Received: 27 December 2018 | Revised: 25 March 2019 | Accepted: 26 April 2019
DOI: 10.1111/exsy.12425

WILEY Expert Systems

SPECIAL ISSUE PAPER

Film mood induction and emotion classification using physiological signals for health and wellness promotion in older adults living alone

Arturo Martínez-Rodrigo[1] | Luz Fernández-Aguilar[2] | Roberto Zangróniz[3] | José M. Latorre[2] | José M. Pastor[1] | Antonio Fernández-Caballero[4,5]

[1]Departamento de Sistemas Informáticos, Universidad de Castilla-La Mancha, Cuenca, Spain
[2]Departamento de Psicología, Universidad de Castilla-La Mancha, Albacete, Spain
[3]Departamento de Ingeniería Eléctrica, Electrónica, Automática y Comunicaciones, Universidad de Castilla-La Mancha, Cuenca, Spain
[4]Departamento de Sistemas Informáticos, Universidad de Castilla-La Mancha, Albacete, Spain
[5]CIBERSAM (Biomedical Research Networking Centre in Mental Health), Madrid, Spain

Correspondence
Arturo Martínez-Rodrigo, Departmento de Sistemas Informáticos, Universidad de Castilla-La Mancha, Cuenca 16071, Spain.
Email: arturo.martinez@uclm.es

Funding information
Castilla-La Mancha Regional Government / FEDER, UE, Grant/Award Number: SBPLY/17/180501/000192; FEDER, UE, Grant/Award Number: 2018/11744; Spanish Agencia Estatal de Investigación / European Regional Development Fund, Grant/Award Number: DPI2016-80894-R

Abstract

This paper introduces a wearable hardware/software system specifically tailored to detect seven emotions (neutral, tenderness, amusement, anger, disgust, fear, and sadness) aimed at promoting health and wellness in older adults living alone at home. The complete software and hardware architectures acquiring and processing electrodermal activity and photoplethysmography signals are introduced. The wearable emotion detection system is trained by eliciting the desired emotions on 39 older adults through a film mood induction procedure. Seventeen features are calculated on skin conductance response and heart rate variability data, grouped into five statistical, four temporal, and eight morphological features. Then, these features are used to run emotion classification considering support vector machines, decision trees, and quadratic discriminant analysis. In line with psychological findings, the results offer a global accuracy of 82% in negative emotion (anger, disgust, fear, and sadness) classification. For positive emotions (tenderness and amusement), also in conformity with previous psychological outcomes, amusement shows the highest ratio of hits (92%) but tenderness the lowest one (66%). These results demonstrate that our wearable emotion detection system can be used by ageing adults, especially for detecting negative emotions that usually damage health and wellness and lead to social isolation.

KEYWORDS

emotion classification, health promotion, mood induction, older adult, physiological signals, wearable, wellness promotion

1 | INTRODUCTION

Improvements in health care systems, progress in medical treatments, and growth in long-term care assistance are augmenting life expectancy. The population is ageing and the number of people exceeding 85 years has doubled during the past decade. The ageing population will double again by year 2050, according to the World Health Organization (World Health Organization, 2011). This fact has important consequences on health care cost, affecting the quality of life of millions of people around the world (Mowafey & Gardner, 2012). In addition, an increasing number of older adults decide to stay at home, often experiencing social isolation and/or exclusion (Moore et al., 2013). Therefore, intensive research is being carried out on smart environments, home automation, ambient intelligence, e-health equipment, fitness devices, and entertainment systems for ambient assisted living (Fernández-Caballero, González, & Navarro, 2017; Hanson et al., 2009; Lago, Jiménez-Guarín, & Roncancio, 2017). Moreover, expert systems evaluating emotions would improve the ageing adult's quality of life and self-perception, saving money to health care institutions. Hence, emotion detection is currently a relevant field of study for the promotion of health and wellness.

Emotion detection requires continuous monitoring of physical or physiological variables that provide useful information about the ageing adult's emotional state (Charles & Luong, 2013). It becomes essential in ageing adults who suffer from loneliness and isolation because they are prone to develop mental disorders because of experimenting negative emotions such as disgust, sadness, anger, and/or fear. Health care institutions

and researchers are investigating potential solutions to mitigate this growing problem. Hence, wearable technology delivering physiological information (Martínez-Rodrigo et al., 2017) is a growing field of consumer electronics called to revolutionize personal health care, as it provides an inexpensive and non-intrusive alternative to continuous health monitoring. Although an important number of wearable devices are able to detect some specific emotions (mostly, distress), only a few works have evaluated a wider range of emotions (Liu & Sourina, 2014; Yuvaraj & Murugappan, 2016; Bong et al., 2017).

Unfortunately, few efforts have been made to regulate the ageing adults' emotional condition (Castillo et al., 2016; Fernández-Caballero, Latorre, Pastor, & Fernández-Sotos, 2014) or improve their autobiographical memory (Ros et al., 2018; Serrano et al., 2017), which would play a key role in health and wellness promotion. However, the affective computing paradigm enables to decide an adequate response to human emotional stimuli (Peralta, Fernández-Caballero, Latorre, Olivas, & Conde, 2018; Poria, Cambria, Bajpai, & Hussain, 2017; Sokolova & Fernández-Caballero, 2015). New architectures and methodologies are providing significant advances in this research area (Castillo, Fernández-Caballero, Castro-Gonzalez, Salichs, & López, 2014; Fernández-Caballero et al., 2016).

Different physiological variables and biopotentials captured from the human body are measured using wearables. Electrodermal activity (EDA), heart rate variability (HRV)—through processing photoplethysmography (PPG) signals—electromyography, and skin temperature are only some examples (Zangróniz, Martínez-Rodrigo, López, Pastor, & Fernández-Caballero, 2018; Zangróniz, Martínez-Rodrigo, Pastor, López, & Fernández-Caballero, 2017). EDA and HRV complementary physiological variables are linked to the central nervous system through sympathetic and parasympathetic components (Malik et al., 1996). It has been reported that EDA is exclusively linked to the sympathetic component and, therefore, with the level of alert (Zangróniz et al., 2017; Martínez-Rodrigo, Fernández-Caballero, Silva, & Novais, 2016; Martínez-Rodrigo, Zangróniz, Pastor, & Fernández-Caballero, 2015). On the other hand, HRV seems to be related to the parasympathetic component (Zangróniz et al., 2018).

This study introduces a wearable technology and an emotional model, specifically tailored to detect a series of emotions in elders. The proposed system is trained through eliciting emotions using films that have demonstrated to induce mood in ageing adults. This paper complements a previous one in which the subjective response given by participants was interpreted (Fernández-Aguilar, Ricarte, Ros, & Latorre, 2018) by means of self-assessment mannequin questionnaires (Bradley & Lang, 1994). The present study includes objective and quantitative measures using EDA and HRV variables in order to analyse the physiological emotional responses in ageing adults.

The remainder of the paper is organized as follows. Section 2 introduces the methods provided. First, the emotional model and the physiological variables are presented. Then, the acquisition and processing of physiological signals is commented. Then, the nature of the features computed on the signals are explained from time, frequency, and morphological viewpoints. Next, the features are used to feed a classification model to provide the performance of the system. Section 3 offers the results obtained from the study. Finally, Section 4 provides some conclusions.

2 | METHODS

2.1 | System architecture

2.1.1 | Software architecture

The software architecture establishes a set of levels to provide a processing hierarchy with clearly defined input/output interfaces. Each level consists of a set of functions that define the algorithms to process at each level. In this sense, the levels of the architecture establish a hierarchy from the lowest level where signals are acquired to the highest level where a decision is made about the emotion felt. The systems transmits the information throughout the levels by connecting the inputs of the immediately upper level through the outputs of the lower level, as can be observed in Figure 1a.

