Biomedical Signal Processing and Control 75 (2022) 103580

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

A new data augmentation convolutional neural network for human emotion recognition based on ECG signals

Sihem Nita [a], Salim Bitam [b], Matthieu Heidet [c], Abdelhamid Mellouk [d,*]

[a] *Department of Computer Science, LESIA Laboratory, University of Biskra, Po. Box 145, R.P. 07000 Biskra, Algeria*
[b] *Department of Computer Science, University of Biskra, Po. Box 145, R.P. 07000 Biskra, Algeria*
[c] *SAMU 94 et Urgences, GHU Henri Mondor, AP-HP, Créteil, France*
[d] *University of Paris-Est Creteil, LISSI-TincNET(CIR), F-94400, Vitry-sur-Seine, France*

A R T I C L E I N F O

Keywords:
Human emotion detection
Electrocardiogram (ECG)
Heart Rate Variability (HRV)
Data augmentation
Convolutional Neural Network (CNN)

A B S T R A C T

Nowadays, human emotion recognition based on electrocardiogram (ECG) signal is considered as a hot topic applied in many sensitive domains such as healthcare, social security, and transportation systems. In the literature, various machine learning algorithms were proposed to this purpose however, the recognition accuracy of these techniques is hampered by the hardness of acquiring huge and balanced number of ECG dataset samples, which is considered as a major challenge in this topic. Therefore, we propose in this paper, a new data augmentation convolutional neural network (CNN) for human emotion recognition based on ECG signal. Specifically, we suggest to enrich the ECG dataset by a significant number of representative ECG samples, generated according to randomize, concatenate and resample realistic ECG episodes process. Hence, a new seven-layer CNN classifier is suggested, consisting of seven layers to detect human emotions in terms of valence, arousal, and dominance levels. Experiments that have been carried out using our proposal for Data Augmentation Convolutional Neural Network strategy on benchmark DREAMER database resulted in an accuracy rate of 95.16% to detect valence, 85.56% for arousal and 77.54% for dominance.

1. Introduction

The automatic digitalization and recognition of human emotions are considered as a novel and fast-growing domain of research, which makes a combination between knowledge in the fields of psychophysiology, biomedical engineering, artificial intelligence and computer science. Therefore, many researchers have revealed two ways to recognize human emotions; the former is from facial and/or voice expression [1–3], where the latter is from physiological signals analysis like Electrocardiogram (ECG) [4–6]. Human emotion detecting from voice or facial expression is very difficult, subjective and, unreliable because the human can mask it humans[7]. Moreover, the true effective states of a human being are not always true in the case when only facial expressions and human voices are considered however, the physiological signals generated by the Autonomic Nervous System (ANS) cannot hide it, as it reflects the real emotional state of the person [7]. Several studies like [8–10] have successfully used physiological signal analysis to identify human emotion. For example, in [11], the authors used the electromyography (EMG), ECG and galvanic skin response (GSR) to detect stress when driving cars. Moreover, the researchers proved that the ECG is a reliable and effective source of information for human emotion recognition systems [12,13] and has considerable potential for recognizing, and predicting human emotions such as anger, joy, trust, sadness, anticipation and surprise [5]. More specifically, to detect these emotions, the Heart Rate Variability (HRV) values, extracted from ECG are required. In fact, HRV analysis is defined as a simple noninvasive and effective metric, reflecting the activity of sympathetic and parasympathetic components of the ANS on the sinoatrial node (known also as a sinus node -SA-), located in the wall of the right atrium of the heart. Therefore, HRV helps to differentiate among multiple emotions such as neutrality, happiness, disgust, fear, sadness and anger [12,14].

To deal with this issue, several machine learning-based techniques were suggested to establish emotion recognition model. These techniques have been applied on a limited size of datasets. This limited size of dataset is due to the high cost of these sensitive data collection. However, the availability of a rich dataset has a big impact on the performance of machine learning. Furthermore, applying deep learning methods to recognize human emotions from ECG signals is still in its

* Corresponding author.
E-mail addresses: sihem.nita@univ-biskra.dz (S. Nita), s.bitam@univ-biskra.dz (S. Bitam), matthieu.heidet@aphp.fr (M. Heidet), mellouk@u-pec.fr (A. Mellouk).

https://doi.org/10.1016/j.bspc.2022.103580
Received 11 August 2021; Received in revised form 26 January 2022; Accepted 12 February 2022
Available online 15 March 2022

infancy. Because of the cost and limitation related to data collection, the labeled ECG samples used in previous works are significantly insufficient to train the deep learning models.

Therefore, we propose in this paper a new data augmentation convolutional neural network for human emotion recognition based on HRV, extracted from ECG signal. Notice that the proposed method increases and diversifies the considered samples and ensures a balanced number of samples in each ECG category or class. It worth noting that the existing studies regarding data augmentation for ECG signal were dedicated to detect heart disease like atrial fibrillation (AF) and not for emotion detection.

In this study, we are interested in one of the most frequently machine learning issues which is the imbalance and scarcity of the quantity of the training data [15]. Consequently, we suggest to enrich the ECG dataset by a significant number of representative ECG samples, generated according to randomize, concatenate and resample realistic ECG episodes process. Indeed, this data augmentation does not replace original data, the process is about increasing the amount of training data by creating new data from existing ones. The obtained dataset is then considered to train a new introduced seven-layer CNN classifier to detect human emotions in terms of valence, arousal, and dominance levels. The performance of this proposal is evaluated over various simulations based DREAMER dataset, based on both original and synthesized data. The obtained results were compared against those reached using different classifiers namely, Neural network (NN) and SVM.

