the input time sequence.

5.2. Classification results

In this section, the pre-processing steps performed on the raw EDA dataset are first explained. Afterwards, the classification results with different modalities using SVM and KNN classifiers are

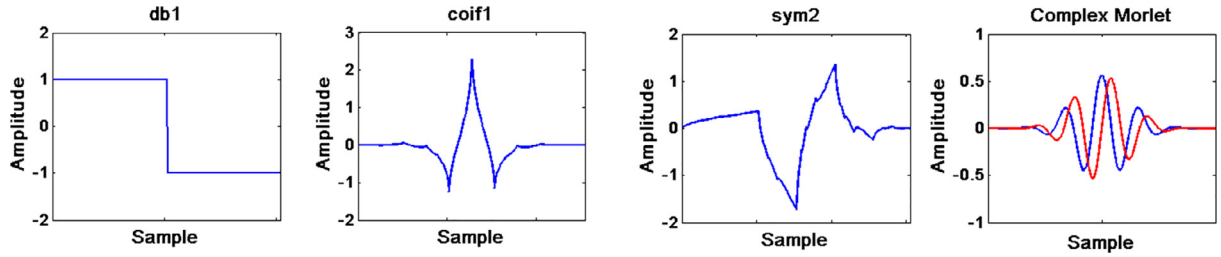

Fig. 6. Different mother wavelets used for feature extraction in this paper. The "db1", "coif1" and "sym2" wavelets are real-valued functions, while the "C-Morlet" function is complex-valued. The corresponding imaginary part of this wavelet function is highlighted in red inside the figure.

Table 3
Comparison of classification accuracy (%) of SVM classifier with different kernel functions using the presented wavelet-based feature extraction, the raw EDA data, and the raw EDA data + statistical features. The results of 64 subjects and 2 and 3-class classification cases are reported. The abbreviations "ACC", "JOY", and "BOR" respectively stand for the emotions "Acceptance", "Joy", and "Boredom". The best value is highlighted in each case. .

		Wavelet-based				Statistics-based feature + Raw data				Raw data			
	Kernels	Accuracy	AUC	Precision	Recall	Accuracy	AUC	Precision	Recall	Accuracy	AUC	Precision	Recall
ACC-BOR	Linear	**75**	**75**	**81**	**71**	56	55	57	56	56	55	58	43
ACC-JOY		**69**	**84**	**78**	**79**	47	56	47	47	49	57	49	39
BOR-JOY		**90**	**87**	**82**	**88**	55	55	54	55	53	54	52	67
ACC-JOY-BOR		**66**				35				36			
ACC-BOR	Polynomial	64	68	70	64	**74**	**82**	**83**	**74**	70	78	79	70
ACC-JOY		**81**	83	78	**81**	77	**90**	**84**	77	60	68	72	60
BOR-JOY		**86**	**87**	**85**	**86**	60	66	59	60	57	62	57	57
ACC-JOY-BOR		**61**				55				46			
ACC-BOR	RBF	**74**	**79**	**84**	**74**	53	58	54	53	57	59	60	57
ACC-JOY		**84**	**89**	**85**	**84**	42	61	37	42	42	65	30	42
BOR-JOY		**89**	**90**	**85**	**89**	51	54	51	51	50	55	50	50
ACC-JOY-BOR		**69**				34				34			

presented. We compare the classification performance of the proposed wavelet-based feature extraction method with the raw EDA signal. Also, the performance of a statistical feature extraction method used for EDA signal (Liu, Fan, Zhang, & Gong, 2016; Mera & Ichimura, 2004) is compared with the proposed method**.** Note that the extracted features with this method are mainly based on the statistical moments of the acquired EDA time sequence such as "the means of the raw signals", "the standard deviations of the raw signals", "the means of the absolute values of the first differences of the raw signals", "the means of the absolute values of the first differences of the normalized signals", "the means of the absolute values of the second differences of the raw signals", and "the means of the absolute values of the second differences of the normalized signals".