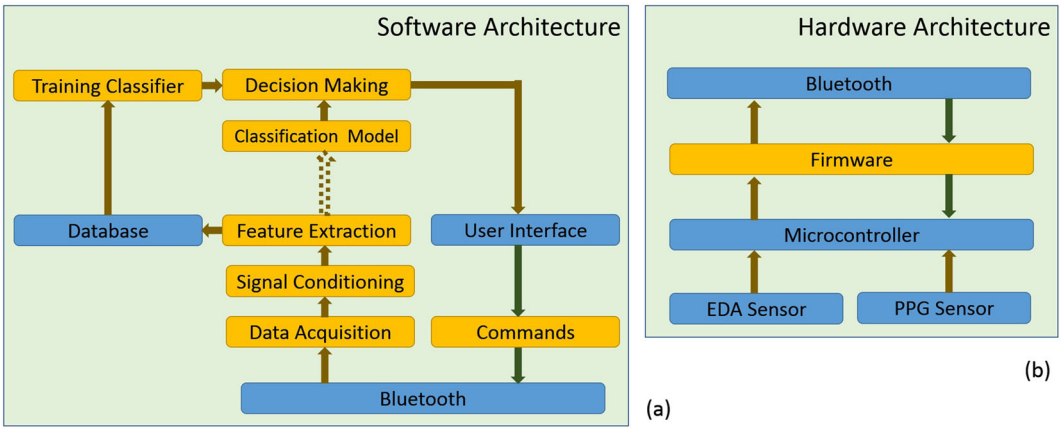

FIGURE 1 System architecture. (a) Software architecture. (b) Hardware architecture. EDA, electrodermal activity; PPG, photoplethysmography

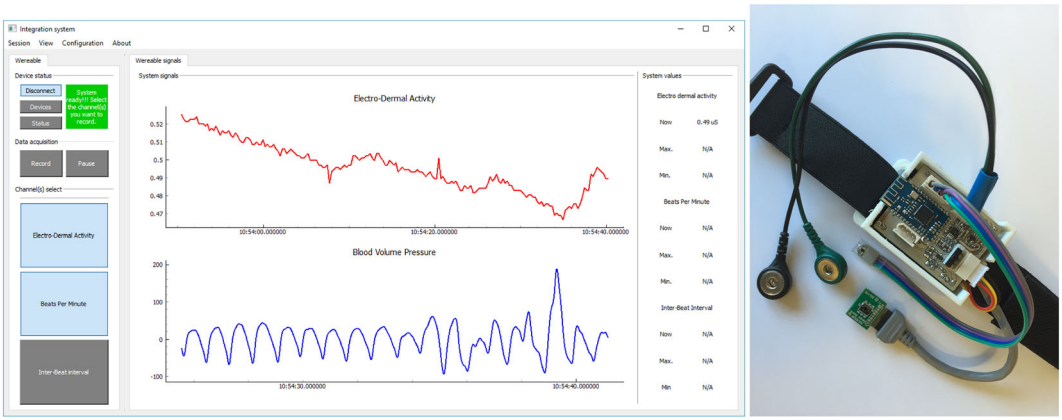

FIGURE 2 Software/hardware systems. (a) Graphical user interface. (b) Wearable device

The data acquisition level is responsible for the control and synchronization of a Bluetooth interface to receive physiological information from the wearable as well as sending control information to a microcontroller. The following level consists of signal conditioning and processing, where the signals are filtered and transformed using different techniques depending on their nature. Once the signals have been processed, they are characterized from several points of view, computing statistical, morphological, and frequency features on the physiological variables. Then, the features are stored in a database or they are used to feed the final classification model, depending on whether the system is in learning or test mode.

In learning mode, features are stored in a local database as well as the nature of the stimulus elicited on the subject. Thus, it is possible to use this information to train an advanced classifier or investigate the information that each feature is providing. In test mode, the classifier has already been trained and the features are directly used to feed the final classifier. The highest level corresponds to decision making, where a decision is taken, depending on the results provided by the classifier and the reliability of the signals, and shown through a user interface.

It is worth noting that information also flows in the opposite direction, because commands are transmitted to the wearable in order to establish connection, stop communication, or disable specific variables. Figure 2a shows the graphical user interface designed ad hoc to control the wearable device.

2.1.2 | Hardware architecture

A tailored hardware architecture, deployed in a wearable device, is proposed for the acquisition of EDA and PPG physiological signals. Figure 1b shows a general scheme of the hardware architecture. A low-power 32-bit ARM Cortex-M4 micro-controller acts as a system control unit. Regarding the acquisition sensors, EDA is measured by quantifying the direct current exosomatic EDA by means of a couple of Ag/AgCl disc electrodes with 10 mm contact diameter. The electrodes are attached to the medial phalanges in palm sides of left index and middle fingers. A single-supply, rail-to-rail input/output, precision operational amplifier implements a voltage-controlled linear current source. On the other hand, PPG signals are measured through an optical plethysmogram technique, using a light-emitting diode (LED), LED driver circuitry, and photo-detector sensors (Kim, Kim, & Ko, 2015).

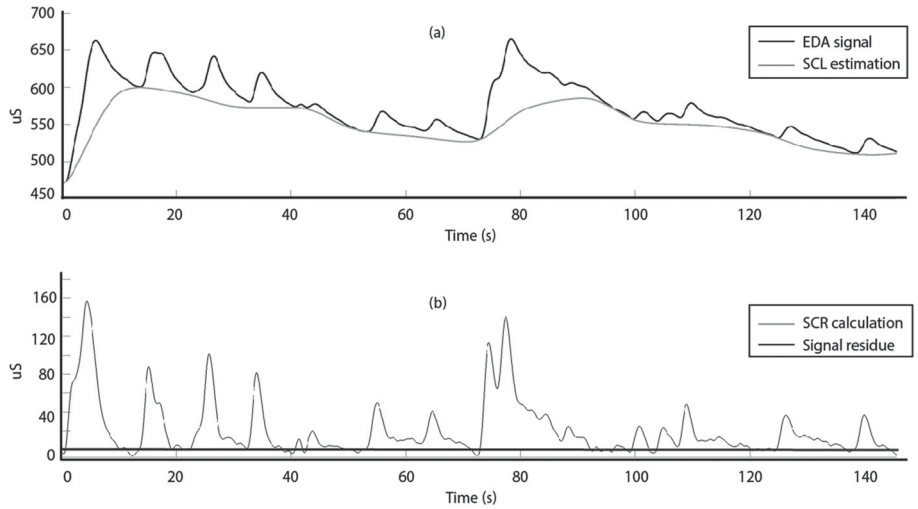

FIGURE 3 Electrodermal activity signal. (a) SCL component. (b) SCR component after deconvolution. EDA, electrodermal activity; SCL, skin conductance level; SCR, skin conductance response

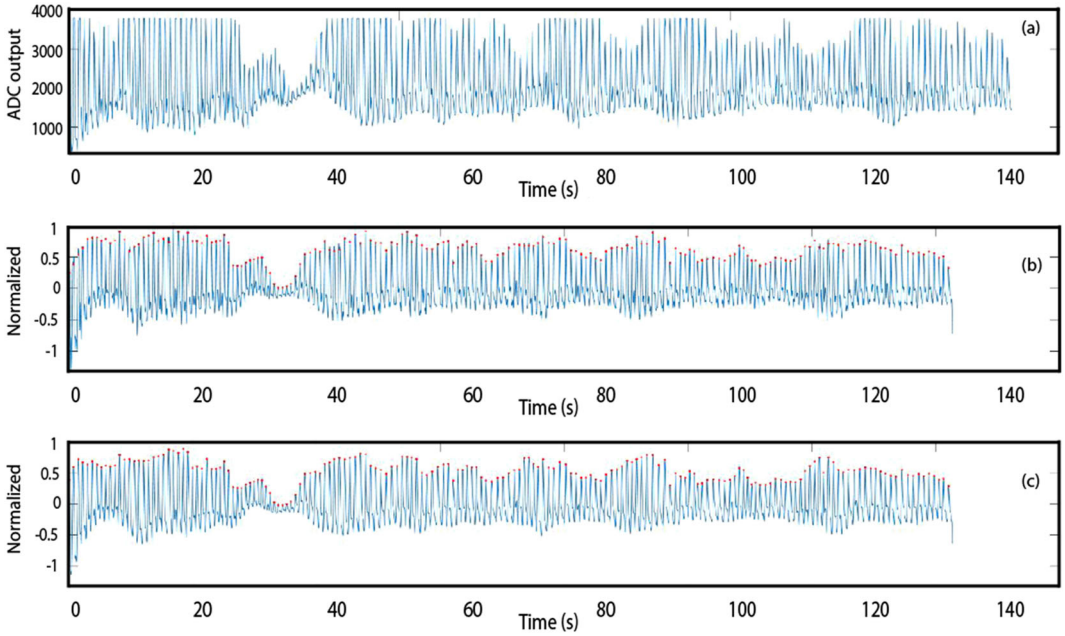

FIGURE 4 Plethysmography signal. (a) Raw photoplethysmography signal from the analog-to-digital converter (ADC). (b) Filtered, normalized, and baseline removed signal. (c) Denoised photoplethysmography signal. Peak detection enabled