The rest of this paper is organized as follows. In Section 2, the basic concepts and motivation of this work are introduced. Section 3 presents related works in human emotion recognition from the ECG signal and the use of data augmentation in affective computing. In Section 4, we illustrate our proposed method in detail. After that, Section 5 suggests a series of experiments to validate the proposed method on the DREAMER dataset [16]. The results are then presented and discussed in Section 6. Finally, this article is concluded with some future research directions.

2. Basic concepts and motivation

2.1. Human emotion

Human emotion is not just a psychological activity, which is very much related to physiological senses like vision, hearing, and others. So, human emotion is defined as a complex behavioral phenomenon that involves different levels of neural and chemical integration [17]. This complex state of feeling could be the cause of many physical and psychological changes that affect thought and behavior. The study presented by Russell [18] is the first one that described the human emotions in two-dimensional model of valence and arousal. Recently, another study presented by Warriner et al. [19], which added the third dimension dominance to form the Valence-Arousal-Dominance space (VAD).

So, Three main affective qualities were considered by the psychologists to describe the human emotions which are: valence, arousal, and dominance [20]. Fig. 1 illustrates a 3 dimensional emotion model, showing Valence, Arousal and Dominance (VAD in short). These emotions are explained below.

- **Valence:** represents the fear or happiness; it is the positivity or negativity of an emotion.
- **Arousal:** represents the intensity of emotion elicited by a stimulus. Precisely, a high arousal is an anger status and a low arousal represent a sadness status.
- **Dominance:** constitutes the state and level of control do by the stimulus, we distinguish two states such as dominant (i.e. with control) or submissive (i.e. without control).

As shown in Fig. 1, being Emotions are categorized on a 3D plan, the x-axis represents the valence, y-axis is for arousal and z-axis illustrates the dominance. There are 15 emotions grouped into 5 clusters: C1 represents the happy group with the emotions of Happy, Joy, Fun, Exciting, then, C2 represents is for the love group comprising: Love, Cheerful, Lovely, C3 shows the Sentimental Group which includes Depressing, Sentimental, Mellows, the C4 group is the sad group, with the emotions: Sad, Melancholy, and Terrible. C5 is the hate group with the emotion of Shock and Hate. Therefore, a human emotion is formed a combination of the three dimensions valence, arousal and dominance. In the happy group (Joy, Exciting, Happy and Fun), the valence and arousal are relatively high. However, in the sad group composed of Sad, Depressing and Melancholy the valence in low. By this way, positive and negative emotions could be expressed by numerical values [21]. For example, when we have a perpendicular plan at valence 6.5, all the emotions on the right side are called positive emotions, however, on the other side, the emotions are considered as a negative emotions. We can also see the values of an emotional state of fun, which are 6.85, 5.85, and 6 for valence, arousal and dominance, respectively. According to [21], the range of the valence, arousal and dominance are fixed as following:

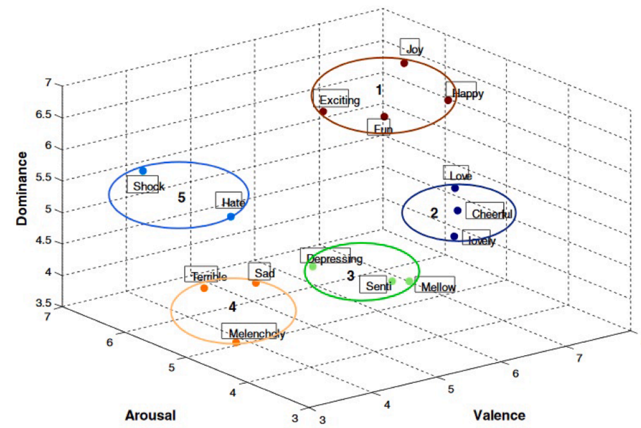


Fig. 1. VAD Emotions representation in 3D space [21].

- **Valence value range:** (Low: 1–4.5), (Medium: 4.5 – 5.5), and (High: 5.5 – 9)
- **Arousal value range:** (Low: 1–4.5), (Medium: 4.5 – 5.5), and (High: (5.5–9)
- **Dominance value range:** (Low: 1–4.5), (Medium: 4.5 – 5.5), and (High: 5.5 – 9)

Human emotions can be the reason of many physiological changes through various involuntary neurological responses such as: respiration [22,23], skin electricity [24,25], temperature [26], muscular activity, [27] and cardiologic activity. [28,30]. Various studies proved that ECG can be efficiently used for detecting human emotion [29,31]. In the next subsection, we explain how ECG can help to detect human emotions.

2.2. ECG and emotion recognition

In cardiology, heart rate (HR) is defined as the number of beats (or systolic contractions) per minute, the ECG records this cardiac electric activity responsible for the myocardial contraction, by the number of ventricular electric QRS waves see Fig. 2). HR is measured by counting the number of R waves registered in a minute. The time interval between two electrical R waves is called the R-R interval, and relates to the clinical interbeat interval. The physiological R-R interval is not constant, and varies depending on several factors, such as breathing, hormonal stimulation, or emotion (e.g. stress).