The segments of the EDA signals obtained from the annotation step are passed through a median filter of size 10 to smooth the signal, eliminating some existing impulsive noise that probably happens due to the quick movements of the subjects during the experiments. Then, the amplitude of the wavelet coefficients are calculated for the frequency range of (0.5, 50)Hz. Such a wide frequency range was empirically chosen to guarantee that all detailed components of the EDA signal are taken into account (see Fig. 3).

Principal component analysis (PCA) (Abdi & Williams, 2010) is then applied on the extracted wavelet-based features to decrease the dimensionality of the data, and, therefore, reduce the computational burden. PCA is a well-known dimensionality reduction approach which is extensively used for data analysis before classification. As such, it can reduce the possibility of overfitting, which may occur due to the large size of the feature vectors. Note that, in our experiments, 95% of the eigen-values corresponding to the maximum variance directions are kept. Since the spectrogram of the raw EDA data (see Fig. 4) is calculated for 100 scales (e.g., Frequency range (0.5, 50)Hz with a resolution of 0.5 Hz), for a fair comparison, we first down-sample the spectrogram by a factor of 100 to make the length of the wavelet-based features equal to the length of the raw data. Then, PCA is applied on it. As a result, on average, the length of the wavelet-based feature vector before and after PCA is respectively ∼1000 and ∼35 samples, while these lengths are ∼1000 and ∼ 15 samples for the raw data.

A leave-one-out cross validation (LOOCV) approach is used to generate the training and test sets for the classification step. In terms of the SVM classifier (*LibSVM* library (Chang & Lin, 2011)), the parameters of the hyperplane are set to achieve the best classification accuracy on the validation set. The following parameters are used for each kernel function: 1. Linear kernel $C = 0.01$, 2. RBF kernel $C = 0.01$, $\gamma = 0.001$, and 3. Polynomial kernel $C = 0.01$, $d = 2$. These values are experimentally set so as to obtain the best classification performance.

Table 3 shows the classification accuracy for the SVM classifier with different kernel functions and the dataset acquired from 64 annotated subjects. In terms of the binary classification cases (i.e., "Acceptance vs Boredom", "Acceptance vs Joy", and "Boredom vs Joy"), besides the classification accuracy, the quantitative measures precision (true positive / (true positive + false positive)), recall (true positive / (true positive + false negative)), and AUC (area under the receiver operating characteristic curve),which are also given in this table. To calculate the precision and recall, first, we take one of the emotions as the positive class and calculate the precision and recall values. Then, we change the order and use the other emotion as the positive class and calculate the precision and recall values. Finally, the average value of the calculated precision and recall is given in the table. As can be seen, the proposed wavelet-based features lead to a higher classification performance as compared to other methods in almost all cases. For example, while the classification accuracy of three-class classification case with the linear kernel SVM and raw EDA signal is about 38%, the introduced feature space achieves an accuracy of 69%. Furthermore, while the proposed wavelet-based features lead to a stable classification performance for all kernel functions, the raw EDA data, and the combination of the raw data and the statistical features show

Table 4
Comparison of classification accuracy (%) of KNN classifier with different K values using the presented wavelet-based feature extraction, the raw EDA data, and the raw EDA data + statistical features. The results of 64 subjects and 2 and 3-class classification cases are reported. The abbreviations "ACC", "JOY", and "BOR" respectively stand for the emotions "Acceptance", "Joy", and "Boredom". The best value is highlighted in each case.