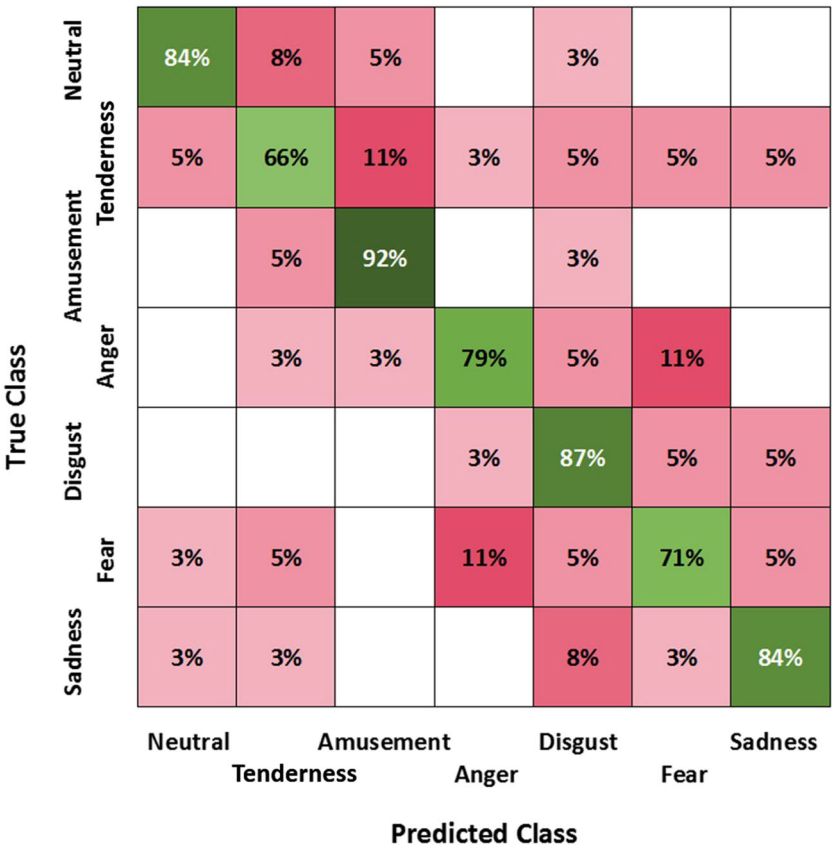

FIGURE 5 Confusion matrix

Although PPG acquisition and signal conditioning require several processing and complex circuits, there are commercial fully integrated analogue front ends and optical sensors suitable for PPG signals that simplify the hardware design. In this regard, a commercial analogue front-end and a commercial optical sensor are used in this work. Figure 2b shows the wearable prototype used in this study. Finally, raw EDA and PPG data are transmitted to the software system by means of a Bluetooth interface.

2.2 | Processing of physiological variables

2.2.1 | EDA processing

EDA signals evaluate the continuous variation in the electrical conductance of the skin. An increase in arousal level produces an activation of the sweat glands, facilitating electrical skin conduction when a subject is altered and preventing the conduction when the person is calm. Therefore, EDA signals feature a series of consecutive peaks and valleys that depend on the alert level, corresponding to the degree of activation in the sympathetic nervous system (Boucsein, 2012).

An EDA signal is composed of the overlapping of two components. On the one hand, the skin conductance level (SCL) represents the baseline of the skin conductance, which varies among people, depending on genetic aspects, physiological states, or autonomic regulation (Valenza, Lanata, & Scilingo, 2012). On the other hand, the skin conductance response (SCR) occurs when the sudomotor nerve is active, thus representing the sympathetic nervous activity. Different approaches have been evaluated in the last years to characterize EDA signals. Most methodologies are based on finding EDA peaks to measure the differences among them and the signal baseline to discriminate between SCR and SCL components. However, the response time of EDA against a specific stimulus is slow, which facilitates that consecutive stimuli occur in EDA signal as a burst of overlapped peaks. This prevents the discrimination of events using detection of peaks and valleys, because SCR boundaries persist as masked by the previous response.

Considering this inconvenience, recent studies have proposed the decomposition of EDA signals into their two components by using a deconvolution operation (Benedek & Kaernbach, 2010). This approach has achieved better performance than other methods to characterize EDA raw signals. Therefore, this operation has been performed in our study. In short, the raw EDA data $y[n]$ are first filtered by applying a 1.5- Hz cut-off low-pass finite impulse response (FIR) filter with order $N = 32$ to decrease the noise generated during the acquisition. Afterwards, a continuous decomposition analysis is performed to separate SCR and SCL, such that

$$\hat{y}[n] = (r * l)[n] = \sum_{i=0}^{N} r[n-i]l[i], \tag{1}$$

being $*$ the convolution operator in time domain, $\hat{y}[n]$ the filtered EDA signal, $r[n]$ the SCR component, and $l[n]$ the SCL component. It is worth noting that $l[n]$ corresponds to the transfer function in Equation (1), such that, if $l[n]$ is known or estimated, a deterministic deconvolution is used to compute the desired component $r[n]$ (Benedek & Kaernbach, 2010). In this sense, *Ledalab*, a MATLAB-based software toolbox for the analysis of skin conductance data, has been used to perform this operation (Ledalab, 2018). Figure 3 shows a random EDA signal filtered and decomposed into its two components.

2.2.2 | PPG processing

PPG signals consist of a blood volumetric measurement using a LED and light sensor to track heart activity. The LED sensor illuminates the skin such that a different amount of light is reflected to the light sensor, depending on the volume of arteries and arterioles caused by the pressure pulse. The resulting signal corresponds to the changes of blood volume during the cardiac cycle. Blood volume fluctuations are highly correlated with heart ventricular depolarization and repolarization, enabling the measurement of the heart's rhythm (Malik et al., 1996). Figure 4a shows a typical example of raw PPG signals.

In this work, PPG signals are acquired from the wearable at a sampling rate of 60 Hz and a 22-bit resolution to perform a correct signal acquisition without distortion, which fulfils the Nyquist criteria (Nitzan, Babchenko, & Khanokh, 1999). Several factors such as sensor location, electrical sources, ambient lights, skin properties, and temperature affect the quality of PPG signals (Elgendi, 2012). These factors add several artefacts and noisy components to the waveform, which intensifies difficulty to signal explanation. Indeed, power line interference is one of the most common noise sources, which is present in the ambience, modulating the PPG signal over a sinusoidal component at its fundamental frequency. Besides, variation in temperature or poor contact of the photo-sensor are only some sources that affect PPG signal. Hence, baseline wander, high-frequency noise, and power-line interference are eliminated by performing forward/backward filtering. Specifically, baseline drift is removed by carrying out a 0.5-Hz cut-off high-pass, linear-phase FIR filter. Similarly, a 30- Hz cut-off low-pass, linear-phase FIR filter eliminates high-frequency noise and power-line interferences. Figure 4b shows the postprocessed signal when different filtering and processing techniques are used to eliminate potential interferences.

