IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 18, 2019, accepted June 4, 2019, date of publication June 10, 2019, date of current version June 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2922037

Examining Human-Horse Interaction by Means of Affect Recognition via Physiological Signals

TURKE ALTHOBAITI[1,2], STAMOS KATSIGIANNIS[1], (Member, IEEE), DAUNE WEST[1], AND NAEEM RAMZAN[1], (Senior Member, IEEE)
[1]School of Computing, Engineering and Physical Sciences, University of the West of Scotland, Paisley PA1 2BE, U.K.
[2]Rafha Community College, Northern Border University, Rafha, Saudi Arabia

Corresponding author: Turke Althobaiti (Turke.Althobaiti@uws.ac.uk)

ABSTRACT For some time, equine-assisted therapy (EAT), i.e., the use of horse-related activities for therapeutic reasons, has been recognised as a useful approach in the treatment of many mental health issues such as post-traumatic stress disorder (PTSD), depression, and anxiety. However, despite the interest in EAT, few scientific studies have focused on understanding the complex emotional response that horses seem to elicit in human riders and handlers. In this work, the potential use of affect recognition techniques based on physiological signals is examined for the task of assessing the interaction between humans and horses in terms of the emotional response of the humans to this interaction. Electroencephalography (EEG), electrocardiography (ECG), and electromyography (EMG) signals were captured from humans interacting with horses, and machine learning techniques were applied in order to predict the self-reported emotional states of the human subjects in terms of valence and arousal. Supervised classification experiments demonstrated the potential of this approach for affect recognition during human–horse interaction, reaching an F1-score of 78.27% for valence and 65.49% for arousal.

INDEX TERMS Affective computing, ECG, EEG, EMG, emotion recognition, equine assisted therapy (EAT), human-horse interaction, physiological signals.

I. INTRODUCTION

In the field of computing, emotion recognition falls within the area of affective computing, which focuses on the recognition and interpretation of emotions, and the processing methods used to interpret emotions in accordance with the needs of the user [1]. Emotion recognition using physiological signals has also become a subject of considerable interest in recent years in the field of human computer interaction. Physiological signals consist of signals originating from the central nervous system (CNS) and the peripheral nervous system (PNS) and have been shown to include information that can be exploited for the assessment of emotion [2]. For example, studies have shown that there is a relation between physiological signals and the Arousal and Valence dimensions of a felt emotion [2], [3], and multiple studies have examined the use of multi-modal systems for capturing physiological signals and mapping them to an emotional state [3], [4].

A field that can potentially benefit from the use of affective computing techniques is the field of animal-assisted intervention (AAI). In recent years, AAIs have been used increasingly as a complement to conventional mental health treatment [5]. However, research has not kept pace with practice, since until relatively recently, there was little scientific research evidence to support the effectiveness of AAIs. However, this is beginning to change, and there has been an increasing number of rigorously designed studies in recent years [6]. Although the importance and relevance of human-animal interaction has been recognised for some time [7], it is only recently that horses have been incorporated into mental health treatment. There are many forms of equine assisted therapy (EAT), such as therapeutic riding, hippotherapy, equine-facilitated therapy, and equine-assisted learning therapy [8].

The first recognition of the horse as an agent of healing can be encountered in early mythology, when it is reported that a physician suggested horse riding to people with untreatable conditions on the premise that it would ''raise spirits'' [9]. Similarly, reports from the eighteenth century mention that the Pope's physician recommended that he ride horses in order to help with his health problems [10]. Mayberry [11] reports that in 1870, a Scottish physician recommended that the riding of a spirited horse should be recognised as a treatment for people with depression because it ''stimulated life forces.''

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

Since the mid-twentieth century, there has been increasing recognition of horse-riding as a type of therapy for disadvantaged people and/or people with disabilities [12]. Riding horses for sport and recreation and riding as therapy differ in many respects. Therapeutic riding involves the use of equine-oriented activities to achieve any of a number of positive results, such as physical, emotional, behavioural, social, cognitive, and educational goals. There are, however, considerable differences in how various organisations carry out EAT or horse related therapeutic activities. The therapeutic notion is based on the principle that horses have been necessary and significant partners for humans since time immemorial and that the part they play in emotional and physical recovery has long been acknowledged [8].

Horses are sensitive to non-verbal communication of both the humans and animals surrounding them [8]. However, they do not have the same socio-cultural standards and restrictions that guide the way in which humans react to each other. Kendall et al. [8] suggest that this offers a secure setting in which people, in particular those with disabilities, can establish trust. To this end, equine-assisted therapy has found many applications, such as the treatment of both children and adults with PTSD [13], [14], depression, internalising and externalising behaviour [15], as well as the treatment of adults combat trauma [16]. Other studies have been conducted for specific purposes, such as treating children with autism depending on the measurement of brain activity [17], while the measurement of heart rate has also been exploited in order to evaluate the interaction between human and horses [18].

While various studies have been conducted examining EAT, such as [19], [20], and [21], relatively few scientific studies have focused on understanding the complex emotional response that horses seem to elicit in human riders and handlers. Evidence about the emotional response to human-horse interaction dates back at least to the earliest existing writing about horsemanship [22]–[24]. Nevertheless, these sources rely on empirical observations and a subjective personal experience as the basis of their conclusions. The use of analytical methods to quantitatively assess human-horse interaction has the potential to provide powerful insights on the benefits of such interaction in psychotherapy intervention. Modern computing capabilities and health sensors can provide the means for a quantitative examination of the emotional responses elicited through human-horse interaction.

In this work, the authors aim to study the potential use of affect recognition techniques, based on physiological signals, in order to assess the interaction between humans and horses in terms of the emotional response of the humans to this interaction. Electroencephalography (EEG), electrocardiography (ECG), and electromyography (EMG) signals were captured from humans while they were interacting with horses following a predefined protocol. The emotional state of each participant during each activity involving the horses was self-reported by selecting the most relevant emotions out of a list. The reported emotions were then transformed into their associated Valence and Arousal dimensions, and features extracted from the physiological signal recordings were used in order to conduct supervised classification experiments with the aim to map these signals to the associated Valence and Arousal values, and consequently to the emotional state of the participant.

The rest of this paper is organised in five sections. Section II provides an overview of the most recent works in the field of affect recognition and human-horse interaction. Section III provides a detailed description of the proposed experimental protocol, while the proposed methodology is described in Section IV. Section V presents and discusses the experimental results. Finally conclusions are drawn in Section VI.

II. BACKGROUND

The task of emotion recognition is of utmost importance in the field of affective computing. As a result, multiple research works have examined the use of physiological signals to achieve this aim. An extensive survey on emotion recognition techniques relying on various stimuli for affect elicitation is provided by Zeng et al. [25], while a more recent survey focusing on the task of continuous affect detection is provided in [26]. Various research works examined the use of features extracted from physiological signals in order to train machine learning models for distinguishing between different emotional states, in terms of Valence and Arousal.

Soleymani et al. [27] examined the use of peripheral physiological signals and eye gaze data along with Support Vector Machines (SVM) for affect recognition using film clips as stimulus. Koelstra et al. [2] evaluated the use of EEG and peripheral physiological signals along the Naive Bayes classifier for the detection of affect when subjects watched music video clips, while Arnau-González et al. [28] examined the performance of connectivity-based and channel-based EEG features on the Koelstra et al. [2] DEAP dataset. Katsigiannis and Ramzan [3] used an SVM classifier with a Radial Basis Function (RBF) kernel and EEG and ECG-based features for emotion recognition under film clip stimulus and using portable wireless devices for signal acquisition, while for music video stimulus, Abadi et al. [29] examined the performance of a wide range of physiological signal modalities, such as magnetoencephalography (MEG), electrooculography (hEOG), ECG, and EMG, as well as near-infra-red (NIR) facial videos. In a similar work using film clips as stimulus, Correa et al. [30] evaluated the use of EEG, ECG and galvanic skin response (GSR) signals, as well as facial and full body video for the task of emotion recognition.