		Wavelet-based				Statistics-based feature + Raw data				Raw data			
		Accuracy	AUC	Precision	Recall	Accuracy	AUC	Precision	Recall	Accuracy	AUC	Precision	Recall
ACC-BOR	K = 1	70	50	70	70	**73**	**53**	**72**	**73**	68	52	68	68
ACC-JOY		**76**	58	76	**76**	76	**63**	**77**	76	62	60	62	62
BOR-JOY		**80**	64	**76**	**80**	57	66	57	57	56	**68**	56	56
ACC-JOY-BOR		**60**				56				43			
ACC-BOR	K = 3	70	63	71	70	**81**	**77**	**79**	**81**	71	68	75	71
ACC-JOY		**82**	81	80	**82**	82	**82**	**81**	82	61	50	61	61
BOR-JOY		**85**	**84**	**78**	**85**	60	53	61	60	55	56	55	55
ACC-JOY-BOR		**65**				56				46			
ACC-BOR	K = 5	64	64	67	64	**75**	**78**	74	**75**	71	67	**75**	71
ACC-JOY		**77**	**84**	73	**77**	77	82	**77**	77	57	50	58	57
BOR-JOY		**85**	**86**	**77**	**85**	61	50	63	61	46	57	46	46
ACC-JOY-BOR		**66**				58				43			

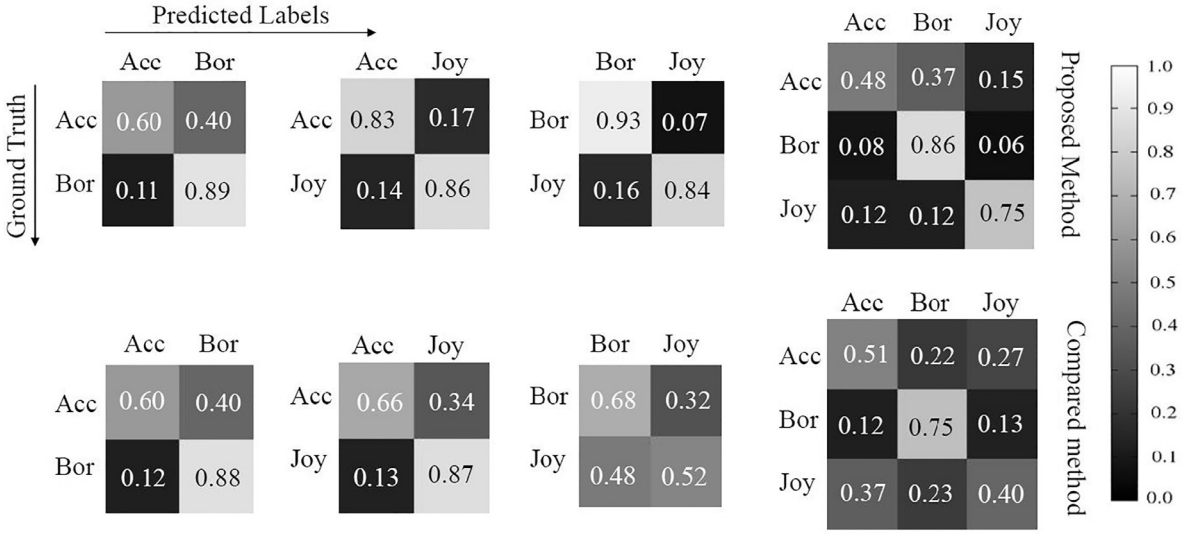

Fig. 7. The confusion matrix of SVM classifier using the proposed wavelet-based feature extraction method (top row) and the results obtained by combining the raw EDA signal and statistical features (bottom row). For the sake of brevity, the best result for each case is given here (e.g., the results of the RBF kernel for the proposed approach, and the results of the polynomial kernel for the other method). The abbreviations "Acc", "Joy", and "Bor" respectively stand for the emotions "Acceptance", "Joy", and "Boredom".

competitive classification performance only for the polynomial kernel. Note that one major problem when analyzing physiological signals is noise interference. In particular, the EDA signal is non-stationary and may include random artifacts, which makes it unsuitable to use the raw time sequence for practical signal processing approaches. Prior studies have represented stochastic physiological signals using statistical features to classify emotional states (Mera & Ichimura, 2004). Unfortunately, information can be lost with such features as simplifying assumptions are made, including knowledge of the probability density function of the data. Furthermore, there may be signal features that have the potential to improve emotion classification accuracy, but are not yet identified (Swangnetr et al., 2013).