PPG local maximums are related to pulse pumping. They are widely used to measure the cardiac pulse in a long-term fashion Allen (2007). This is why procedures based on adaptive thresholding are extensively performed to detect the peaks on PPG signals (Shin et al., 2009). Nonetheless, PPG signals may contain inherent noise, sudden amplitude changes and different morphologies due to premature ventricular contractions or movements. Therefore, a robust and reliable peak detection algorithm, capable of dealing with different morphologies, is necessary to face these questions. In our design, a peak detector based on the phasor transform is put into use (Martínez-Rodrigo et al., 2010). The algorithm normalizes the signal, regardless of its amplitude and morphology, to be suitable to assess PPG signals. Figure 4c shows the detection of PPG signal peaks by using this technique. Finally, HRV is estimated by measuring the time variation in consecutive PPG peaks. The characteristics extracted by analysing HRV have been widely used to investigate the sympathetic and parasympathetic function of the autonomic nervous system (Malik et al., 1996; Malik, 1996).

2.3 | Experiment description

2.3.1 | Participants

Thirty-nine older adults (26 women and 13 men) aged 60 to 84 ($M = 68.51, SD = 6.66$) were enrolled in this study. Participants were recruited from University of Experience, Albacete, Spain (academic courses for ageing adults). The group had similar education years ($M = 13.33, SD = 3.08$). Moreover, they were receiving no psychotropic or drug treatment and had no previous history of psychological, psychiatric, or neurological disorder, according to the criteria of the *Diagnostic and Statistical Manual of Mental Disorders*, *Fifth Edition* (American Psychiatric Association, 2013). Moreover, none of the participants were taking any medication that could affect the results of the task. They presented no auditory or visual impairments other than requiring corrective lenses. All were of Caucasian ethnicity and native Spanish speakers. All participants gave voluntary consent to taking part in the study, according to the requirements of the approved ethics procedure granted from Clinical Research Ethics Committee of the University Hospital at Albacete, Spain.

As cognitive impairment and depressive symptoms may affect the emotional response, a paper-and-pencil version of tests was administered prior to the experiment. Mini-mental state examination (MMSE; Folstein et al. 1975) was administered to assess cognitive impairment. MMSE is a screening tool used to measure decline in cognitive abilities such as alterations in memory. People with cognitive impairment (MMSE score lower than 27) were excluded from our study. Moreover, to assess depressive symptoms, we administered the Beck Depression Inventory II (Beck et al., 1961). Beck Depression Inventory II is a self-report questionnaire that assesses symptoms of depression such as anhedonia and sadness. The positive and negative affect schedule-state version (Watson, Clark, & Tellegen, 1988) was used to assess the general affect (positive and negative) before the experimental session.

2.3.2 | Mood induction procedure

We selected 54 excerpts from high-definition clips previously validated in young Spanish adults (Fernández, Mateos, Ribaudi, & Fernández-Abascal, 2011). The selected scenes maintained the same features used in similar studies (e.g., Rottenberg, Ray, & Gross, 2007; Schaefer, Nils, Sanchez, & Philippot, 2010). We selected film clips to elicit the following seven target emotions: sadness, disgust, fear, anger, amusement, tenderness, and neutral. The entire duration of the experimental task was approximately 1 hr. The task was programmed and administered using E-prime 2 Professional (Psychology Software Tools, Inc.). Following the recommendations of previous studies (Rottenberg et al., 2007), the experiment was conducted in a room with dimmed lighting on a 20'' computer screen situated in a laboratory (8 m^2).

First, participants were informed that some film clips might portray potentially shocking scenes and were informed that they were free to withdraw from the experiment at any time. At the start of the session, the wearable was placed while the experiment procedure was explained. Participants were informed that the aim of the study is to broaden knowledge about emotions. Thus, they should express what they felt during the exposure to each film clip.

A practice film clip was used to demonstrate how the rating would be performed. After this trial, participants started an individual session by watching a neutral clip to establish the baseline of the physiological parameters. In total, each participant watched a set of nine film clips including six emotional targets (sadness, disgust, fear, anger, amusement, and tenderness; Fernández-Aguilar et al. 2018; Fernández-Aguilar et al. 2016). The order of the film clips was counterbalanced. Each trial started with a fixation cross in the centre of the screen for 500 ms, followed by an emotional clip. Also, they had to complete a distraction task to prevent the excitation transfer from one emotion to another. The distraction task consisting of distinguishing geometric figures on the screen. At the end of the session, another neutral clip was presented to recover a relaxed state.

2.4 | Feature extraction

With the aim of characterizing the processed information, different features were calculated on SCR and HRV data series. Table 1 shows a summary of the feature nomenclature and a brief description of their meaning. These features were chosen with the aim of covering a wide range of approaches, because the most relevant features for constructing the model are unknown a priori. Hence, several features were selected from statistical, temporal, and morphological perspectives to maximize the chance of finding independent and informative discriminants.

It is worth noting that human reaction against a specific stimulus is reflected in the SCR signal as a peak or a burst of peaks, according to the degree of alert. The higher the alteration caused, the higher the peak produced in SCR data, and the more reactions against a continuous stimulus, the more peaks in the SCR signal. In a similar way, when valence alterations are produced, they are reflected in the HRV signal as peaks or burst of peaks quantifying the degree of time differences throughout consecutive heartbeats. Higher differences in heart variability cause more pronounced peaks, and the more episodes are produced in a given time interval, the more peaks appear in the HRV signal. From a physiological point of view, both signals are quite similar, as no reaction is identified like a continuous baseline, and reactions against stimuli are drawn in the signals as peaks proportional to the intensity, duration, and number of emotional events. Considering the data to be analysed as

$$r[n] = (y[n_0], y[n_0 + 1], \ldots, y[n_e]), \quad for \quad n = 1, 2, \ldots, L, \tag{2}$$

being L its total length in samples, the statistical measures such as mean (MEN), standard deviation (STD), and dynamic range (DRG) can provide global relevant information on the average and variability of data series, respectively. Moreover, maximum (MAX) or minimum (MIN) values can

TABLE 1 Features and their description

	Feature	Description
Statistical feature	MEN	Mean
	STD	Standard deviation
	MAX	Maximum
	MIN	Minimum
	DRG	Dynamic range
	ASY	Asymmetry
	KUR	Kurtosis
Temporal feature	FRD	Mean of the first derivative
	FRS	Standard deviation of the first derivative
	SDM	Mean of the second derivative
	SDS	Standard deviation of the second derivative
Morphological feature	ARC	Arc-length
	INT	Integral
	POT	Potency
	RMS	Root mean square
	APR	Area-perimeter ratio
	EPR	Energy-perimeter ratio

TABLE 2 Partial statistical significance contribution of each feature to the final model

Signal	Feature	Partial *p* value	Global *p* value
EDA	ASY	1.50×10^{-6}	
HRV	FRD	1.67×10^{-6}	
	ARC	0.007	
	RMS	0.013	
	APR	0.011	
EDA + HRV	Combination		1.39×10^{-19}

Abbreviations: EDA, electrodermal activity; HRV, heart rate variability; ASY, asymmetry; KUR, kurtosis; FRD, mean of the first derivative; ARC, arc length; RMS, root mean square; APR, area-perimeter ratio.

TABLE 3 Classifiers' performance

Classifiers		Emotions						
		Neutral	Tenderness	Amusement	Anger	Disgust	Fear	Sadness
QDA	TPR (%)	79	37	71	32	24	45	34
(Acc = 45.9%)	FNR (%)	21	63	29	68	76	55	66
DTC	TPR (%)	89	68	84	58	53	71	61
(Acc = 69.2%)	FNR (%)	11	32	16	42	47	29	39
SVM	TPR (%)	84	66	92	79	87	71	84
(Acc = 80.5%)	FNR (%)	16	34	8	21	13	29	16

Abbreviations: QDA, quadratic discriminant analysis; DTC, decision tree-based classifier; SVM, support vector machine; TPR, true positive ratio; FNR, false negative ratio.

produce specific information on the highest or lowest reaction obtained across the data, which may differ depending of the stimulus nature. Finally, third and fourth statistics can provide information about the location and variability of data series, respectively, which are computed by means of skewness (ASY) and kurtosis (KUR).