Despite the extensive literature on affect recognition using physiological signals, to date, few studies have been carried out to explore the interaction between humans and horses using physiological signals. One early study, conducted in Japan by Hama et al. [31], examined the heart rates (HR) of both humans and horses when the humans groomed the horses for ninety seconds. The Tohoku Activation Deactivation Adjective Check List [31] was used to measure the subjective arousal levels of the humans before and after grooming the horses. Six male subjects with a positive attitude toward companion animals and six male subjects with a negative attitude towards them were chosen by their scores on the Pet Attitude Scale. These two groups were joined by a third group of six male subjects, who were members of the Doshisha University Riding Club, to take part in this experiment. The HRs of the human subjects during the first 10 seconds after beginning the grooming were significantly higher than those obtained after that period, but slowly returned to baseline levels, a tendency which was more noticeable in the negative attitude group. The HRs of the horses increased during the first 20 seconds after the human subjects of the negative attitude group began to groom them, but slowly reduced as the grooming continued. The Hama et al. study [31] concluded that grooming horses results in a reduction in tension.

Almost twenty years after the Hama et al. study [31], Chen et al. [17] conducted a considerably more sophisticated study, the purpose of which was to determine the relationship between autism spectrum disorder (ASD) and resting frontal EEG brain activity in young children during interaction with a horse. Resting frontal EEG brain activity was used based on findings that greater left frontal activity was associated with high levels of outward expressions of anxiety and anger in higher functioning children with ASD [32], [33] and that resting frontal EEG alpha asymmetry may therefore be a valid measure to define individual differences in children with ASD. The results showed that children with ASD exhibited higher left frontal dominance during the baseline condition, but right frontal dominance while grooming the horse. Chen et al. inferred that the horse conveyed its calmness to the children with ASD during grooming. Hence, Chen et al. [17] concluded that this change in attentional focus of the children with ASD may be attributed to the interaction with the horse.

In another study, Guidi et al. [18] attempted to explore the reliability of using wearable systems for monitoring physiological signals in horses and to estimate and quantify human-horse interaction during a particular experimental protocol. A preliminary estimation of human-horse interaction was carried out using the dynamic time warping (DTW) algorithm to analyse the heart rate variability (HRV) in both human and horse in a group of fourteen human subjects and one horse. The HRV was monitored by a wearable, e-textile-based system developed by the authors and whose performance was compared to an existing widely-used system in terms of movement artefact (MA) percentage. Regarding the human-horse interaction, the three classes of interaction (positive, negative, neutral) were recognised by an SVM classifier with an accuracy of almost 79%. Guidi et al. [18] concluded that it was viable to measure human-horse interaction quantitatively, and that such a measure could be very useful in many areas of application.

Lanata et al. [34] focused on the use of ECG signals to examine human-horse interaction. The signals were acquired in three phases: 1) before interaction, 2) visual-olfactory interaction, and 3) grooming, for the purpose of distinguishing the interaction activity between the subject and the horse. In that study, the use of the Nearest Mean classifier with ECG-based features reached a classification accuracy of 70.87%. Another study by the same researchers on the same topic, using the same activities, but with an SVM classifier, reached an accuracy rate of 90.95% [35].

In their previous work [36], the authors of this paper examined the use of ECG-based features in order to distinguish between negative and positive emotion during human-horse interaction. The experimental evaluation provided promising results on the efficiency of ECG signals in distinguishing between negative and positive emotions, reaching a classification accuracy of 74.21% [36]. However, it is clear that further research on human-horse interaction using various physiological signals is needed in order to validate the feasibility of using such signals in the field of human-horse interaction.

III. EXPERIMENTAL PROTOCOL

In order to study the affective responses elicited by human-horse interaction, participants were asked to interact with horses under a predefined scenario while physiological signals were recorded. After each experiment, participants were asked to provide feedback in relation to the emotions they experienced during the interaction with the horse. It must be noted that this study, including the acquisition and publication of anonymised data, was approved by the University of the West of Scotland University Ethics Committee (UWS UEC).

A. EXPERIMENTAL SETTING

The experiment took place in a small livery yard in *Ayrshire, Scotland, UK,* from late May through early August 2018. At the beginning of the experiment, participants were first given a consent form to sign, were then briefed about the experimental procedure, and were given the opportunity to ask any questions that they may have had. Instructions about horse handling, as well as about safety were also provided by the horse handler. Then, the researcher supervising the experiment proceeded to attach the physiological sensors on the participant and test signal acquisition. The experiment commenced afterwards. Two healthy stallion horses were selected by the handler (*Max*, 20 years old, and *Braga*, 8 years old) based on their friendliness and calmness towards unknown people and people with no previous experience with horses. During the experiment, each participant interacted with both horses by following the same protocol.

The experimental protocol consisted of an approximately 10 minute interaction with the horse in a small indoor sand arena. The interaction was divided into three consecutive phases based on the performed activity, namely *Looking, Grooming, and Leading*. *Looking* was the first phase of the protocol and lasted for 4 minutes. During this phase, participants were asked to sit on a chair within the sand arena while the horse was left free to move. The objective of this activity was to let both the participant and the horse to become comfortable with the presence of each other, as well as to allow the horse to familiarise itself with the setting which included the research team and equipment [18]. For the second phase, *Grooming*, participants were asked to groom the horse with a brush for 2 minutes. Prior research has shown that grooming horses leads to a decrease in the heart rate of both the human and the horse when they are both comfortable [31]. It must be noted that the horse was tied to a pole during the second activity. For the final phase, *Leading*, participants were asked to lead the horse around a predetermined path within the sand arena. The duration of this phase varied depending on each participant's experience with horses and their ability to control the horse, having a 4 minutes maximum duration. After the three activities, participants were asked to complete a questionnaire regarding their emotional state associated with each activity. A 10 minutes break was set up between each iteration of the experiment in order to allow the handler to bring the next horse to the arena and place the physiological sensors on the participant if needed.

B. DATA ACQUISITION

Three different physiological signals were captured during this study, namely electroencephalography (EEG), electrocardiography (ECG), and electromyography (EMG). Wireless portable low-cost and low-weight sensors were used for acquiring all the physiological signals, and a laptop computer was used for signal recording. The portable sensors were selected so as to accommodate the requirements of capturing outdoors, without the need for cables between the sensors and the recording device, in order to not restrict the ability of the users to move while interacting with the horses. Furthermore, the small size and the wearable nature of the used sensors ensured that the sensors would not be visible by the participants, while the laptop computer used for signal recording was set to not emit any sounds or show movement on the screen, and its size was minimal in relation to the area where the experiment was conducted, thus avoiding bias caused by the presence of equipment.

A SHIMMER™ v2 [37] wireless sensor was used in order to capture the ECG signals at a 256 Hz sampling rate, using four standard electrodes positioned on both lower ribs and clavicle. A SHIMMER™ v2 wireless sensor was also used in order to acquire the EMG signals at a 256 Hz sampling rate, using three standard electrodes positioned on the upper trapezius muscles. An Emotiv EPOC+ wireless headset [38] was used for the acquisition of 14-channel EEG signals at a sampling rate of 256 Hz. The Emotiv EPOC+ headset utilises 16 gold plated contact sensors that are fixed to flexible plastic arms and are placed against the head of the user on locations that align with the AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, M1 and M2 locations [39]. Out of the 16 contact sensors, the sensors at positions M1 and M2 were used as reference and the remaining 14 were used for capturing the EEG data. In addition to the captured ECG, EMG, and EEG signals, all the captured samples were also accompanied by timestamps with millisecond precision.