More specifically, the researchers mentioned that there are innervations of the ANS within the heart four chambers (i.e. two atria and two ventricles). These innervations, play a major role on the cardiac output. Indeed, they have an effect on both the physiological pacemaker

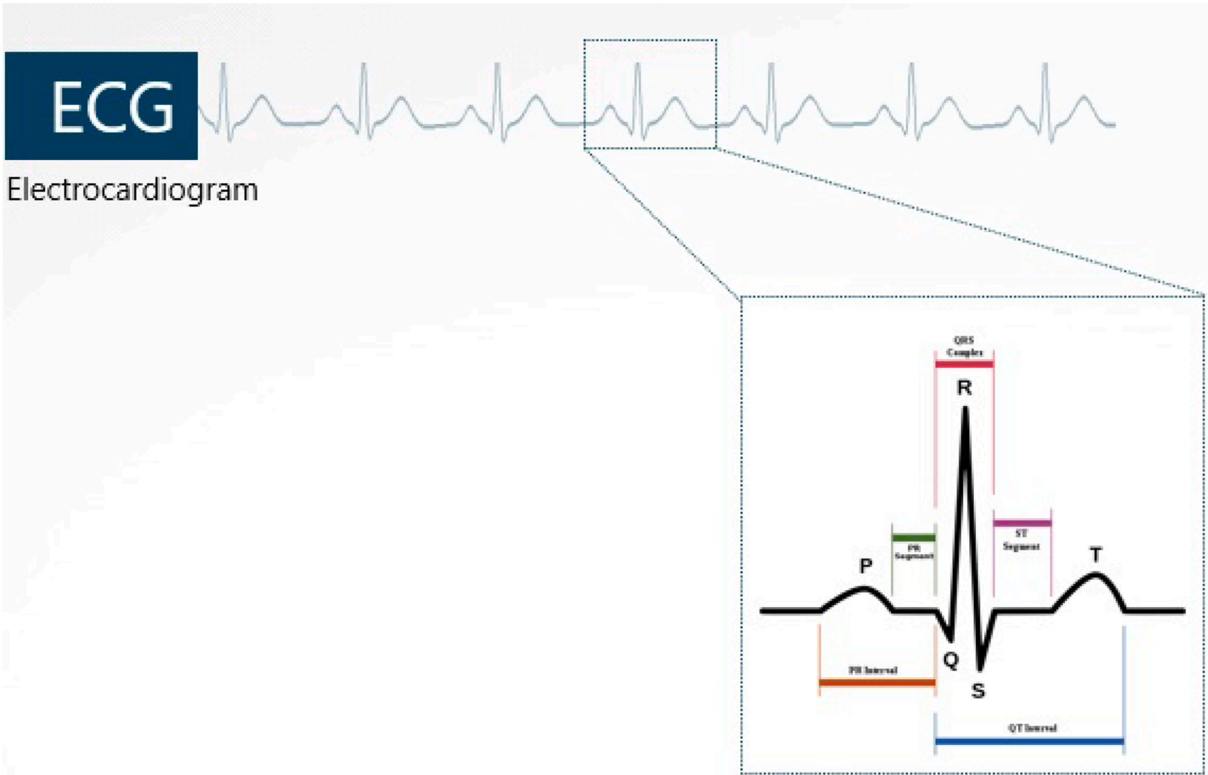


Fig. 2. ECG signal.

(sinus node), which controls the HR (chronotropic effect), and on the conduction of the electric signal running from this node through the rest of the heart (dromotropic effect). The sympathetic nervous system (SNS) has positive chronotropic and dromotropic effects (increases HR and speed of electric signal conduction), whereas the parasympathetic nervous system (PNS) has negative effect on both functions (slows pace and delays electric conduction) [32]. According to the intensity of a specific emotion, the sympathetic system is stimulated to prepare the body against a strong activity (fight-or-flight response). The parasympathetic system dominates in calm, resting activities (rest-and-digest response). Moreover, the HRV is very linked to the ANS and it is responsible for keeping the balance between the two systems: parasympathetic branches which are defined as the rest and digest response and sympathetic branches which are defined as the “fight and flight” responses [33].

3. Related work

This section reviews previous works related presented in the literature, to detect human emotion using ECG and EEG signals. Furthermore, the different machine learning techniques conceived to augment datasets, which are gathered from ECG are also presented and discussed.

3.1. Emotion detection methods

Up to now, several techniques have been conceived to recognize emotion in the basis of ECG signal and EEG signal. This section presents and discusses these works.

3.1.1. ECG-based emotion detection methods

In the past decades, ECG-based emotion recognition was considered as one of the most important branches of emotion recognition [5,34,39].

In order to identify emotions in response to music, Kim et al. in [35] used different physiological signals namely, ECG, EMG, respiration, and skin conductivity. 110 features are calculated from various analysis domains, including HRV/breathing rate variability (BRV), geometric analysis, entropy, multiscale entropy, time/frequency, sub-band spectra, etc. These calculated features have been used to detect the best emotion-related features then to relate them to emotional states using the backward feature selection method. Furthermore, a novel scheme of emotion-specific multilevel dichotomous classification (EMDC) was developed to ameliorate the accuracy of four musical emotions in terms of arousal dimension such as positive/low arousal, positive/high arousal, negative/low arousal and negative/high arousal. As a classifier, the authors used an extended linear discriminant analysis (pLDA). After a set of experiments, the authors achieved a recognition accuracy of 70% for subject-independent classification against 95% for subject-dependent classification. Nardelli et al. proposed a novel approach to identify emotions automatically, these emotions being evoked by emotional sounds. The HRV features extracted from ECG signals were used as input for the automatic emotion detection system. In this study, the emotions are expressed on valence with two classes and on arousal with four classes. The results obtained using the quadratic discriminant classifier for arousal and valence achieved a recognition accuracy rates of 84.26% and 84.72%, respectively [12].