When a stimulus is more intensive, it is reflected as a sudden rise slope in the signal. On the contrary, when the stimulus is less intensive, it is reflected as a raising smoother slope. Once the reaction has reached its maximum, the recovery time needed to reach the zero line is represented by the down slope that is different from one stimulus to another. Consequently, quantifying the velocity and acceleration of the slopes forming the bumps can provide important information to discriminate among different reactions. Thus, temporal raw data were transformed into its first and second derivative in order to highlight the different changes in the forming slopes, such that

$$d1 = \sum_{n=2}^{L} (r[n] - r[n-1]), \tag{3}$$

$$d2 = \sum_{n=2}^{L} (d1[n] - d1[n-1]). \tag{4}$$

Notice that, as the duration of the stimulus can variate, mean and standard deviation of derivatives were computed in order to normalize the results regardless of their longitude, as shown below.

$$FRD = \frac{1}{L}\sum_{n=1}^{L} d1[n], \tag{5}$$

$$SMD = \frac{1}{L}\sum_{n=1}^{L} d2[n], \tag{6}$$

$$FRS = \sqrt{\frac{1}{L}\sum_{n=1}^{L}(d1[n] - \bar{d1})}, \tag{7}$$

$$SDS = \sqrt{\frac{1}{L}\sum_{n=1}^{L}(d2[n] - \bar{d2})}, \tag{8}$$

where $\bar{d1}$ and $\bar{d2}$ correspond to the mean of $d1$ and $d2$, respectively. Another way to measure the level of reaction consists in characterizing the SCR events or peaks by using shape or morphological markers. Thus, the more intensive and numerous the emotional events are, the more the SRC and HRV signal are altered, regarding to their zero line. In this regard, arc length or perimeter (ARC) may provide significant information on the level of alteration, as this parameter measures the rectified signal length. It can be mathematically expressed as

$$ARC = \frac{1}{L}\sum_{n=2}^{L}\sqrt{1 + (r[n] - r[n-1])^2}. \tag{9}$$

In addition, some additional parameters related to the events' amplitude were used to quantify other possible alterations. Thus, the normalized area (INT), the normalized potency (POT), and the normalized root mean square value (RMS) were computed as

$$INT = \frac{1}{L}\sum_{n=1}^{L}|r[n]|, \tag{10}$$

$$POT = \frac{1}{L}\sum_{n=1}^{L} r[n]^2, \tag{11}$$

$$RMS = \sqrt{\frac{1}{L}\sum_{n=1}^{L} r[n]^2}. \tag{12}$$

Finally, considering that the signals-event amplitude can directly affect its arc length and its energy, their relationship was analysed by normalizing RMS and INT with respect to the rectified signal length by quantifying the area-perimeter (APR) and energy-perimeter (EPR) rates as shown below.

$$APR = \frac{INT}{ARC}, \tag{13}$$

$$EPR = \frac{RMS^2}{ARC}. \tag{14}$$

2.5 | Statistical analysis

A multiparametric model had to be considered to build an algorithm able to discriminate among these different emotions. Seventeen features were calculated for each physiological variable, which supposed a considerable amount of information to manage simultaneously. In order to reduce the amount of information and to design a more efficient algorithm, it was necessary to test the single information provided by each parameter to the final classification model. Indeed, it is desirable to discard redundant information and retain interconnected data. Therefore, a step-wise regression (SWR) scheme was used to choose the more relevant features in the multivariate model. An SWR consists of a systematic method for adding and removing terms from a multilinear model based on their statistical significance in a regression. In this study, the SWR scheme was configured with values of statistical significance $p \leq 0.05$ and $p \geq 0.1$ for introducing and removing variables from the resulting discriminant model, respectively. Indeed, the highly correlated variables would stay out from the analysis, and thus exclusively, the most relevant features were kept, providing additional information to the regression model.

Furthermore, these variables were latter used to feed advanced classifiers (Jha, Pan, Elahi, & Patel, 2018) and run emotion classification. More concretely, a support vector machine (SVM), a decision tree classifier (DTC), and a quadratic discriminant analysis (QDA) were considered. It is worth noting that the SVM classifier was run using a cubic kernel function and kernel scale one. Regarding the DTC classifier, the nodes' growth was stopped when each node solely contained either fragments from only one group or a sample number less than 20% of the entire data set. Moreover, every node was split by using an impurity-based Gini index. Finally, it is important to remark that machine learning was performed using MATLAB® and Statistics and Machine Learning Toolbox™, release 2018a.

3 | RESULTS

3.1 | Study of optimal features

Different analyses were performed to examine the relevant features that provide meaningful information to the model. First, an SWR was performed using exclusively features extracted from EDA signals. In this case, only morphological parameters ASY ($p = 1.5 \times 10^{-6}$) and KUR ($p = 0.0001$) were considered as significant in the model, which scored a total statistical significance of $p = 5.03 \times 10^{-9}$. Then, a new model was constructed considering exclusively features extracted from HRV data. In this case, four out of 17 features were included in the model. More concretely, we included temporal feature FRD ($p = 1.70 \times 10^{-10}$), morphological characteristics ARC ($p = 0.0017$), RMS ($p = 0.0076$), and APR ($p = 0.016$), which achieve a combined statistical significance of $p = 2.14 \times 10^{-17}$. Finally, another model was tested by including the 34 parameters from EDA and HRV data together. In this circumstance, a total of five parameters were chosen, raising the combined statistical significance up to $p = 1.39 \times 10^{-19}$.

Table 2 shows the chosen features and their partial statistical significance contribution to the final model. As can be observed in bold, morphological parameter ASY ($p = 1.5 \times 10^{-6}$) was used from EDA signal, whereas the rest of chosen parameters were computed from HRV series, ranging from temporal features as FRD ($p = 1.67 \times 10^{-6}$) to morphological characteristics as ARC ($p = 0.007$), RMS ($p = 0.013$), and APR ($p = 0.011$). This combination shows the existence of complementary information among EDA and HRV variables when explaining the different emotional states. Indeed, the final model almost consists of a combination of the first two models, where the features provide additional information among different features and the two physiological variables.

3.2 | Multiparametric analysis and advanced classification

The most significant five features chosen in the regression model were used as input to several advanced classifiers with the aim to improve the overall classification performance. Nevertheless, the nonequal sample distribution of participants (26 women and 13 men) could lead to bias due to the imbalance caused by the different contribution of both genres. For this reason, a one-way multivariate analysis of variance for independent samples was performed for each emotion, grouping the samples according with their gender. Results obtained after performing this analysis showed that there were no significant differences between men and women for any emotion ($p < 0.05$).

Table 3 shows the results of global accuracy for QDA, DTC, and SVM classifiers in terms of true positive rates (TPR) and false negative rates (FNR) for each emotion under study. As can be observed, QDA classifier scored 45.9% global accuracy, where only "neutral" and "amusement" achieved an acceptable 79% and 71% ratio of success, respectively. However, the TPRs on the rest of emotional states are below 50%. On the other hand, DTC classifier achieved 69.2% global performance, being all TPRs above 50%. It is worth noting that "neutral" and "amusement" reached 89% and 84% TPR, respectively. It also deserves mention that "fear" increased 16% regarding the QDA classifier, reaching 71% TPR. However, the rest of emotions achieved a more discreet success ratio, ranging from 53% to 68%. Finally, SVM classifier scored 80.5% global accuracy, improving more than 11% regarding the DTC classifier. "Neutral," "tenderness," and "fear" achieved a performance comparable with the DTC classifier, scoring TPRs of 84%, 66%, and 71%, respectively. Whereas "amusement" increased up to 92% TPR, valuable emotions as "anger" and "sadness" increased their success scoring 21% and 23%, respectively. In this line, it is remarkable that "disgust" achieved a 87% TPR, increasing more than 30% in regard to the DTC classifier.