TABLE 1. List of emotions included in the self-assessment questionnaire, arranged by the level of arousal (low / high) and the level of valence (negative / positive).

	Low Arousal	High Arousal
Negative Valence	Miserable, Sad, Depressed, Gloomy, Bored, Droopy	Alarmed, Afraid, Angry, Tense, Frustrated, Annoyed, Distressed
Positive Valence	Content, Satisfied, At ease, Serene, Calm, Relaxed, Sleepy, Tired	Astonished, Excited, Aroused, Happy, Delighted, Glad, Pleased

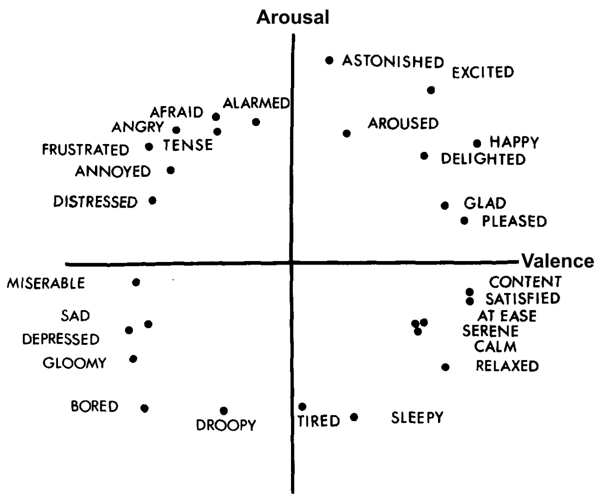


FIGURE 1. Positioning of emotions in the valence/arousal space according to Russel's *Circumplex Model of Affect* [40].

Apart from the recording of physiological signals, the whole experiment was video-recorded for reference and validation purposes and in order to accurately extract the timestamps associated with each activity.

C. PARTICIPANTS AND SELF-ASSESSMENT FEEDBACK

Twenty three healthy subjects were recruited for this study. Unfortunately, the acquired recordings for four of them had to be discarded due to missing data in some of the recordings. Out of the 19 subjects that were used for the analysis of the acquired data, 12 were male and 7 female, aged between 19 and 64 years old ($\mu_{age} = 38.05$, $\sigma_{age} = 13.14$). The participants prior experience with horses varied from no experience at all (8 participants), to prior experience but not with the specific horses (5 participants), and to prior experience with the specific horses (6 participants including the two owners).

After finishing the three activities with each horse, participants were given a questionnaire to fill. The questionnaire was divided into three parts, one for each activity, in each of which the participants had to select the emotions they felt before, during, and after the activity, out of a list of 28 emotions. Table 1 and FIGURE 1 show these 28 emotions arranged in terms of Valence and Arousal, as proposed by Russel in the *Circumplex Model of Affect* [40]. Russel's

Valence/Arousal model [41] characterises emotion in terms of dimensions that correspond to the main aspects of emotions as follows: the Valence dimension provides a measurement of the positiveness of a human's feelings, spanning from negative to positive, whereas the Arousal dimension provides a measurement of excitement, spanning from bored to excited. As a result, each perceived emotional state can be depicted on a 2-dimensional plane with Valence and Arousal at each axis respectively. It must be noted that participants were only asked to select the emotions that they felt and did not have to consider the Valence and Arousal scale, thus avoiding any bias due to misunderstanding the rating scale.

IV. DATA ANALYSIS

A. SIGNAL PRE-PROCESSING

In order to analyse the captured physiological signals in relation to each activity, the timestamps from the signals and the timestamps from the video recording of each experiment were used in order to divide each recording to segments referring to each activity and horse. As a result, the signals recorded for each participant were divided into six segments (3 activities $\times$ 2 horses). Then, in order to reduce the effects of noise and artefacts due to the participants movement during the experiment, pre-processing was applied to the ECG, EMG, and EEG signals.

ECG signals commonly suffer from artefacts such as baseline wander due to movement or respiration and from high-frequency noise such as the electromyographic noise caused by muscle activity [42]. To cope with these issues, baseline wander reduction was first applied as follows: The ECG signal was first filtered by applying a median filter with a 200 *ms* window, followed by applying a median filter on the filtered signal with a 600 *ms* window, and finally subtracting the filtered signal from the original signal [43]. After baseline wander reduction, a bandpass filter between 0.7 - 20 Hz was applied to reduce noise.

EMG signals were pre-processed as proposed in the Augsburg Biosignal Toolbox (AuBT) [44], as follows: the peaks with values within the 3% of the lowest or highest values within the EMG signal were first cut, and then, a 3rd order Butterworth FIR lowpass filter with a cutoff frequency of 0.4 Hz was applied, followed by normalising the result within the range [0, 1].

EEG signals were pre-processed by first applying a Butterworth bandpass filter between 0.4 and 65 Hz and then by using the EEGLAB toolbox [45] in order to apply the PREP EEG data pre-processing pipeline [46], which includes the following steps: First, filtering is applied in order to remove line-noise and then the EEG signal is referenced relative to an estimate of the "true" average reference. Finally, bad channels are detected and interpolated relative to the reference.

B. FEATURE EXTRACTION

After pre-processing, various features were extracted from each segment of the acquired physiological signals in order to be used for training machine learning models for the prediction of the associated emotional state. To compensate for the variation in the duration of the activities across different participants, and in order to avoid any resulting bias in the extracted features, only the last 30 *sec* of each activity were taken into consideration for feature extraction. The variation in the duration of affective stimulus is a common issue in affective computing studies and arises due to practical reasons such as film clip duration, music clip duration, participants solving cognitive tasks faster or slower, etc. The most common strategies for addressing this issue are performing any analysis in moving windows of fixed duration (e.g. [47]), or taking into consideration a fixed-length window from the end of the recording (e.g. [2], [3], [29]). In this work, we opted for the latter, similarly to [2] and [48], since the former option is more suited for real-time applications. The extracted features were the following:

1) ECG-BASED FEATURES

In the literature, features extracted from ECG signals have been shown to correlate with changes in the affective state of a person [2], [27]. The most commonly used ECG features are heart rate (HR) and heart rate variability (HRV) specific parameters in the time and frequency domain respectively. Rainville et al. [49] showed that heart rate variability may decrease with fear, sadness and happiness, while pleasantness may lead to an increase in the peak heart rate [50].

Based on these findings, HR and HRV features were computed from the acquired ECG signals. The Pan-Tompkins QRS detection algorithm [51] was used to detect QRS complexes and R-peaks within the ECG signals, and the Augsburg Biosignal Toolbox (AuBT) [44] was used in order to compute 84 statistical features from each part of the PQRST complexes. The extracted features were the maxima, minima, mean, median, standard deviation and range from the raw signal and the derivative of PQ, QS and ST complexes, the number of intervals with latency $>$ 50 *ms* from HRV, the Power Spectral Density (PSD) from HRV between the intervals [0 , 0.2], [0.2 , 0.4], [0.4 , 0.6] and [0.6 , 0.8], and the maxima, minima, mean, median, standard deviation and range from the HRV histogram. After computing the aforementioned features, the final feature vector F_{ECG} was created as their concatenation.