Katsigiannis and Ramzan in [4] presented a multi-modal database called DREAMER, this latter consists of various ECG signals to recognize emotions that are evoked by audio-visual stimuli. The authors considered HR and HRV as features to detect the emotion in terms of valence, arousal and dominance using support vector machine (SVM) classifier, where the classification accuracy achieved are 62.37%, 62.37% and 61.57%, respectively.

In [40], Correa et al. presented a new dataset named A dataset for Multimodal research of affect, personality traits and mood on Individuals and GrOupS (AMIGOS). AMIGOS is conceived to detect human emotions using neurophysiological signals. This dataset contains different multimodal records of the participants as well as their reaction emotional videos. The first set of videos is classified into one to four quadrants of the valence-arousal (VA) space such as: LVLA, LVHA, HVLA and HVHA (where: L/H: Low/high and V/A: Valence/Arousal).

However, the second set of videos contains eight video extracts from movies according to their score in Internet Movie Database (IMDb) Top Rated Movies 3 list. Those selected videos from movies are affective multimedia content and do not demand a prior knowledge of the participants to be understood. In this case, the authors used short and long videos to generate different types of emotions, this is in two social situations: one for individual viewing and the other for viewing with groups of viewers. Using the wearable sensors, they collected participants' physiological signals such as: EEG, ECG and Galvanic Skin Response (GSR). The experimental evaluation showed that there are important correlations between the internal and external effect of valence and arousal which is good to predict the affective state of participants.

Sarkar and Etemad in [5,41] applied an existing self-supervised deep multi-task learning framework on ECG recordings for emotion detection. In this research activity, the authors used four public datasets namely, SWELL, WESAD, DREAMER and AMIGOS. Compared to a fully-supervised method, the results obtained showed that the proposed method is able to improve classification performance. Specifically, arousal and valence detection is performed on DREAMER dataset, achieving an accuracy of 77.1% and 74.9%, respectively, where the accuracy obtained by WESAD is 95.0% for 4 affective states: amused, meditated, stressed and neutral. Also, AMIGOS has given an accuracy of 78.3% and 79.6% for valence and arousal, respectively. Finally, with SWELL, the achieved accuracy is of 92.6%, 93.8% and 90.2% for arousal, valence, and stress.

L.Granados et al. in [42] used a dataset of physiological signals (ECG, GSR and galvanic skin response) from AMIGOS dataset. They suggested a Convolutional Neural Network (CNN) as an automatic feature extractor of GSR and ECG to detect valence and arousal. The experimental results of this research activity showed a better precision to classify different emotional states, compared with the results obtained by [40], when using classic algorithms of machine learning: Gaussian Naïve Bayes and SVM. In [30], Nita et al. used an ERF classification approach to detect and diagnose driver's stress level while driving on the road. This approach was tested on MIT-BIH physioNet dataset [43]. The results proved that the ERF is more efficient than SVM in terms of recognition accuracy.

Despite physiological signal analysis being considered as an effective method to recognize the emotion of humans using HRV of ECG signal, the proposed techniques applied a short dataset like AMIGOS, DREAMER, WESAD and SWELL in terms of size of ECG samples and their diversity. This limitation is due to two main reasons; the first one concerns the limited number of participants for measurement due to some difficult conditions (e.g. persons in the cars), and the second reason is that the only person authorized to interpret and annotate each sample is the cardiologist; it is a very hard task.

3.1.2. EEG-based emotion detection methods

As a physiological signal, the EEG can provide important and complex information about a person's emotional state. In the past decades, EEG-based emotion recognition has received great interest from researchers.

In order to detect autism spectrum disorder (ASD) in children, Aslam et al.[44] developed an emotion recognition processor, based on an eight-channel EEG signal. This research activity combines a patient special SVM classifier with a hardware-efficient feature extraction engine realized to discriminate the emotions in real-time. The accuracy results in valence and arousal were 63% and 60%, respectively. Aslam et al. [45], proposed a processor for Chronic neurological disorders (CND's) to detects human emotions using eight EEG channels. Using linear SVM (LSVM) as a classifier, they achieved an accuracy of 70.71% on SEED dataset for valence, in addition, the classification accuracy achieved with DEAP dataset is about 72.96% and 73.14% for valence and arousal, respectively. To solve the problem of limited EEG data in order to use deep learning methods to identify emotions from EEG signals, the authors of [46] applied a simple data augmentation method on MAHNOB-HCI dataset, aiming at generating more EEG training samples. The obtained results showed the effectiveness of the data augmentation method to improve the performance of deep models.

Table 1 summaries recent ECG and EEG-based emotion recognition studies proposed in the literature as well as various comparisons classification criteria.