It is important to remark that the comparison of results among different works has to be considered with caution, because different ways to elicit emotions may trigger different cognitive processes (Valenza, Lanata, & Scilingo, 2012). Thus, there is no agreement in determining when a specific stimulus is suitable for the elicitation of a concrete emotion. In this sense, it has been not corroborated what type of stimuli is the best to provoke a clear response linked to the nature of the stimulus. Commonly, the reaction of each person against a specific stimulus depends on life experiences or memories and may be thoroughly opposite from one person to another. Considering this preamble, most works based on wearables found in the literature that used EDA and HRV variables have been designed exclusively to detect emotions that provoke negative stress, achieving accuracy rates that range from 78% to 97% (Healey & Picard, 2005; Sharma & Gedeon, 2013; Sandulescu, Andrews, Ellis, Bellotto, & Mozos, 2015; Salafi & Kah, 2015).

Nevertheless, different methodologies, number of features, variables, and classifiers have been used in the previous works. Thus, hand and foot EDA, together with HRV, electromyography, and respiration were used (Healey & Picard, 2005). Similarly, EDA, blood pressure, HRV, eye gaze, and pupil dilation were used (Sharma & Gedeon, 2013). The single paper from the previous relation that only based on EDA and HRV (Sandulescu et al., 2015) offers an accuracy of 78%. Considering together the negative emotions that may cause negative stress, the present system achieved a global accuracy of 82%. It is important to remark that unlike other works, four negative emotions (sadness, fear, disgust, and anger) have been considered to perform this calculation, thus providing a better approximation to a real scenario.

Moreover, by means of a confusion matrix, Figure 5 shows the performce of the SVM classifier. "Tenderness" emotion had the worst performance, where several FNRs were spread out among the rest of emotions, especially to "amusement" where 11% individuals were misclassified. In this regard, it seems logical that "amusement" and "tenderness" were badly classified, because both emotions share a high degree of valence component and they differ slightly in arousal level. Indeed, "amusement" scored the highest rate of success, but the model misclassified 5% individuals as experiencing "tenderness" emotion. Similarly, "fear" and "anger" shared a high component of arousal level and

both shared a displeasure component, making that 11% of the cases had mixed each other. Finally, ''sadness'' and ''disgust'' shared almost similar degrees of unpleasantness with some variations in their alert level. This circumstance was revealed in the number of misclassified shared cases, where 8% emotions elicited with ''sadness'' were classified as ''disgust'' and 5% emotions elicited as ''disgust'' were classified as ''sadness.''

These results should be addressed through psychological findings and evidences. Indeed, responses from the autonomic nervous system vary as a function of the valence and arousal level of affective stimuli. Even beyond the dimensions associated with the emotional response, there is a substantial support to the association between specific autonomic response patterns and discrete emotional states (McGinley & Friedman, 2017; Stephens, Christie, & Friedman, 2010).

The results have shown that the present system achieved a global accuracy of 82% in negative emotion classification. Specific emotions related to negative social interactions have shown a remarkable discrimination. These results can be related to many of the theoretical models in ageing specificity in predictive emotional response (Carstensen, 2006; Charles & Luong, 2013). For example, several studies indicate that older people have a greater emotional response to stimuli that involve irreversible personal losses or social injustice (Kunzmann & Grühn, 2005; Seider et al., 2010). Thus, when older adults watch clips of social or emotional injustice, they respond with intense unpleasantness, for example, in the presence of sadness and anger stimuli. They respond in that way because they are more sensitive to socioemotional information than other negative information (Fernández-Aguilar et al., 2018).

The results have shown different trends for positive emotions (tenderness and amusement). Both represent two distinct positive categories evidenced in previous research on emotion categorization (Shaver et al., 1987). The study of positive emotions through mood induction procedures is scarce, but some conclusions can be derived. Previous research has shown that amusement is an emotion with a high rate of correct classification (e.g., McGinley & Friedman, 2017). However, tenderness has not been correctly classified. In part, this result may be related to the fact that tenderness is an attachment-related emotion. It is the representation of affect, closeness, and affiliation, among others (Depue & Morrone-Strupinsky, 2005; Schaefer et al., 2010). Thus, tenderness could imply a mix of physiological emotional responses harder to discriminate.

4 | CONCLUSIONS

In this work, a complete system composed of hardware, control software, signal processing, and classification model has been considered to deploy a wearable with a remarkable ability to discriminate among seven considered emotional states (neutral, tenderness, amusement, anger, disgust, fear, and sadness). Moreover, contrary to most algorithms usually trained with a nontargeted sample, in this study, the experimentation and results are completely focused on older people. Indeed, it has been already reported that a reaction against a specific stimulus of a young individual is different from the reaction shown by an older person (Fernández-Aguilar et al., 2018). Therefore, most models by other authors cannot be extrapolated. On the other hand, the ageing subjects selected in this study may have a low risk of isolation, which could be considered a limitation of the study. These participants form up a representative sample of active and socially engaged adult people. In our opinion, this is important in a first step to demonstrate the proper functioning of the system in healthy ageing adults. Future studies using this system should focus in older adults at true risk of isolation to better understand differences in emotional reactions between more active and passive older adults.

Different features and classifiers have been considered in this study. However, only five out of 34 features have demonstrated to be significant in our multiparametric model. Indeed, much effort has been dedicated to keep the design as simple as possible to decrease the computational burden of the global system. It is crucial to reduce the amount of information to be simultaneously processed. Thus, it would be possible to develop a real time system, not only in a distributed architecture, but also embedded in a single wearable device. Regarding the classifiers, SVM has shown the highest accuracy and performance in classifying correctly the emotions. However, beyond the achievement of this classifier, it has been important to maximize the number of hits, especially in emotions with a negative context. Certainly, a true detection of negative emotions is crucial to prevent that ageing adults fall into negative mental states. Consequently, an increase in the true positive ratio of negative emotions such as sadness, fear, disgust, and anger is desirable, much more than other neutral or positive emotions such as tenderness and amusement.

The classification results for each detected emotion have been explained through previous psychological evidences. The results have shown that the present system achieved 82% global accuracy for negative emotions. This value is in line with some theoretical models predicting emotional responses in ageing adults. Nevertheless, the results have shown a different trend for positive emotions. As in some previous research, amusement reached a high number of hits, whereas tenderness received the lowest rate of hits. The previous results demonstrate that the proposed wearable emotion detection system can be used by ageing adults, especially for detecting negative emotions that usually deteriorate health and wellness and lead to social isolation.

ACKNOWLEDGEMENTS

This work was partially supported by Spanish Ministerio de Ciencia, Innovación y Universidades, Agencia Estatal de Investigación (AEI) / European Regional Development Fund (FEDER, UE) under DPI2016-80894-R grant, by Castilla-La Mancha Regional Government / FEDER, UE, under SBPLY/17/180501/000192 and SBPLY/17/180501/000411 grants, by European Regional Development Fund (FEDER, EU) under 2018/11744 grant, and by Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM) of the Instituto de Salud Carlos III.

CONFLICT OF INTEREST

The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results.

ORCID

Arturo Martínez-Rodrigo https://orcid.org/0000-0003-2343-3186
Luz Fernández-Aguilar https://orcid.org/0000-0002-0411-0587
Roberto Zangróniz https://orcid.org/0000-0002-0334-8155
José M. Latorre https://orcid.org/0000-0002-6159-5074
José M. Pastor https://orcid.org/0000-0001-5346-996X
Antonio Fernández-Caballero https://orcid.org/0000-0002-8211-0398

REFERENCES

Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, *28*(3), R1.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders: DSM-5*. Washington, DC: APA Publishing.

Beck, A., Ward, C., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, *4*(6), 561–571.

Benedek, M., & Kaernbach, C. (2010). A continuous measure of phasic electrodermal activity. *Journal of Neuroscience Methods*, *190*(1), 80–91.