The results of the KNN classifier with three different values for $K = 1$, 3, 5 are also given in Table 4. As shown, the proposed method outperforms in most of the cases. The results obtained by the combination of the raw data and the statistical features surpass the proposed method for some classification tasks. For instance, it obtains a classification accuracy of 73% for $K = 1$ and "Acceptance vs Boredom" classification task, while the proposed method reaches 70% for the same task. However, in other cases, the proposed shows better performance. An interesting comparison can be done for 3-class classification, where the proposed method achieves ~64% classification accuracy on average for all K values, and two other feature extraction approaches result in ~57% and ~44% respectively, showing the superiority of the proposed method for the complicated classification tasks.

Fig. 7 shows the confusion matrices of the SVM classifier with the proposed method and combination of the raw EDA data and statistical features. For the sake of brevity, the best result for each method is shown in this figure. As one can see, the classification accuracy of each single emotion increases with the proposed approach in most of the cases. For instance, the recognition rates of emotions "Acceptance", "Joy" and "Boredom" are respectively 51%, 40% and 75% using the raw EDA signal together with the statistical features. On the other hand, the corresponding recognition rates of the same emotions are 48%, 86% and 75% with the proposed feature extraction scheme.

6. Discussion, Conclusions, and future work

This paper presented an emotion classification approach based on the EDA signals acquired from Q-sensors. The EDA dataset used in this paper consists of a set of multimodal recordings of social and communicative behavior of 100 children younger than 30 months old provided by the Georgia Institute of Technology. We

considered the time-frequency analysis of the input signals to generate a more reliable feature space for emotion recognition. To this end, the continuous wavelet transform (CWT) of the data assuming complex Morlet (C-Morlet) wavelet function was used inside a frequency range of (0.5, 50)Hz. A SVM classifier was then employed to classify different emotions using the wavelet-based features.

To set up the experiments, the dataset was first annotated in order to segment the time sequence of the EDA signal in accordance with the facial expression of the children. Three different emotions were recognized in the annotation step including Acceptance, Joy, and Boredom. Various experiments were carried out on the dataset using either the raw segmented EDA signal or its corresponding time-frequency representation as features. The quantitative results show that the emotion classification performance is remarkably improved when the proposed wavelet-based features are used with the SVM classifier.

Apart from the "C-Morlet", we have also evaluated the effect of different wavelet functions such as "Symlets", "Daubechies", and "Coiflets" on the feature extraction stage, and therefore, the classification performance. The experimental results confirmed the superiority of the "C-Morlet" wavelet function.

Developing an automated system capable of real-time processing of the data can be an interesting extension to this work. This enables us to detect emotions and give feedback to the participants during the experimental sessions. Moreover, due to the limitation of available datasets, creating a more comprehensive dataset would be necessary for the future research, so that we can compare the difference between non-autistic and autistic groups regarding the emotion reaction during different social stimuli.

As another future work, it would be relevant to involve various emotions in the experiments such as anger, sadness and surprise. It would be interesting to expand current knowledge on the numerous emotion patterns as outcome of EDA signals from children with autism.

Acknowledgement

The authors would like to express their sincere thanks to the Georgia Institute of Technology, Atlanta, Georgia, USA, for providing us with the dataset and the response files of the subjects used in this research. This research is partially supported by grant IIS-1450933 from the National Science Foundation.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis.. *Wiley Interdisciplinary Reviews: Computational Statistics, 2*(4), 433–459.

Amershi, S., Conati, C., & McLaren, H. (2006). Using feature selection and unsupervised clustering to identify affective expressions in educational games. *8th international conference on intelligent tutoring systems workshop in motivational and affective issues in ITS*.