2) EMG-BASED FEATURES

The Augsburg Biosignal Toolbox [44] was used to extract 21 statistical features from the acquired EMG signals. The computed features consisted of the mean, median, standard deviation, minima, maxima, and the number of times per time unit that the signal reached the minima and the maxima, extracted from the raw EMG signal, the first derivative of the EMG signal, as well as its second derivative. The final feature vector F_{EMG} was created by concatenating all the computed features:

$$F_{EMG} = \begin{bmatrix} F_{EMG(t)} & F_{\frac{d}{dt}EMG(t)} & F_{\frac{d^2}{dt^2}EMG(t)} \end{bmatrix} \tag{1}$$

3) EEG-BASED FEATURES (AVERAGE PSD)

The Power Spectral Density (PSD) of different frequency bands has been shown to correlate with the affective state of humans and has been commonly applied to explain patterns in EEG signals [27], [52]. To this end, PSD features were computed from the theta (θ: 4-8 Hz), low alpha ($\tilde{\alpha}$: 8-10 Hz), alpha (α: 8-13 Hz), beta (β: 13-30 Hz) and gamma (γ: 30-64 Hz) frequency bands of each of the 14 channels of the acquired EEG signals. These features were computed using Welch's estimate of spectral power and by averaging across the components belonging to the frequency band as follows: for each channel, the FFT is computed over a Hamming window of 2 *sec* (512 samples) with 75% overlapping (384 samples) and is then averaged to produce the final PSD estimate. The logarithm of the PSD is then used as the extracted feature. The final feature vector $F_{EEG_{avg}}$ contains 70 features (5 frequency band PSDs $\times$ 14 channels) and is computed as:

$$F_{EEG_{avg}} = [\, F_{1,\theta} \quad F_{1,\tilde{\alpha}} \quad F_{1,\alpha} \quad F_{1,\beta} \quad F_{1,\gamma} \quad \ldots \quad F_{14,\theta} \quad F_{14,\tilde{\alpha}} \quad F_{14,\alpha} \quad F_{14,\beta} \quad F_{14,\gamma} \,] \tag{2}$$

where $F_{i,\theta}$, $F_{i,\tilde{\alpha}}$, $F_{i,\alpha}$, $F_{i,\beta}$, and $F_{i,\gamma}$ are the logarithms of the PSD of the i-th channel, $i = 1, 2, \ldots, 14$, for the theta, low alpha, alpha, beta, and gamma bands respectively.

4) EEG-BASED FEATURES (SPECTRAL)

PSD-based spectral EEG features were also computed as described by Monge-Álvarez et al. [53]. Five spectral features were computed for each of the theta, alpha, beta, and gamma bands of each channel of the EEG signal. The computed features were the *Spectral Bandwidth* (*SB*), the *Spectral Crest Factor* (*SCF*), the *Spectral Flatness* (*SF*), the *Spectral Roll-off* (*SRO*), and the *Ratio f50 vs f90* (*R*5090). After computing the spectral features, the feature vector is computed by concatenating the computed features for all EEG channels and all frequency bands, leading to 280 features (5 features $\times$ 4 frequency bands $\times$ 14 channels).

$$F_{Spectral} = [\, F^j_{1,\theta} \quad F^j_{1,\alpha} \quad F^j_{1,\beta} \quad F^j_{1,\gamma} \quad \ldots \quad F^j_{14,\theta} \quad F^j_{14,\alpha} \quad F^j_{14,\beta} \quad F^j_{14,\gamma} \,] \tag{3}$$

where $F^j_{i,\theta}$, $F^j_{i,\alpha}$, $F^j_{i,\beta}$, and $F^j_{i,\gamma}$ are the j-th spectral feature, $j = \{SB, SCF, SF, SRO, R5090\}$, of the i-th channel, $i = 1, 2, \ldots, 14$, for the theta, alpha, beta, and gamma bands respectively.

5) EEG-BASED FEATURES (MFCC)

Mel Frequency Cepstral Coefficients (MFCCs) provide a parametric representation of the Fourier Spectrum and have been recently applied on EEG signal analysis with promising results [54], [55]. MFCC features were computed from each channel of the EEG signal using 18 filterbanks, leading to 12 cepstral coefficients per channel, as proposed by Piciucco et al. [54]. The final feature vector $F_{EEG_{MFCC}}$ was computed as the concatenation of the cepstral coefficients of all channels and included 168 features (12 cepstral coefficients $\times$ 14 channels).

$$F_{EEG_{MFCC}} = \left[F_{EEG_{MFCC},1} \; F_{EEG_{MFCC},2} \ldots F_{EEG_{MFCC},14}\right] \tag{4}$$

where $F_{EEG_{MFCC},i}$ is the feature vector of the i-th EEG channel. Four different sets of EEG-based MFCC features were computed, depending on the frequency band of the EEG signal over which they were computed: [0.5-40 Hz], [4-40 Hz], [0.5-30 Hz], and [4-30 Hz].

6) FUSION OF FEATURES

Previous research on emotion recognition via physiological signals has shown that approaches utilising features based on multiple modalities led to increased classification accuracy compared to single-modality approaches [27], [29]. To this end, feature fusion was also examined by concatenating the feature vectors computed by each physiological signal, after normalising them to the range [0, 1] in order to compensate for the difference in their numerical range.

C. SELF-ASSESSMENT LABELS

The participants of this study provided a self-assessment of their emotional state in relation to each activity by selecting one or more emotions from a predefined list (Table 1). Russel's *Circumplex Model of Affect* [40] was then used in order to map the reported emotions to their associated Valence and Arousal values. FIGURE 1 ([40]) was used in order to extract the vectors (V, A) associated with each reported emotion, with V denoting the Valence value and A the Arousal value. Then, the values of V and A were normalised to the range $[-1, 1]$, with $V > 0$ referring to positive Valence, $V < 0$ to negative Valence, $A > 0$ to high Arousal, and $A < 0$ to low Arousal. When multiple emotions were reported by a participant for an activity, the vector (V, A) was computed as the sum of the vectors (V_i, A_i), $i = 1, 2, \ldots, N$, with i denoting the i-th reported emotion and N being the number of different emotions reported. Thresholding was then applied in order to compute the final Valence and Arousal labels associated with each activity. For $V > 0$, the Valance label was set to *Positive* and for $V < 0$ the Valance label was set to *Negative*. Similarly, the Arousal label was set to *High* for $A > 0$ and to *Low* for $A < 0$. It must be noted that no reported emotion had a V or A equal to 0, thus equality was not considered during thresholding.

V. RESULTS AND DISCUSSION

A. VALENCE AND AROUSAL RATINGS

The acquired self-assessment labels were first analysed in order to examine their distribution and consequently the class balance of the examined problem. By examining the Valance and Arousal labels of the samples in the dataset, it is evident that the dataset is moderately unbalanced for Arousal, with 70.2% of the samples associated with High Arousal (HA) and 29.8% with Low Arousal (LA), and highly unbalanced for Valence, with 87.7% of the samples associated with Positive Valence (PV) and only 12.3% with Negative Valence (NV). Furthermore, from the bar plot in FIGURE 2, it is evident

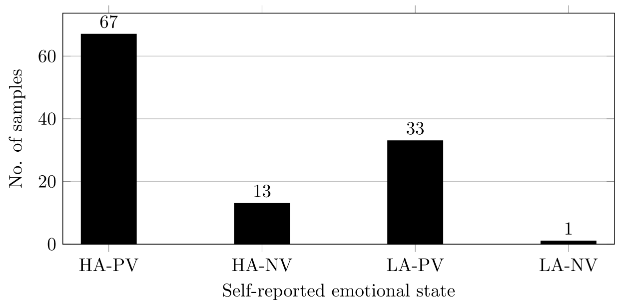


FIGURE 2. Histogram of self-reported emotional states of the participants. H:High, L:Low, P:Positive, N:Negative, A:Arousal, V:Valence.