Bong, S. Z., Wan, K., Murugappan, M., Ibrahim, N. M., Rajamanickam, Y., & Mohamad, K. (2017). Implementation of wavelet packet transform and non linear analysis for emotion classification in stroke patient using brain signals. *Biomedical Signal Processing and Control*, *36*, 102–112.

Boucsein, W. (2012). *Electrodermal Activity*. New York, NY, US: Springer Science & Business Media.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*(1), 49–59.

Carstensen, L. L. (2006). The influence of a sense of time on human development. *Science*, *312*(5782), 1913–1915.

Castillo, J. C., Castro-González, A., Fernández-Caballero, A., Latorre, J. M., Pastor, J. M., Fernández-Sotos, A., & Salichs, M. A. (2016). Software architecture for smart emotion recognition and regulation of the ageing adult. *Cognitive Computation*, *8*(2), 357–367.

Castillo, J. C., Fernández-Caballero, A., Castro-Gonzalez, A., Salichs, M., & López, M. (2014). A framework for recognizing and regulating emotions in the elderly. In Pecchia, L., Chen, L. L., Nugent, C., & Bravo, J. (Eds.), *Ambient Assisted Living and Daily Activities*. Cham: Springer International Publishing, pp. 320–327.

Charles, S. T., & Luong, G. (2013). Emotional experience across adulthood: The theoretical model of strength and vulnerability integration. *Current Directions in Psychological Science*, *22*(6), 443–448.

Depue, R. A., & Morrone-Strupinsky, J. V. (2005). A neurobehavioral model of affiliative bonding: Implications for conceptualizing a human trait of affiliation. *Behavioral and Brain Sciences*, *28*(3), 313–349.

Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. *Current Cardiology Reviews*, *8*(1), 14–25.

Fernández, C., Mateos, J. C., Ribaudi, J., & Fernández-Abascal, E. (2011). Spanish validation of an emotion-eliciting set of films. *Psicothema*, *23*(4), 778–785.

Fernández-Aguilar, L., Latorre, J. M., Ros, L., Serrano, J. P., Ricarte, J. J., Martínez-Rodrigo, A., ..., & Fernández-Caballero, A. (2016). Emotional induction through films: A model for the regulation of emotions. In Chen, Y.-W., Tanaka, S., Howlett, R. J., & Jain, L. C. (Eds.), *Innovation in Medicine and Healthcare 2016*. Cham: Springer International Publishing, pp. 15–23.

Fernández-Aguilar, L., Ricarte, J., Ros, L., & Latorre, J. M. (2018). Emotional differences in young and older adults: Films as mood induction procedure. *Frontiers in Psychology*, *9*, 1110.

Fernández-Caballero, A., González, P., & Navarro, E. (2017). Gerontechnologies—Current achievements and future trends. *Expert Systems*, *34*(2), e12203.

Fernández-Caballero, A., Latorre, J. M., Pastor, J. M., & Fernández-Sotos, A. (2014). Improvement of the elderly quality of life and care through smart emotion regulation. In Pecchia, L., Chen, L. L., Nugent, C., & Bravo, J. (Eds.), *Ambient Assisted Living and Daily Activities*. Cham: Springer International Publishing, pp. 348–355.

Fernández-Caballero, A., Martínez-Rodrigo, A., Pastor, J. M., Castillo, J. C., Lozano-Monasor, E., López, M. T., ..., & Fernández-Sotos, A. (2016). Smart environment architecture for emotion recognition and regulation. *Journal of Biomedical Informatics*, *64*, 55–73.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Hanson, M. A., Powell, H. C. Jr., Barth, A. T., Ringgenberg, K., Calhoun, B. H., Aylor, J. H., & Lach, J. (2009). Body area sensor networks: Challenges and opportunities. *Computer*, *42*(1), 58–65.

Healey, J. A., & Picard, R. W. (2005). Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, *6*(2), 156–166.

Jha, S. K., Pan, Z., Elahi, E., & Patel, N. (2018). A comprehensive search for expert classification methods in disease diagnosis and prediction. *Expert Systems*, e12343.

Kim, J., Kim, J., & Ko, H. (2015). Low-power photoplethysmogram acquisition integrated circuit with robust light interference compensation. *Sensors*, *16*(1), 46.

Kunzmann, U., & Grühn, D. (2005). Age differences in emotional reactivity: The sample case of sadness. *Psychology and Aging*, *20*(1), 47–59.

Lago, P., Jiménez-Guarín, C., & Roncancio, C. (2017). Contextualized behavior patterns for change reasoning in ambient assisted living: A formal model. *Expert Systems*, *34*(2), e12163.

Ledalab (2018). Introduction. http://www.ledalab.de/

Liu, Y., & Sourina, O. (2014). EEG-based subject-dependent emotion recognition algorithm using fractal dimension. In *IEEE International Conference on Systems, Man and Cybernetics*, IEEE. Proceedings. Vol. 4: San Diego, California, USA, 5 - 8 October 2014. pp. 3166–3171.

Malik, M. (1996). Heart rate variability. *Annals of Noninvasive Electrocardiology*, *1*(2), 151–181.

Malik, M., Bigger, J. T., Camm, A. J., Kleiger, R. E., Malliani, A., Moss, A. J., & Schwartz, P. J. (1996). Heart rate variability standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, *17*, 354–381.

Martínez-Rodrigo, A., Alcaraz, R., & Rieta, J. J. (2010). Application of the phasor transform for automatic delineation of single-lead ecg fiducial points. *Physiological Measurement*, *31*(11), 1467.

Martínez-Rodrigo, A., Zangróniz, R., Pastor, J. M., & Sokolova, M. V. (2017). Arousal level classification of the aging adult from electro-dermal activity: From hardware development to software architecture. *Pervasive and Mobile Computing*, *34*, 46–59. Pervasive Computing for Gerontechnology.

Martínez-Rodrigo, A., Fernández-Caballero, A., Silva, Fábio, & Novais, Paulo (2016). Monitoring electrodermal activity for stress recognition using a wearable, *Ambient Intelligence and Smart Environments*. Cham: Springer International Publishing, pp. 416–425.

Martínez-Rodrigo, A., Zangróniz, R., Pastor, J. M., & Fernández-Caballero, A. (2015). Arousal level classification in the ageing adult by measuring electrodermal skin conductivity. In Bravo, J., Hervás, R., & Villarreal, V. (Eds.), *Ambient Intelligence for Health*. Cham: Springer International Publishing, pp. 213–223.

McGinley, J. J., & Friedman, B. H. (2017). Autonomic specificity in emotion: The induction method matters. *International Journal of Psychophysiology*, *118*, 48–57.

Moore, P., Xhafa, F., Barolli, L., & Thomas, A. (2013). Monitoring and detection of agitation in dementia: Towards real-time and big-data solutions. In *2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing*, Compiegne, France 28-30 October 2013. pp. 128–135.

Mowafey, S., & Gardner, S. (2012). A novel adaptive approach for home care ambient intelligent environments with an emotion-aware system. In *Proceedings of 2012 UKACC International Conference on Control*, Cardiff, UK, 3-5 September 2012. pp. 771–777.

Nitzan, M., Babchenko, A., & Khanokh, B. (1999). Very low frequency variability in arterial blood pressure and blood volume pulse. *Medical & Biological Engineering & Computing*, *37*(1), 54–58.

Peralta, A., Fernández-Caballero, A., Latorre, J. M., Olivas, J. A., & Conde, E. (2018). Use of soft-computing techniques to study the influence of external factors during the emotional evaluation of visual stimuli. *Journal of Electrical Engineering*, *6*, 116–123.

Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). A review of affective computing: From unimodal analysis to multimodal fusion. *Information Fusion*, *37*, 98–125.

Ros, L., Romero, D., Ricarte, J. J., Serrano, J. P., Nieto, M., & Latorre, J. M. (2018). Measurement of overgeneral autobiographical memory: Psychometric properties of the autobiographical memory test in young and older populations. *PLOS ONE*, *13*(4), 1–18.