Atkinson, J., & Campos, D. (2016). Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert Systems with Applications, 47*, 35–41.

Blain, S., Mihailidis, A., & Chau, T. (2008). Assessing the potential of electrodermal activity as an alternative access pathway. *Medical Engineering & Physics, 30*(4), 498–505.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. *Handbook of Emotions, 2*, 173–191 Edition.

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1*(1), 18–37.

Canento, F., Fred, A., Silva, H., Gamboa, H., & Lourenço, A. (2011). Multimodal biosignal sensor data handling for emotion recognition. In *IEEE conf on sensors* (pp. 647–650).

Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST), 2*(3), 27.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297.

Gamboa, H. (2008). *Doctoral dissertation, PhD thesis*. Universidade Técnica de Lisboa, Instituto Superior Técnico.

Godfrey, A., Conway, R., Leonard, M., Meagher, D., & Ólaighin, G. M. (2009). A continuous wavelet transform and classification method for delirium motoric subtyping. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 17*(3), 298–307.

Golshan, H. M., Hebb, A. O., Hanrahan, S. J., Nedrud, J., & Mahoor, M. H. (2016). A Multiple Kernel Learning approach for human behavioral task classification using STN-LFP signal. In 2016 IEEE 38th annual international conference of the engineering in medicine and biology society *(EMBC)* (pp. 1030–1033).

Golshan, H. M., Hebb, A. O., Hanrahan, S. J., Nedrud, J., & Mahoor, M. H. (2017). An FFT-based synchronization approach to recognize human behaviors using STN-LFP signal. In 2017 IEEE international conference on acoustics, speech and signal processing *(ICASSP)* (pp. 979–983).

Golshan, H. M., Hebb, A. O., Hanrahan, S. J., Nedrud, J., & Mahoor, M. H. (2018). A hierarchical structure for human behavior classification using STN local field potentials. *Journal of Neuroscience Methods, 293*, 254–263.

Goodwin, M. S. (2016). 28.2 laboratory and home-based assessment of electrodermal activity in individuals with autism spectrum disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(10), S301–S302.

Greco, A., Valenza, G., Citi, L., & Scilingo, E. P. (2017). Arousal and valence recognition of affective sounds based on electrodermal activity. *IEEE Sensors Journal, 17*(3), 716–725.

Grossmann, A., & Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis, 15*(4), 723–736.

Haag, A., Goronzy, S., Schaich, P., & Williams, J. (2004). Emotion recognition using bio-sensors: First steps towards an automatic system. *Tutorial and Research Workshop on Affective Dialogue Systems*, 36–48.

Jang, E. H., Park, B. J., Kim, S. H., & Sohn, J. H. (2012). Emotion classification based on physiological signals induced by negative emotions: Discrimination of negative emotions by machine learning algorithm. In 2012 9th IEEE international conference on networking, sensing and control *(ICNSC)* (pp. 283–288).

Jang, E. H., Park, B. J., Kim, S. H., Chung, M. A., Park, M. S., & Sohn, J. H. (2014). Emotion classification based on bio-signals emotion recognition using machine learning algorithms. In 2014 international conference on information science, electronics and electrical engineering *(ISEEE): 3* (pp. 1373–1376).

Kappas, A., Küster, D., Basedow, C., & Dente, P. (2013). A validation study of the Affective Q-Sensor in different social laboratory situations. *53rd annual meeting of the society for psychophysiological research*.

Kim, K. H., Bang, S. W., & Kim, S. R. (2004). Emotion recognition system using short--term monitoring of physiological signals. *Medical and Biological Engineering and Computing, 42*(3), 419–427.

Kwon, J., Kim, D. H., Park, W., & Kim, L. (2016). A wearable device for emotional recognition using facial expression and physiological response. In 2016 IEEE 38th annual international conference of the engineering in medicine and biology society *(EMBC)* (pp. 5765–5768).