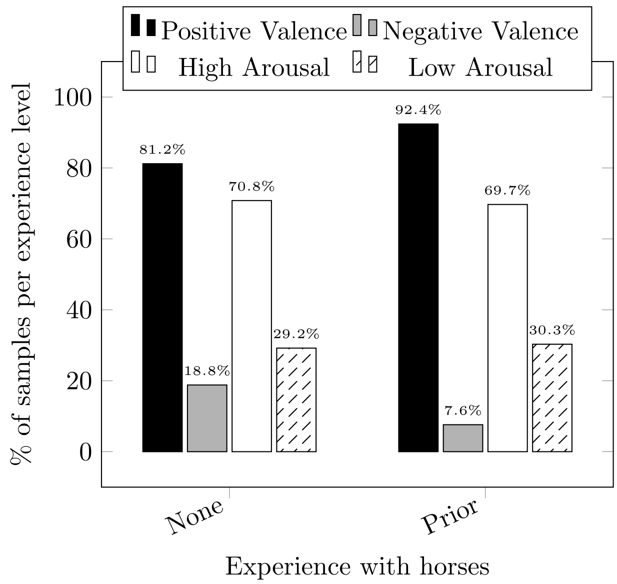


FIGURE 3. Valence and arousal ratings in relation to the participants' prior experience with horses.

that the self-reported emotional states associated with both Low Arousal and Negative Valence (LA-NV) were very rare compared to the other cases (HA-PV, LA-PV, HA-NV), with only one participant reporting a LA-NV state in one activity. This finding is consistent with the findings of previous research work [31], suggesting that the interaction between humans and horses is usually pleasant and leads to positive emotions.

FIGURE 3 shows the distribution of Valence and Arousal ratings for the participants with no prior experience with horses and for the participants with prior experience (either in general or with the specific horses used in this study). From this figure, it is evident that Arousal ratings were similar for both participant categories, with ~70% referring to High Arousal and ~30% to Low Arousal. On the contrary, the interaction with the horses seems to have elicited more positive emotions to participants with prior experience with horses, with 92.4% of the ratings of the experienced participants referring to Positive Valence, compared to 81.2% of the ratings of participants with no prior experience.

The distribution of Valence and Arousal ratings was also examined in relation to the performed activity, as shown

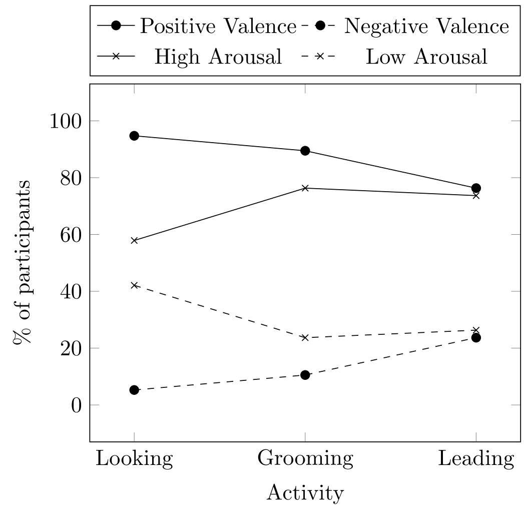


FIGURE 4. Valence and arousal ratings in relation to the performed activity.

in FIGURE 4. It is evident that for the first activity (Looking), the majority of participants (94.74%) reported Positive Valence, with their number gradually decreasing for the next two activities (89.47% for Grooming and 76.32% for Leading). Both the Looking and Grooming activities elicited mostly pleasant emotions to the participants, with pleasantness dropping for the Leading activity, where participants had to "handle" the horse, a task that can be challenging and even scary for inexperienced people. The opposite behaviour was observed for Arousal. For the Looking activity, participants were more evenly distributed, with 57.89% reporting High Arousal and their number increasing for the next two activities (76.32% for Grooming and 73.68% for Leading).

B. CLASSIFICATION EXPERIMENTS

Supervised classification experiments were conducted in order to evaluate the ability of the features extracted from the recorded physiological signals to characterise the emotional state of the participants during their interaction with the horses, in terms of Valence and Arousal. Both problems were set up as binary problems (Negative vs Positive Valence, Low vs High Arousal) and the features computed in Section IV-B were used in order to train machine learning models using various classification algorithms. The examined classification algorithms were the k-Nearest Neighbour (kNN) for $k = 1, 3, 5$, Linear Support Vector Machines (LSVM), Support Vector Machines using a Radial Basis Function kernel (SVM-RBF), Decision Trees (DT), and Linear Discriminant Analysis (LDA). A *Leave-One-Out* cross validation procedure was followed in order to avoid over-fitting the trained models and provide a fair performance evaluation. To this end, at each iteration of the cross validation, one sample is used for testing the classification model and the rest for training the model. The average performance across all the iterations of the cross validation procedure is then reported as the overall performance of the model.

TABLE 2. Classification performance (%) for valence, in terms of accuracy and F1-score, for each set of features and each classification algorithm tested.

Features	Classifier	Accuracy	F1-score
ECG	LSVM	84.21	58.09 *†‡
EMG	LDA	86.84	52.33 *
EEG-PSDavg	1NN	90.35	**78.27** *†‡
EEG-Spectral	1NN	87.72	71.50 *†‡
EEG-MFCC [4-40]	DT	86.84	68.49 *†‡
EEG-MFCC [0.5-40]	LSVM	86.84	68.49 *†‡
EEG-MFCC [4-30]	LSVM	80.70	55.21 *†‡
EEG-MFCC [0.5-30]	DT	85.09	68.27 *†‡
ALL	LSVM	88.60	65.91 *†‡
ECG/EMG/EEG-PSDavg	1NN	91.23	76.72 *†‡
ECG/EMG/EEG-Spectral	1NN	91.23	76.72 *†‡
ECG/EMG	5NN	89.47	59.67 *
EEG (ALL)	LSVM	87.72	67.40 *†‡
n/a	Random	50.00	41.71
n/a	Majority	87.72	46.73
n/a	Class Ratio	78.45	50.00

*†‡Statistically significant difference compared to random voting ($\star$), majority voting ($\dagger$), and voting according to the class ratio ($\ddagger$), $p < 0.05$.

TABLE 3. Classification performance (%) for arousal, in terms of accuracy and F1-score, for each set of features and each classification algorithm tested.

Features	Classifier	Accuracy	F1-score
ECG	1NN	56.14	44.62 *†‡
EMG	5NN	65.79	50.75 *†
EEG-PSDavg	DT	63.16	56.71 *†‡
EEG-Spectral	DT	70.18	64.38 *†‡
EEG-MFCC [4-40]	LSVM	70.18	63.75 *†‡
EEG-MFCC [0.5-40]	LDA	70.18	**65.49** *†‡
EEG-MFCC [4-30]	LDA	60.53	55.88 †‡
EEG-MFCC [0.5-30]	LDA	66.67	61.99 *†‡
ALL	LSVM	68.42	59.29 *†‡
ECG/EMG/EEG-PSDavg	5NN	67.54	54.52 *†
ECG/EMG/EEG-Spectral	3NN	66.67	55.02 *†‡
ECG/EMG	3NN	62.28	49.71 *†‡
EEG (ALL)	LSVM	69.30	61.62 *†‡
n/a	Random	50.00	47.88
n/a	Majority	70.18	41.24
n/a	Class Ratio	58.14	50.00

*†‡Statistically significant difference compared to random voting ($\star$), majority voting ($\dagger$), and voting according to the class ratio ($\ddagger$), $p < 0.05$.

Due the the moderate number of available samples and the practical difficulty to obtain more, the unbalanced dataset was used for the supervised classification experiments but the F1-score was used as a measure of classification performance instead of classification Accuracy, in order to compensate for the class imbalance. The F1-score is the harmonic mean of Precision and Recall, and constitutes a better classification performance metric in cases of uneven class distribution. Furthermore, since the F1-score is different depending on which class is considered positive, the reported F1-score was computed as the average F1-score between the two classes.