Rottenberg, J., Ray, R. D., & Gross, J. J. (2007). Emotion elicitation using films. In Coan, J. A., & Allen, J. J. B. (Eds.), *The handbook of emotion elicitation and assessment*. London: Oxford University Press.

Salafi, T., & Kah, J. C. Y. (2015). Design of unobtrusive wearable mental stress monitoring device using physiological sensor. In Goh, J., & Lim, C. T. (Eds.), *7th WACBE World Congress on Bioengineering 2015*. Cham: Springer International Publishing, pp. 11–14.

Sandulescu, V., Andrews, S., Ellis, D., Bellotto, N., & Mozos, O. (2015). Stress detection using wearable physiological sensors. In Ferrández, J. M., Álvarez-Sánchez, J. R., de la Paz, F., Toledo-Moreo, F. J., & Adeli, H. (Eds.), *Artificial Computation in Biology and Medicine*. Cham: Springer International Publishing, pp. 526–532.

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition and Emotion*, *24*(7), 1153–1172.

Seider, B. H., Shiota, M. N., Whalen, P., & Levenson, R. W. (2010). Greater sadness reactivity in late life. *Social Cognitive and Affective Neuroscience*, *6*(2), 186–194.

Serrano, J. P., Fernández-Aguilar, L., Ros, L., Ricarte, J. J., Nieto, M., & Latorre, J. M. (2017). New technologies to improve mood by eliciting autobiographical memories in older adults. *Innovation in Aging*, *1*, 746.

Sharma, N., & Gedeon, T. (2013). Hybrid genetic algorithms for stress recognition in reading. In Vanneschi, L., Bush, W. S., & Giacobini, M. (Eds.), *Evolutionary computation, machine learning and data mining in bioinformatics*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 117–128.

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology*, *52*(6), 1061–1086.

Shin, H. S., Lee, C., & Lee, M. (2009). Adaptive threshold method for the peak detection of photoplethysmographic waveform. *Computers in Biology and Medicine*, *39*(12), 1145–1152.

Sokolova, M. V., & Fernández-Caballero, A. (2015). A review on the role of color and light in affective computing. *Applied Sciences*, *5*(3), 275–293.

Stephens, C. L., Christie, I. C., & Friedman, B. H. (2010). Autonomic specificity of basic emotions: Evidence from pattern classification and cluster analysis. *Biological Psychology*, *84*(3), 463–473.

Valenza, G., Lanata, A., & Scilingo, E. P. (2012). The role of nonlinear dynamics in affective valence and arousal recognition. *IEEE Transactions on Affective Computing*, *3*(2), 237–249.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

World Health Organization (2011). Ageing and life course. http://www.who.int/ageing/en/

Yuvaraj, R., & Murugappan, M. (2016). Hemispheric asymmetry non-linear analysis of eeg during emotional responses from idiopathic Parkinson's disease patients. *Cognitive Neurodynamics*, *10*(3), 225.

Zangróniz, R., Martínez-Rodrigo, A., López, M. T., Pastor, J. M., & Fernández-Caballero, A. (2018). Estimation of mental distress from photoplethysmography. *Applied Sciences*, *8*(1), 69.

Zangróniz, R., Martínez-Rodrigo, A., Pastor, J. M., López, M. T., & Fernández-Caballero, A. (2017). Electrodermal activity sensor for classification of calm/distress condition. *Sensors*, *17*(10), 2324.

AUTHOR BIOGRAPHIES

Arturo Martínez-Rodrigo received his MSc degree in Telecommunications Engineering and his PhD in Medical Care Research from the Universidad de Castilla-La Mancha, Spain, in 2010 and 2013, respectively. He was Assistant Professor with Department of Mathematics at Universidad de Castilla-La Mancha in 2010. Since 2015, he is Assistant Professor with Department of Computer Science, Universidad de Castilla-La Mancha. He has taught several subjects related to computation, statistics, programming, and communications. His research interests include statistics, signal processing applied to biomedical signals, artificial intelligence, biomedical sensors, sensor networks, and communications. He is author of more than 30 publications, including more than 20 peer-reviewed articles, 16 contributions in conference proceedings, and 12 book chapters. He has also been working for private companies related to the military area, designing and developing autopilot systems and system architectures for unmanned aerial vehicle platforms as well as communication systems in hostile environments.

Luz Fernández-Aguilar holds a degree in Psychology from the Universidad de Valencia, Spain, and a Master in Cognitive and Behavioural Neuroscience from the Universidad de Granada, Spain. She is Assistant Professor with Department of Psychology at Universidad de Castilla-La Mancha, Albacete, Spain. Her research focuses on emotion variables related to Mental Health and Gerontology. Currently, she is working on the study of neural, psychological, and physiological factors involved in the induction of positive and negative emotions in older people.

Roberto Zangróniz received his MSc degree in Physics and his MSc degree in Electronics Engineering from the Universidad de Valladolid, Spain, in 1999 and 2001, respectively. He has an extensive experience in the design and development of Low-Power Embedded Systems. His research interests include signal processing applied to biomedical signals, biomedical sensors, and wireless sensor networks. In 2001, he joined the Universidad de Castilla-La Mancha where he is Lecturer with Department of Electrical Engineering, Electronics, Automation and Communications.

José M. Latorre is Bachelor in Psychology and Doctor in Experimental Psychology. He is Full Professor with the Universidad de Castilla-La Mancha, Albacete, Spain, in the area of Basic Psychology. He has experience in Basic and Applied Experimental Psychology in the study of ageing and various transdiagnostic factors in Psychopathology, mainly autobiographical memory. Together with others, he has developed psychotherapy and cognitive training techniques for older people with cognitive impairment or depression and patients with schizophrenia. Currently, he is working on the study of neural and psychological factors involved in the induction of positive emotions in older people.

José M. Pastor received his MSc degree in Electronics and Automatic Control Engineering from the Universidad Politécnica de Madrid (UPM) in 1991. In 1992, he starts his doctoral research in Robotics in Construction Industry. In 1996, he moved to the Engineering Department at Universidad Carlos III de Madrid. He received his PhD in Robotics and Artificial Intelligence in 1997 from UPM. Since 2005, he is Associate Professor with Department of Computer Science at Universidad de Castilla-La Mancha (UCLM). He has taught several subjects related to robotics, computation, programming, and communications. His research interests include signal processing applied to biomedical signals, artificial intelligence, biomedical sensors, sensor networks, wearables, and communications. He is author of more than 40 scientific contributions. He is Director of the Research Institute of Audio-Visual Technologies and Head of the Biomedical, Electronics and Telecommunication Research Group at UCLM.

Antonio Fernández-Caballero received his MSc degree in Computer Science from the Universidad Politécnica de Madrid, Spain, and his PhD degree from Department of Artificial Intelligence of Universidad Nacional de Educación a Distancia, Spain. He is Full Professor at Universidad de Castilla-La Mancha, Albacete, Spain. He is Head of the natural and artificial Interaction Systems (n&aIS) team, belonging to the Laboratory on User Interfaces and Software Engineering (LoUISE) in the Computer Science Research Institute at Albacete (i3A) since 2001. Among his research interests are Computer Vision, Pattern Recognition, Human-Machine Interaction, Affective Computing, Multiagent Systems, and Mobile Robots. He has authored over 370 scientific contributions. He is Topic Editor-in-Chief for Vision Systems of International Journal of Advanced Robotic Systems. He is Associate Editor of Pattern Recognition Letters, Associate Editor of Frontiers in Neuroinformatics, and Specialty Chief Editor for Robot and Machine Vision of Frontiers in Robotics and AI. He is Guest Editor of several special issues of leading international journals.

How to cite this article: Martínez-Rodrigo A, Fernández-Aguilar L, Zangróniz R, Latorre JM, Pastor JM, Fernández-Caballero A. Film mood induction and emotion classification using physiological signals for health and wellness promotion in older adults living alone. *Expert Systems*. 2019;e12425. https://doi.org/10.1111/exsy.12425