3.1.3. Data augmentation methods for ECG

As mentioned above, ECG-based emotion recognition has received great interest from researchers especially by machine learning community. However, the success of the ECG analysis based on machine learning approaches depends mainly on a rich annotated dataset. Additionally, generating an annotated ECG dataset with high quality remains a major challenge. In fact, the model trained with small datasets does not generalize well data from the validation and test sets then, the study results will not be precise. Consequently, this type of models suffers from the problem of overfitting occurred when a good fit is achieved on the training data, while the model does not generalize well on new and unseen data. Data augmentation is one of the best solutions to reduce overfitting on models, it is based on increasing the amount of training data. In the literature, there are several data augmentation methods proposed to extend ECG dataset [38]. For instance, Cao et al. in [36] developed a new data augmentation method to improve deep neural networks (DNN), conceived to detect atrial fibrillation (AF) from ECG recordings. The principle of this method is to concatenate the original ECG episode to the duplicated one, this is based on some characteristic points. After that, this concatenated signal is resampled at random. This algorithm makes a balance in the number of samples between the different classes of the dataset and also increases the variety of the dataset. The results of this work ameliorate the performance of DNN for AF detection. Based on generative adversarial networks (GANs), Haradal et al. [37] proposed the use of a synthetic generation method for time series as well as a related application to increase data for biosignal classification. In order to generate data of time-series, the authors developed every neural network in the GANs using Long short Short-term Term Memories (LSTM) units this is done for its hidden layers based on a Recurrent Neural Network (RNN). The experiment results showed the capability of this method to generate synthetic biosignals using the EEG and ECG datasets and to analyze with better precision the studied system.

The authors of [47] used the permutation method combined with Window Slicing (WS) for the first time as a data augmentation method to monitor Parkinson's disease using wearable sensor data. This proposed method and CNN are applied in order to determine the motor state of Parkinson's Disease patients. The results obtained reached an accuracy of 86.88%. Also, Nonaka al. [48] applied a suitable method of data augmentation in a DNN model in order to classify atrial fibrillation. Using ECG augmentation with a single lead ECG data, the results showed that the proposed method improve classification of atrial fibrillation with an accuracy of 84.27%.

Although the considerable success given by data augmentation methods cited above and applied for various research areas like medical image analysis tasks, this scheme was not applied yet to human emotion recognition, known as a critical domain.

Therefore, we propose in this work a new data augmentation convolutional neural network (CNN) for human emotion recognition based on ECG signal, where the data obtained is further used as an input for the new seven-layer CNN classifier to classify the different types of human emotions.

4. The proposed ECG data augmentation for human emotion recognition using seven-layer CNN model

To recognize human emotion and to cope with the issue of limited size and imbalanced samples used to ML approaches, we detail in this

Table 1
Recent Emotion Recognition Studies based on ECG and EEG signals.

Ref	Recorded signals	Method	Emotions	Detection precision rate	Detection confusion	Amount of used data
			ECG signal			
[5]	SWELLAMIGOS AMIGOS	Self-supervisedapproach approach	Arousal Valence	High	Medium	Low
[4]	ECG signalsfrom 23 subjectsDREAMER	SVM	Valence	62.37%	Low	Low
	from 23 subjects DERAMER		Arousal	62.37%		
			Dominance	61.57%		
[39]	ECG signalsfrom 27 subjectsAMIGOS	Least squares SVMLS-SVM	Positive/negative valence	82.78%	Medium	Low
	from 27 subjects AMIGOS	LS-SVM	high/low arousal	72.91%		
			4 types of emotions	61.52%		
[40]	ECG signalsfrom 40 subjectsAMIGOS	Gaussian Naïve Bayes	Valence	54.5%	Medium	Low
	from 40 subjects AMIGOS	SVM	Arousal	55.1%		
			EEG signal			
Ref	**Recorded signals**	**Method**	**Emotions**	**Detection precision rate**	**Detection confusion**	**Amount of used data**
[49]	DEAP database	Spectral and time features, multiple-fusion-layer based ensemble classifier of stacked auto-encoder (MESAE)	Arousal Valence	77.19% 76.17%	Medium	Low
[50]	EEG signals from 57 subjects	K-nearest Neighbour (KNN), Probabilistic Neural Network (PNN)	Sad, disgust, fear, anger, happy and surprise	82.32% only sad emotion is highly perceived	Medium	High
[51]	EEG signals from 12 subjects	Linear Discrimination Analysis (LDA)	Positive and negative	64.73 %	Medium	Low
[52]	EEG signals from 15 subjects SEED dataset	Regularized graph neural network	Neutral, sad, fear, and happy	73.84 %	Medium	High

section our proposal, presented in Fig. 3. We start with ECG signal acquisition and preprocessing, then the augmented data are generated to increase data samples using our novel approach. In order to extract heart rate variabiliy, we apply a HRV analysis tool named HRV Analysis Software (HRVAS). HRVAS extracts various HRV features such as time-domain features (e.g. RMSSD: Root Mean Square of the Successive