C. CLASSIFICATION RESULTS

Classification results in terms of F1-score and Accuracy are reported in Tables 2 and 3 for Valence and Arousal respectively. While Accuracy values are not taken into consideration due to the class imbalance, they are reported for reference purposes. Both single-modality features and feature fusion approaches were evaluated, and Tables 2 and 3 report the best performing setting for each of the single-modality approaches, as well as some of the best performing fusion approaches. Classification F1-score for Valence reached 78.27% using the EEG-PSDavg features and the 1-NN classifier. For Arousal, the highest classification F1-score (65.49%) was achieved using the EEG-based MFCC features for the [0.5-40 Hz] frequency band and the LDA classifier. The best performance for the feature fusion approaches was slightly lower, reaching an F1-score of 76.72% for Valence using the fusion of the ECG, EMG, and EEG-PSDavg features, as well as the fusion of the ECG, EMG, and EEG-Spectral features, and the 1-NN classifier. Furthermore, the fusion of all the EEG-based features along the Linear SVM classifier provided the highest fusion-based F1-score for Arousal (61.62%).

D. SIGNIFICANCE ANALYSIS

Taking into consideration the class imbalance within the examined dataset and in order to evaluate the statistical significance of the acquired results, the results were compared to the analytically determined expected results for voting randomly (50% probability for each class), voting according to the majority class in the training data (100% probability of the majority class), and voting according to the class ratio (the probability of each class is equal to its ratio of samples within the training set). The overall class ratios of the dataset were used for computing the results for voting according to the class ratio and consequently the computed accuracy and F1-score are slightly overestimated since the class ratio of each training fold of the leave-one-out cross validation would be needed to accurately compute the results. The analytically computed results for Valence and Arousal are reported in Tables 2 and 3 respectively.

Random voting provides an expected accuracy of 50% for both Valence and Arousal, and an F1-score of 41.71% for Valence and 47.88% for Arousal. To test for significance, an unpaired Kruskal-Wallis test was performed, comparing the predicted class labels from random voting to the predicted labels for each experimental setting depicted in Tables 2 and 3. For Valence, all settings (combination of features and classification algorithm) performed significantly better than random voting ($p \leq 1.65 \cdot 10^{-8}$). Similarly for Arousal, all settings performed significantly better than random voting ($p \leq 0.023$), apart from when the EEG-based MFCC features for the [4-30 Hz] frequency band were used ($p = 0.062$).

Class ratio based voting provided an expected F1-score of 50% for both Valence and Arousal, and an Accuracy of 78.45% for Valence and 58.14% for Arousal. An unpaired Kruskal-Wallis test, comparing the predicted class labels

TABLE 4. Classification performance for valence and arousal across the literature, in terms of accuracy (%) and F1-score (%), when features based on physiological signals are used.

Approach	Reference	Stimulus	Brain signal device	Valence		Arousal	
				Accuracy	F1-score	Accuracy	F1-score
AMIGOS	[30]	Film clips	Emotiv EPOC (EEG)	n/a	56.40	n/a	57.70
Arnau et al.	[28]	Music videos	Biosemi Active II (EEG)	69.60	69.20	67.70	66.70
DEAP	[2]	Music videos	Biosemi Active II (EEG)	62.70	60.80	62.00	58.30
DECAF	[29]	Film clips	ELEKTA Neuromag (MEG)	60.00	59.00	62.00	58.00
DREAMER	[3]	Film clips	Emotiv EPOC (EEG)	62.49	53.05	62.32	57.98
MAHNOB-HCI	[27]	Film clips	Biosemi Active II (EEG)	57.00	56.00	52.40	42.00
This work	-	Human-horse interaction	Emotiv EPOC+ (EEG)	90.35	78.27	70.18	65.49

Note: Reported results refer to the highest results achieved using only physiological signals. Accuracy for this work is provided for reference but must not be taken into consideration due to the class imbalance.

from class ratio based voting to the predicted labels for each experimental setting depicted in Tables 2 and 3, was used to test for significance. All settings performed significantly better than class ratio based voting for Valence ($p \leq 0.031$), apart from when the EMG-based features ($p = 0.31$) and the fusion of ECG and EMG-based features ($p = 0.56$) were used. For Arousal, all settings performed significantly better than class ratio based voting ($p \leq 0.022$), apart from when the EMG-based features ($p = 0.15$) and the fusion of ECG, EMG, and EEG-PSDavg based features ($p = 0.074$) were used.

A paired Wilcoxon signed-rank test was used to test significance against majority voting since the predicted class labels can be computed definitely on a one-by-one basis. Results for Valence showed that all settings performed significantly better than majority class voting ($p \leq 0.001$), apart from when the EMG-based features ($p = 0.082$) and the fusion of ECG and EMG-based features ($p = 0.158$) were used. In the case of Arousal, all settings performed significantly better than majority class voting ($p \leq 1.92 \cdot 10^{-5}$).

E. FURTHER DISCUSSION

Table 4 shows the highest classification performance for Valence and Arousal, achieved in other research works ([2], [3], [27]–[30]) using features extracted from physiological signals. It is evident that the results achieved in this work are consistent with the results from other works that employ similar approaches. It must be noted that Accuracy results for this work must not be taken into consideration when comparing the results due to the class imbalance of the examined dataset. The comparable results achieved in this work and the compared approaches provide evidence that the use of features based on physiological signals is suitable for the task of emotion recognition during human-horse interaction.

Furthermore, the results also provide evidence that the use of low-cost portable devices for emotion recognition applications is a viable alternative to expensive and non-portable medical-grade EEG, ECG, and/or EMG devices, such as the ones used in [2], [27], [28], as also evidenced by the results of [3] and [30]. Portable and wireless sensors are necessary for studying human-horse interaction due to the requirement of not restricting the users ability to move. Furthermore, the ability to monitor and detect the emotional response of people interacting with horses can potentially be beneficial to the field of equine assisted therapy by facilitating the study of the complex emotional responses that horses seem to elicit in human riders.

VI. CONCLUSION

In this work, an affect recognition approach based on physiological signals and machine learning was evaluated for the task of detecting the emotional state of people interacting with horses. EEG, ECG, and EMG signals were recorded while human subjects engaged in three different activities (Looking, Grooming, Leading) with two different horses. Portable wireless wearable devices were used for signal acquisition to avoid hindering the participants ability to move. Participants reported their emotional state in terms of distinct emotions, which were then mapped to their associated Valence and Arousal values. Time and frequency domain statistical features were extracted from the acquired signals in order to train machine learning models for the task of distinguishing between positive and negative Valence, and low and high Arousal. Supervised classification experiments using a leave-one-out cross validation procedure and various classification algorithms demonstrated the efficiency of the proposed approach. A 78.27% F1-score was reached for Valence using the ECG-PSDavg features and the 1-NN classifier, whereas a 65.49% F1-score was achieved for Arousal using the EEG-based MFCC features for the [0.5-40 Hz] frequency band and the LDA classifier. The acquired results provide evidence on the suitability of physiological signals for the task of affect recognition in the context of human-horse interaction.

Potential applications of this study could include the detection of the emotional responses of people undergoing equine-assisted therapy (EAT) or potentially general animal-assisted therapy in order to evaluate their effectiveness. The proposed methodology offers a quantitative method for assessing human-horse and potentially human-animal interaction, an approach that can be significantly beneficial to researchers studying such interactions that usually rely on empirical and subjective observations. Furthermore, the portability and the wireless characteristics of the employed approach are ideal for field studies, such as those required when examining human-animal interaction.

Future work will include the study of additional physiological signals and different feature extraction approaches for the task of emotion recognition during human-horse interaction. Furthermore, the proposed approach for the evaluation of the emotional response of humans during human-horse interaction will be evaluated within the context of equine assisted therapy in order to assess the potential benefits.