Kylliäinen, A., & Hietanen, J. K. (2006). Skin conductance responses to another person's gaze in children with autism. *Journal of Autism and Developmental Disorders, 36*(4), 517–525.

Laparra-Hernández, J., Belda-Lois, J. M., Medina, E., Campos, N., & Poveda, R. (2009). EMG and GSR signals for evaluating user's perception of different types of ceramic flooring. *International Journal of Industrial Ergonomics, 39*(2), 326–332.

Legiša, J., Messinger, D. S., Kermol, E., & Marlier, L. (2013). Emotional responses to odors in children with high-functioning autism: Autonomic arousal, facial behavior and self-report. *Journal of Autism and Developmental Disorders, 43*(4), 869–879.

Li, J., Zhang, L., Tao, D., Sun, H., & Zhao, Q. (2009). A prior neurophysiologic knowledge free tensor-based scheme for single trial EEG classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 17*(2), 107–115.

Lidberg, L., & Wallin, B. G. (1981). Sympathetic skin nerve discharges in relation to amplitude of skin resistance responses. *Psychophysiology, 18*(3), 268–270.

Lin, Y. P., Wang, C. H., Wu, T. L., Jeng, S. K., & Chen, J. H. (2007). Multilayer perceptron for EEG signal classification during listening to emotional music. In *TENCON 2007-2007 IEEE region 10 conference* (pp. 1–3). IEEE.

Liu, F., Liu, G., & Lai, X. (2014). Emotional intensity evaluation method based on Galvanic skin response signal. In *2014 seventh international symposium on computational intelligence and design (ISCID): 1* (pp. 257–261).

Liu, M., Fan, D., Zhang, X., & Gong, X. (2016). Human emotion recognition based on galvanic skin response signal feature selection and SVM. In International conference on smart city and systems engineering *(ICSCSE)* (pp. 157–160).

Lord, C., Ruter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*(5), 659–685.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11*(7), 674–693.

Mardaga, S., Laloyaux, O., & Hansenne, M. (2006). Personality traits modulate skin conductance response to emotional pictures: An investigation with Cloninger's model of personality. *Personality and Individual Differences, 40*(8), 1603–1614.

Mera, K., & Ichimura, T. (2004). Emotion analyzing method using physiological state. *Knowledge-Based Intelligent Information and Engineering Systems*, 195–201.

Najafi, B., Aminian, K., Paraschiv-Ionescu, A., Loew, F., Bula, C. J., & Robert, P. (2003). Ambulatory system for human motion analysis using a kinematic sensor: Monitoring of daily physical activity in the elderly. *IEEE Transactions on Biomedical Engineering, 50*(6), 711–723.

Nasoz, F., Alvarez, K., Lisetti, C. L., & Finkelstein, N. (2004). Emotion recognition from physiological signals using wireless sensors for presence technologies. *Cognition, Technology & Work, 6*(1), 4–14.

Oberman, L. M., Winkielman, P., & Ramachandran, V. S. (2009). Slow echo: Facial EMG evidence for the delay of spontaneous, but not voluntary, emotional mimicry in children with autism spectrum disorders. *Developmental Science, 12*(4), 510–520.

Ooi, J. S. K., Ahmad, S. A., Chong, Y. Z., Ali, S. H. M., Ai, G., & Wagatsuma, H. (2016). Driver emotion recognition framework based on electrodermal activity measurements during simulated driving conditions. In *2016 IEEE EMBS conference on biomedical engineering and sciences (IECBES)* (pp. 365–369). IEEE.

Ousley, O. Y., Arriaga, R., Abowd, G. D., & Morrier, M. (2012). Rapid assessment of social-communicative abilities in infants at risk for autism. *Technical Report, Georgia Tech*.

Perez-Gaspar, L-A., Caballero-Morales, S-O., & Trujillo-Romero, F. (2016). Multimodal emotion recognition with evolutionary computation for human-robot interaction. *Expert Systems with Applications, 66*, 42–61.