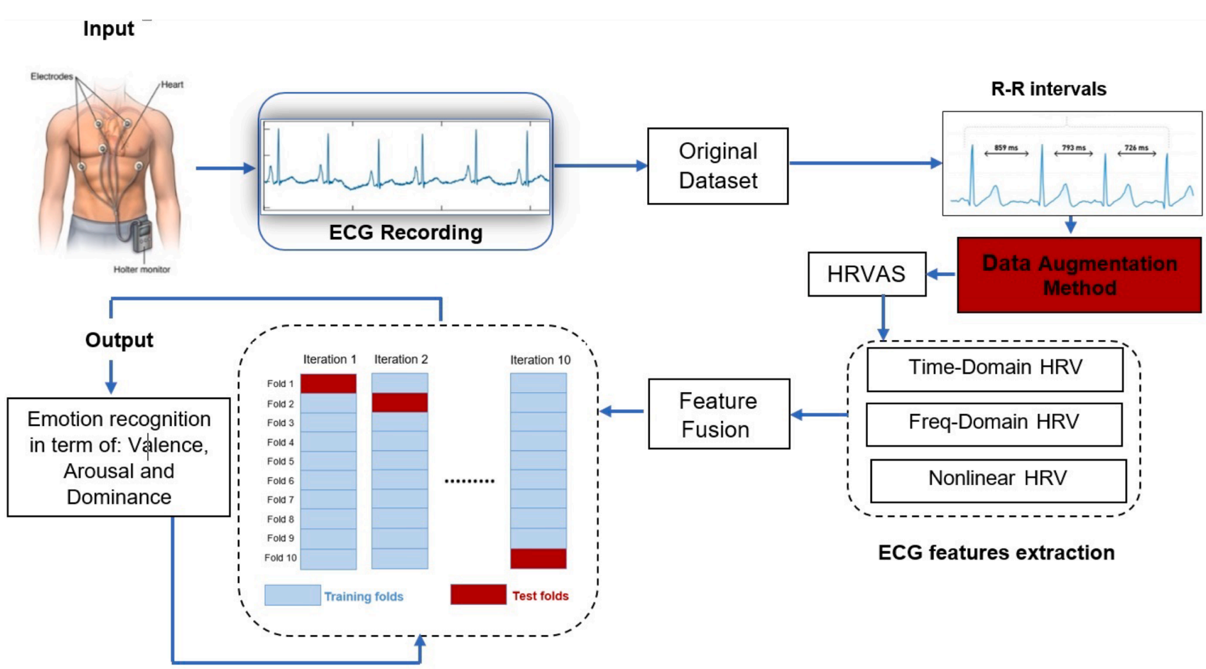


Fig. 3. Overall block scheme of the proposed method for ECG-based emotion recognition system.

Differences and SDNN: standard deviation of normal to normal R-R intervals and), frequency-domain features (e.g. LF: Low Frequency power, HF: High Frequency power and LF/HF) and nonlinear domain features (e.g. Sample entropy (SampEn)). For more details about how these features are calculated, please refer to [54]. According to these extracted features obtained from an ECG signal, the classification is performed by the suggested CNN to recognize the human emotion class such as valence, arousal, or dominance. In the next subsections, we present a detailed description of each step.

4.1. Data acquisition and preprocessing of ECG signal

In data analysis science, the data used for learning are often benchmark data, however, in this healthcare domain, this kind of datasets is limited in records number (i.e. samples quantity) and in diversity (i.e. different types of people like young or old people, men or females, etc.). The data augmentation may shed to this issue. In order to enrich the dataset by introducing unobserved and various samples, we propose to start by the use of a publicly available ECG dataset like DREAMER [4] expressing humans with valence, arousal, or dominance emotions. This dataset is considered as a small ECG datasets where the total size of the ECG for this dataset is 414 (for 23 persons exposed to 18 trials and each example of ECG contain 2 channels) [53]. More specifically, the ECG DREAMER dataset is a multi-modal database for analyzing emotions resulting from a set of audio-visual stimuli which are in the form of movie clips. Furthermore, this dataset includes ECG records collected only from 23 participants, where each participant watches 18 videos chosen to evoke nine special emotions namely calmness happiness, excitement, fear, surprise, sadness, disgust and anger. Therefore, the participant's rated his emotional response in terms of valence, arousal and dominance on a scale from 1–5.

We remind that emotions are recognized in the basis of HRV, which is in its turn detected using ECG R-peaks (i.e. R-R intervals) [56]. It is worth noting that R-R interval corresponds to the interval between two successive R peaks in an ECG. To do so, we used R-wave detection algorithm from the input ECG signals to extract R-R intervals series (see Fig. 4) [57].

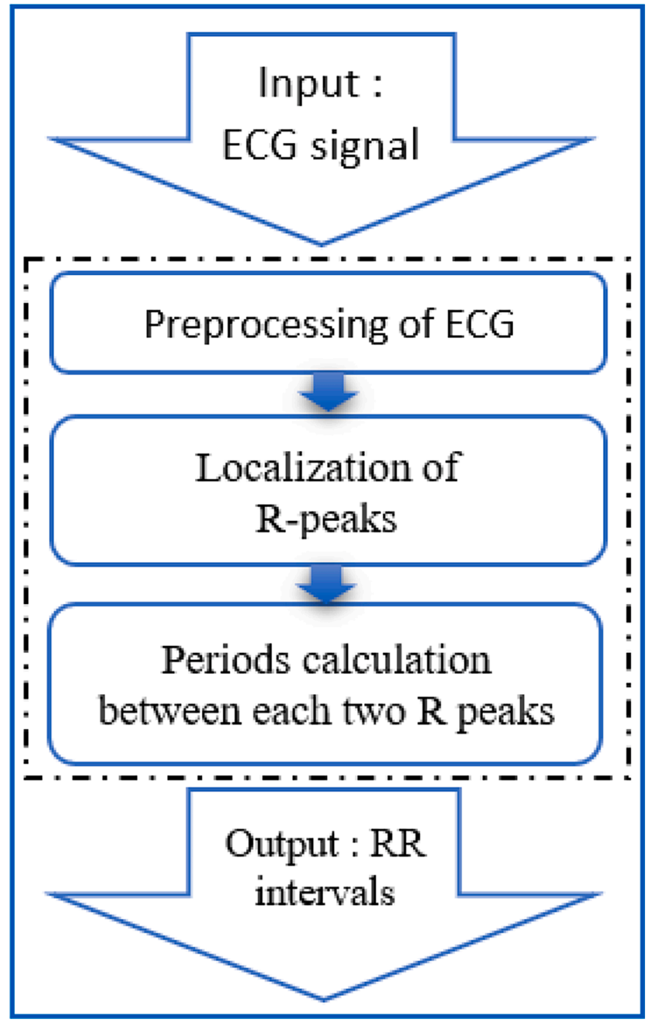


Fig. 4. R-wave detection process.