REFERENCES

[1] R. W. Picard, "Affective computing: From laughter to IEEE," *IEEE Trans. Affective Comput.*, vol. 1, no. 1, pp. 11–17, Jan. 2010.

[2] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 18–31, Jan. 2012.

[3] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 1, pp. 98–107, Jan. 2018.

[4] N. Ramzan, S. Palke, T. Cuntz, R. Gibson, and A. Amira, "Emotion recognition by physiological signals," *Electron. Imag.*, vol. 2016, no. 16, pp. 1–6, 2016.

[5] M. L. Morrison, "Health benefits of animal-assisted interventions," *Complementary Health Pract. Rev.*, vol. 12, no. 1, pp. 51–62, 2007.

[6] M. Acri, K. Hoagwood, M. Morrissey, and S. Zhang, "Equine-assisted activities and therapies: Enhancing the social worker's armamentarium," *Social Work Educ.*, vol. 35, no. 5, pp. 603–612, 2016.

[7] C. E. Amiot and B. Bastian, "Toward a psychology of human–animal relations," *Psychol. Bull.*, vol. 141, no. 1, pp. 6–47, 2015.

[8] E. Kendall, A. Maujean, C. A. Pepping, and J. J. Wright, "Hypotheses about the psychological benefits of horses," *Explore*, vol. 10, no. 2, pp. 81–87, 2014.

[9] M. R. Haskin, W. J. Erdman, J. Bream, and C. G. Mac Avoy, "Therapeutic horseback riding for the handicapped," *Arch. Phys. Med. Rehabil.*, vol. 55, no. 10, pp. 473–474, 1974.

[10] D. Woods, "Horseriding catching on as a therapy for the disabled," *Can. Med. Assoc. J.*, vol. 121, no. 5, pp. 631–650, 1979.

[11] R. P. Mayberry, "The mystique of the horse is strong medicine: Riding as therapeutic recreation," *Rehabil. Literature*, vol. 39, nos. 6–7, pp. 192–196, 1978.

[12] A. Masini, "Equine-assisted psychotherapy in clinical practice," *J. Psychosocial Nursing Mental Health Services*, vol. 48, no. 10, pp. 30–34, 2010.

[13] J. Yorke, W. Nugent, E. Strand, R. Bolen, J. New, and C. Davis, "Equine-assisted therapy and its impact on cortisol levels of children and horses: A pilot study and meta-analysis," *Early Child Develop. Care*, vol. 183, no. 7, pp. 874–894, 2013.

[14] J. L. Earles, L. L. Vernon, and J. P. Yetz, "Equine-assisted therapy for anxiety and posttraumatic stress symptoms," *J. Traumatic Stress*, vol. 28, no. 2, pp. 149–152, 2015.

[15] K. Kemp, T. Signal, H. Botros, N. Taylor, and K. Prentice, "Equine facilitated therapy with children and adolescents who have been sexually abused: A program evaluation study," *J. Child Family Stud.*, vol. 23, no. 3, pp. 558–566, 2014.

[16] K. Schroeder and D. Stroud, "Equine-facilitated group work for women survivors of interpersonal violence," *J. Spec. Group Work*, vol. 40, no. 4, pp. 365–386, 2015.

[17] C.-C. J. Chen, D. Crews, S. Mundt, and S. D. Ringenbach, "Effects of equine interaction on EEG asymmetry in children with autism spectrum disorder: A pilot study," *Int. J. Develop. Disabilities*, vol. 61, no. 1, pp. 56–59, 2015.

[18] A. Guidi, A. Lanata, P. Baragli, G. Valenza, and E. P. Scilingo, "A wearable system for the evaluation of the human-horse interaction: A preliminary study," *Electronics*, vol. 5, no. 4, p. 63, 2016.

[19] W. Benda, N. H. McGibbon, and K. L. Grant, "Improvements in muscle symmetry in children with cerebral palsy after equine-assisted therapy (hippotherapy)," *J. Alternative Complementary Med.*, vol. 9, no. 6, pp. 817–825, 2003.

[20] B. T. Klontz, A. Bivens, D. Leinart, and T. Klontz, "The effectiveness of equine-assisted experiential therapy: Results of an open clinical trial," *Soc. Animals*, vol. 15, no. 3, pp. 257–267, 2007.

[21] A. Selby and A. Smith-Osborne, "A systematic review of effectiveness of complementary and adjunct therapies and interventions involving equines," *Health Psychol.*, vol. 32, no. 4, p. 418, 2013.

[22] X. Xenophon, *The Art of Horsemanship*. New York, NY, USA: Dover, 2006.

[23] D. Duarte, *The Royal Book of Jousting, Horsemanship, and Knightly Combat: A Translation Into English of King Dom Duarte's 1438 Treatise Livro Da Ensinança de Bem Cavalgar Toda Sela (The Art of Riding in Every Saddle)*, A. F. Preto, Ed. Highland Village, TX, USA: Chivalry Bookshelf, 2010.

[24] A. de Pluvinel, *Le Maneige Royal*. Franktown, VA, USA: Xenophon Press, 2010.

[25] Z. Zeng, M. Pantic, G. I. Roisman, and T. S. Huang, "A survey of affect recognition methods: Audio, visual, and spontaneous expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 1, pp. 39–58, Jan. 2009.

[26] H. Gunes and B. Schuller, "Categorical and dimensional affect analysis in continuous input: Current trends and future directions," *Image Vis. Comput.*, vol. 31, no. 2, pp. 120–136, 2013.

[27] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 42–55, Jan. 2012.

[28] P. Arnau-González, M. Arevalillo-Herráez, and N. Ramzan, "Fusing highly dimensional energy and connectivity features to identify affective states from EEG signals," *Neurocomputing*, vol. 244, pp. 81–89, Jun. 2017.

[29] M. K. Abadi, R. Subramanian, S. M. Kia, P. Avesani, I. Patras, and N. Sebe, "DECAF: MEG-based multimodal database for decoding affective physiological responses," *IEEE Trans. Affective Comput.*, vol. 6, no. 3, pp. 209–222, Jul. 2015.

[30] J. A. M. Correa, M. K. Abadi, N. Sebe, and I. Patras, "AMIGOS: A dataset for affect, personality and mood research on individuals and groups," *IEEE Trans. Affective Comput.*, to be published. doi: 10.1109/TAFFC.2018.2884461.

[31] H. Hama, M. Yogo, and Y. Matsuyama, "Effects of stroking horses on both humans' and horses' heart rate responses," *Jpn. Psychol. Res.*, vol. 38, no. 2, pp. 66–73, 1996.

[32] S. K. Sutton, C. P. Burnette, P. C. Mundy, J. Meyer, A. Vaughan, C. Sanders, and M. Yale, "Resting cortical brain activity and social behavior in higher functioning children with autism," *J. Child Psychol. Psychiatry*, vol. 46, no. 2, pp. 211–222, 2005.

[33] C. P. Burnette, H. A. Henderson, A. P. Inge, N. E. Zahka, C. B. Schwartz, and P. C. Mundy, "Anterior EEG asymmetry and the modifier model of autism," *J. Autism Develop. Disorders*, vol. 41, no. 8, pp. 1113–1124, 2011.

[34] A. Lanata, A. Guidi, G. Valenza, P. Baragli, and E. P. Scilingo, "Quantitative heartbeat coupling measures in human-horse interaction," in *Proc. IEEE EMBC*, Orlando, FL, USA, Aug. 2016, pp. 2696–2699.

[35] A. Lanata, A. Guidi, G. Valenza, P. Baragli, and E. P. Scilingo, "The role of nonlinear coupling in human-horse interaction: A preliminary study," in *Proc. IEEE EMBC*, Seogwipo, South Korea, Jul. 2017, pp. 1320–1323.