Poh, M. Z., Swenson, N. C., & Picard, R. W. (2010). A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering, 57*(5), 1243–1252.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology, 17*(3), 715–734.

Presti, L., Sclaroff, S., & Rozga, A. (2013). Joint alignment and modeling of correlated behavior streams. In *Proceedings of the IEEE international conference on computer vision workshops* (pp. 730–737).

Prince, E. B., Kim, E. S., Wall, C. A., Gisin, E., Goodwin, M. S., Simmons, E. S., et al. (2017). The relationship between autism symptoms and arousal level in toddlers with autism spectrum disorder, as measured by electrodermal activity. *Autism, 21*(4), 504–508.

Rajagopalan, S. S., Murthy, O. R., Goecke, R., & Rozga, A. (2015). Play with me—Measuring a child's engagement in a social interaction. In 2015 11th IEEE international conference and workshops on automatic face and gesture recognition *(FG): 1* (pp. 1–8).

Rajagopalan, S., Dhall, A., & Goecke, R. (2013). Self-stimulatory behaviours in the wild for autism diagnosis. In *Proceedings of the IEEE international conference on computer vision workshops* (pp. 755–761).

Rehg, J., Abowd, G., Rozga, A., Romero, M., Clements, M., & Sclaroff, S. (2013). Decoding children's social behavior. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3414–3421).

Rigas, G., Katsis, C. D., Ganiatsas, G., & Fotiadis, D. I. (2007). A user independent, biosignal based, emotion recognition method. In *International conf on user modeling* (pp. 314–318).

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161.

Sano, A., & Picard, R. W. (2011). Toward a taxonomy of autonomic sleep patterns with electrodermal activity. In *2011 Annual international conference of the IEEE engineering in medicine and biology society, EMBC* (pp. 777–780).

Schmidt, S., & Walach, H. (2000). Electrodermal activity (EDA)-State-of-the-art measurement and techniques for parapsychological purposes. *The Journal of Parapsychology, 64*(2), 139.

Shahani, B. T., Halperin, J. J., Boulu, P. H., & Cohen, J. (1984). Sympathetic skin response–a method of assessing unmyelinated axon dysfunction in peripheral neuropathies. *Journal of Neurology, Neurosurgery & Psychiatry, 47*(5), 536–542.

Sharma, V., Prakash, N. R., & Kalra, P. (2016). EDA wavelet features as Social Anxiety Disorder (SAD) estimator in adolescent females. In *2016 international conference on advances in computing, communications and informatics (ICACCI),* (pp. 1843–1846). IEEE.

Stagg, S. D., Davis, R., & Heaton, P. (2013). Associations between language development and skin conductance responses to faces and eye gaze in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 43*(10), 2303–2311.

Swangnetr, M., & Kaber, D. B. (2013). Emotional state classification in patient–robot interaction using wavelet analysis and statistics-based feature selection. *IEEE Transactions on Human-Machine Systems, 43*(1), 63–75.

Tarvainen, M. P., Karjalainen, P. A., Koistinen, A. S., & Valkonen-Korhonen, M. V. (2000). Principal component analysis of galvanic skin responses. In *Proceedings of the 22nd annual international conference of the IEEE engineering in medicine and biology society: 4* (pp. 3011–3014).

Vetterli, M., & Herley, C. (1992). Wavelets and filter banks: Theory and design. *IEEE Transactions on Signal Processing, 40*(9), 2207–2232.

Wetherby, A. M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34*(5), 473–493.

Whang, M., & Lim, J. (2008). A physiological approach to affective computing. *Affective Computing. InTech*, 309–318.

Xu, B., Fu, Y., Jiang, Y. G., Li, B., & Sigal, L. (2016). Heterogeneous knowledge transfer in video emotion recognition, attribution and summarization. *IEEE Transactions on Affective Computing*. doi:10.1109/TAFFC.2016.2622690.