[36] T. Althobaiti, S. Katsigiannis, D. West, M. Bronte-Stewart, and N. Ramzan, "Affect detection for human-horse interaction," in *Proc. 21st Saudi Comput. Soc. Nat. Comput. Conf. (NCC)*, Riyadh, Saudi Arabia, Apr. 2018, pp. 1–6.

[37] A. Burns, B. R. Greene, M. J. McGrath, T. J. O'Shea, B. Kuris, S. M. Ayer, F. Stroiescu, and V. Cionca, "SHIMMER—A wireless sensor platform for noninvasive biomedical research," *IEEE Sensors J.*, vol. 10, no. 9, pp. 1527–1534, Sep. 2010.

[38] N. A. Badcock, P. Mousikou, Y. Mahajan, P. de Lissa, J. Thie, and G. McArthur, "Validation of the emotiv EPOC EEG gaming system for measuring research quality auditory ERPs," *PeerJ*, vol. 1, p. e38, Feb. 2013.

[39] D. Wijayasekara and M. Manic, "Human machine interaction via brain activity monitoring," in *Proc. IEEE HSI*, Sopot, Poland, Jun. 2013, pp. 103–109.

[40] J. A. Russell, "A circumplex model of affect," *J. Personality Social Psychol.*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980.

[41] J. A. Russell and A. Mehrabian, "Evidence for a three-factor theory of emotions," *J. Res. Pers.*, vol. 11, no. 3, pp. 273–294, 1977.

[42] M. Blanco-Velasco, B. Weng, and K. E. Barner, "ECG signal denoising and baseline wander correction based on the empirical mode decomposition," *Comput. Biol. Med.*, vol. 38, no. 1, pp. 1–13, Jan. 2008.

[43] N. Kannathal, U. R. Acharya, K. P. Joseph, L. C. Min, and J. S. Suri, ''Analysis of electrocardiograms,'' in *Advances in Cardiac Signal Processing*. Berlin, Germany: Springer, 2007, ch. 2, pp. 55–82.

[44] J. Wagner, ''Augsburg biosignal toolbox (AuBT),'' Univ. Augsburg, Augsburg, Germany, 2005.

[45] A. Delorme and S. Makeig, ''EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis,'' *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[46] N. Bigdely-Shamlo, T. Mullen, C. Kothe, K. Su, and K. Robbins, ''The PREP pipeline: Standardized preprocessing for large-scale EEG analysis,'' *Frontiers Neuroinf.*, vol. 9, p. 16, Jun. 2015.

[47] O. AlZoubi, S. K. D'Mello, and R. A. Calvo, ''Detecting naturalistic expressions of nonbasic affect using physiological signals,'' *IEEE Trans. Affective Comput.*, vol. 3, no. 3, pp. 298–310, Jul. 2012.

[48] E.-H. Jang, B.-J. Park, S.-H. Kim, M.-A. Chung, M.-S. Park, and J.-H. Sohn, ''Classification of human emotions from physiological signals using machine learning algorithms,'' in *Proc. ACHI*, Nice, France, 2013, pp. 395–400.

[49] P. Rainville, A. Bechara, N. Naqvi, and A. R. Damasio, ''Basic emotions are associated with distinct patterns of cardiorespiratory activity,'' *Int. J. Psychophysiol.*, vol. 61, no. 1, pp. 5–18, Jul. 2006.

[50] P. J. Lang, M. K. Greenwald, M. M. Bradley, and A. O. Hamm, ''Looking at pictures: Affective, facial, visceral, and behavioral reactions,'' *Psychophysiology*, vol. 30, no. 3, pp. 261–273, 1993.

[51] J. Pan and W. J. Tompkins, ''A real-time QRS detection algorithm,'' *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.

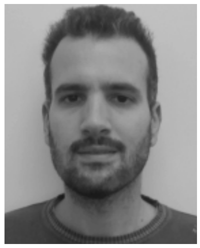
[52] R. J. Davidson, ''Affective neuroscience and psychophysiology: Toward a synthesis,'' *Psychophysiology*, vol. 40, no. 5, pp. 655–665, 2003.

[53] J. Monge-Alvarez, C. Hoyos-Barceló, L. M. S. José-Revuelta, and P. Casaseca-de-la-Higuera, ''A machine hearing system for robust cough detection based on a high-level representation of band-specific audio features,'' *IEEE Trans. Biomed. Eng.*, to be published. doi: 10.1109/TBME.2018.2888998.

[54] E. Piciucco, E. Maiorana, O. Falzon, K. Camilleri, and P. Campisi, ''Steady-state visual evoked potentials for EEG-based biometric identification,'' in *Proc. BIOSIG*. Darmstadt, Germany: Gesellschaft Informatik, Sep. 2017, pp. 1–5.

[55] P. Nguyen, D. Tran, X. Huang, and D. Sharma, ''A proposed feature extraction method for EEG-based person identification,'' in *Proc. Int. Conf. Artif. Intell. (ICAI)*, 2012, pp. 1–6.

TURKE ALTHOBAITI received the B.Sc. degree in computer science from Taif University, Saudi Arabia, in 2009, the M.Sc. degree in computer science from Ball State University, USA, in 2014, and the Ph.D. degree in computer science from the University of the West of Scotland, U.K. He is a Lecturer with Northern Borders University, Saudi Arabia. His research interests include affective computing and machine learning.



STAMOS KATSIGIANNIS (M'19) received the B.Sc. (Hons.) degree in informatics and telecommunications from the National and Kapodistrian University of Athens, Greece, in 2009, the M.Sc. degree in computer science from the Athens University of Economics and Business, Greece, in 2011, and the Ph.D. degree in computer science (biomedical image and general purpose video processing) from the National and Kapodistrian University of Athens, Greece, in 2016. He is currently a Postdoctoral Research Fellow with the School of Computing, Engineering, and Physical Sciences, University of the West of Scotland, U.K. He has participated in six national and international research projects, and has authored and coauthored over 30 research publications, including peer-reviewed journals, book chapters, and conference proceedings. His research interests include affective computing, image analysis, machine learning, video coding, image and video quality, and GPU computing.

DAUNE WEST received the B.A. (Hons.) degree in classical studies and the M.A. degree in classics from the University College of Wales, Aberystwyth, and the PgDIS and Ph.D. degrees in information systems from Portsmouth University. She is a Senior Lecturer with the Information Systems Engineering Group, School of Computing, Engineering, and Physical Sciences, University of the West of Scotland. Her research interests include systems concepts to support the development of computer-based information systems that ''fit'' user requirements and in the use of action research to undertake information systems research and consultancy. She has also published articles concerning the application of systems ideas to help explore human–horse interaction, particularly in relation to classical dressage.

NAEEM RAMZAN (S'04–M'08–SM'13) received the M.Sc. degree in telecommunications from the University of Brest, France, in 2004, and the Ph.D. degree in electronics engineering from Queen Mary University of London, London, U.K., in 2008. Currently, he is a Full Professor with the School of Computing, Engineering, and Physical Sciences, University of the West of Scotland, U.K. He has authored or coauthored over 110 research publications, including journals, book chapters, and standardisation contributions. He coedited a book titled ''Social Media Retrieval'' (Springer, 2013). He is a Fellow of the Higher Education Academy. He served as a Guest Editor for a number of special issues in technical journals. He has organised and co-chaired three ACM Multimedia Workshops, and served as the Session Chair/Co-Chair for a number of conferences. He is the Co-Chair of the Ultra HD Group of the Video Quality Experts Group (VQEG) and the Co-Editor-in-Chief of VQEG E-Letter. He has participated in more than 20 projects funded by European and U.K. research councils.