4.2. Data augmentation strategy

In the aim generating many and balanced ECG samples, we suggest a data augmentation process over four steps, presented in Fig. 3 and illustrated as follows.

4.2.1. Step 1. Detecting R-waves

Based on the R-wave detection algorithm, the R-waves of the QRS complex of the normalized ECG signals are detected R1, R2,…, Rn, as shown in Fig. 5.

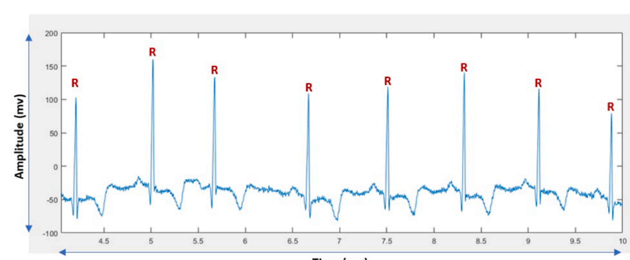


Fig. 5. Detecting R-waves.

4.2.2. Step 2. Periods calculation of R-R intervals

Different periods between successive R-waves (R-R intervals) are then calculated, (R-R interval(1), R-R interval(2),…, R-Rinterval(n)) as shown in Fig. 6.

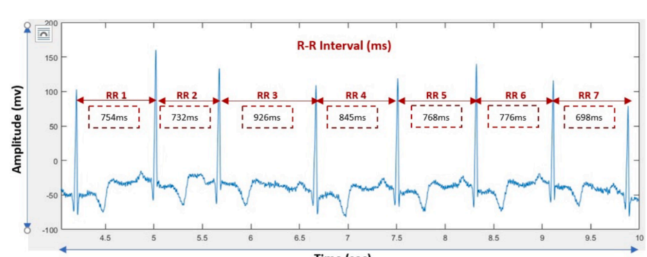


Fig. 6. Periods calculation of R-R interval.

4.2.3. Step 3. Random selection of new R-R intervals

To extend ECG samples and records, all the nn extracted R-R intervals from each signal is arranged and resampled in one signal with a new order. As shown in Fig. 6, RR1, RR2,…, RRn are selected to be arranged and resampled as following: RR7, RR6, RR1,…, RR4. These new R-R intervals are put in the selected ECG signal at random as shown in Fig. 7.

4.2.4. Step 4. R-R intervals concatenation

In this step, the generated and selected R-R intervals are concatenated together and added to previous ECG records which giving rise to a new ECG signal like: RR7, RR6, RR1, RR2, RR3, RR5, RR4, as depicted in

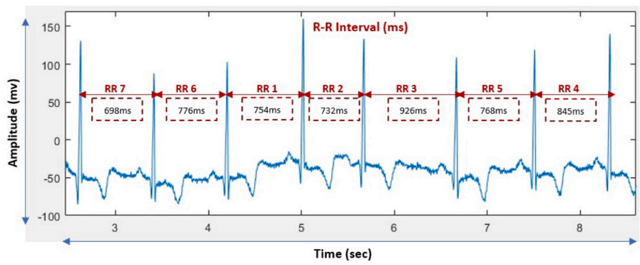


Fig. 7. Random selection of new R-R interval.

Fig. 7.

Our system starts with the use of an ECG dataset benchmark like DREAMER dataset to detect the human emotion, this dataset consists of only 414 samples (records) of ECG signal. The number of samples of the different ECG signal, could be balanced by adding the same number of samples in each signal of any of categories in order to enrich each category. To do so, we randomly extracted 24 samples from each ECG signal, the total number of samples is above 10000 samples. It should be noted that the resampled signals are the same length as the original's ECG recording. In this research activity, we try to examine the effects of various values of dataset size (like dataset containing 5000 or 10000 samples, or more) on neural network training phase.

4.3. HRV features extraction

Standard HRV Measures: In the literature, the ECG in specific HRV signals play a crucial role in research on emotion assessment. In fact, the HRV is a measure that indicates variation in the HR over time as shown in Fig. 8. The ANS is responsible for control this variation through the balance between the SNS and PNS, in order to react to the daily stressors as well as to control the human body's most important systems such as respiration, digestion and heart rate. In summary, HRV is one of the most interesting and noninvasive way to recognize the ANS imbalances of the human because it has a direct impact on the activity of the heart [58,59]. Standard HRV analysis is to extract a variety of parameters that are specified in the frequency domain and also in the time domain [60–62]. In the field of ECG-based emotion recognition, Zhao et al. in [63] studied the differences of HRV indices between six different emotions: happiness, sadness, anger, disgust, fear and neutral, the obtained results proved that there are important differences of HRV indices between these emotions. In this work, we relied on ECG signals to extract a series of features as well as its derived HRV features extracted from frequency-domain, time-domain, and nonlinear domains. Measuring these features is illustrated as follows:

- **Time-Domain Features:** in this domain, we measure the variation in heart rate over time (i.e. the intervals between successive normal cardiac cycles). This variation reflects to do some easy calculations such as: calculate the mean normal-to-normal (NN) intervals, the variance between NN intervals, the standard deviation of NN (SDNN) and the root mean square of differences between adjacent R-R intervals (RMSSD).
- **Frequency-Domain Features:** it is a complex analysis technique, the role of this technique is to show the amount of signal that lies one or more frequency bands (ranges). For the HRV, the technique uses the frequency bands that could tend to correlate with some physiological phenomenon (e. g. Parasympathetic nervous system activity) [65]. In addition, the influence of sympathetic and parasympathetic nerves on HRV can be distinguished by this method very well. From the power spectral density (PSD) analysis, several features are calculated in the frequency domain analysis. However, the Power Spectral Density (PSD) analysis is used to understand HRV, furthermore, there are three spectral bands of the PSD which are: Very Low Frequency (VLF) with spectral components less than 0.04 Hz; Low Frequency (LF) belongs to the interval [0.04, 0.15 Hz] and High Frequency (HF) defined in the interval [0.15, 0.4 Hz].
- **Nonlinear Features:** In the nonlinear analysis, several features are calculated, including ECG-derived respiration (EDR) related parameters, nonlinear dynamics related parameters, Poincaré plot related parameters, and self-correlation related parameters.

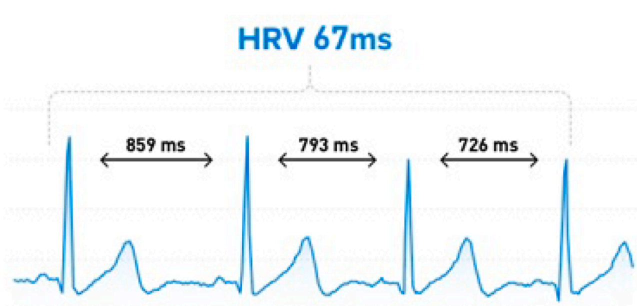


Fig. 8. Heart Rate Variability [64].

Table 2
Notation of features extracted from ECG [68].

Features group	Variable	Unit	Description of the extracted features
Time-Domain features	SDNN	ms	standard deviation of all NN intervals
	SDANN	ms	Standard deviation of the averages of NN intervals in all 5-minute segments of the entire recording
	RMSSD	ms	The square root of the mean of the sum of the squares of differences between adjacent NN intervals
	SDNN index	ms	Mean of the standard deviations of all NN intervals for all 5-minute segments of the entire recording
	SDSD	ms	Standard deviation of differences between adjacent NN intervals
	NN50 count		Number of pairs of adjacent NN intervals differing by more than 50 ms in the entire recording; three variants are possible counting all such NN intervals pairs or only pairs in which the first or the second interval is longer
	pNN50		NN50 count divided by the total number of all NN intervals
	HRV		Total number of all NN intervals divided by the height of the histogram of all NN
	TINN	ms	Baseline width of the minimum square difference triangular interpolation of the highest peak of the histogram of all NN intervals
Frequency-Domain features	Total power	ms^2	Variance of all NN intervals <0.4 Hz
	ULF	ms^2	Ultra low frequency <0.003 Hz
	VLF	ms^2	Very low frequency < 0.003–0.04 H
	LF	ms^2	Low frequency power 0.04–0.15 Hz
	HF	ms^2	High frequency power 0.15–0.4 Hz
	LF/HF ratio	/	ratio of low-high frequency power
Nonlinear features	SD1	/	The standard deviation of the Poincare plot perpendicular to the line-of-identity
	SD2	/	The standard deviation of the Poincare plot along the line-of-identity
	SD1/SD2	/	Ratio of SD1/SD2
	ApEn	/	Measures the complexity or irregularity of the RR series
	SampleEn	/	Sample entropy : a tolerance (r) of 0.2 standard deviation of the R-R interval and an embedding dimension (m) of 2.

The extracted HRV features groups are listed in Table 2. The use of HRV becomes an increasingly popular and important tool to identify the emotion of a human. In this phase, we suggest the use of the HRV Analysis Software (HRVAS) [66] which is a HRV analysis tool developed using Matlab software [67]. It is used to extract HRV features including the time-domain (e.g. SDNN and RMSSD, the frequency-domain (ex: Low-Frequency power (LF), High-Frequency power (HF) and LF/HF) and nonlinear domain (ex: SampEn), more details about the calculation of these parameters are presented in [54].

4.4. Architecture of seven-layer CNN model for ECG emotion recognition system

CNN is usually consisted of two main parts; the former is a feature extractor, which is responsible for an automatic features learning from raw input data, while the latter is considered as a fully connected multi-layer perceptron (MLP). This MLP is responsible for the classification according to the learned features, realized by the first part. The architecture of seven-layer CNN model for ECG emotion recognition system is presented in Fig. 9. The network consists of seven layers, including four convolutional layers (Conv1D), two max pooling layers, one fully connected layers, and one Softmax layer. The convolution operations are performed by the convolutional layers C1, C2, C3 and C4 according to Equation 1, where each layer uses the output of the previous one using the current convolution kernel.

$$x_k^l = f(\sum_{i \in M_k} x_i^{l-1} \ * \ w_{ik} + b_k)(1)$$

Where, x_k^l is the output of the k-th neuron in layer l, M_k represents the effective range of the convolution kernel, w_{ik} is the weight kernel between the i-th neurons in layer $l-1, b_k$ represents the bias of the k-th neuron in layer l and the k-th neuron in layer l, in addition, $f()$ is used as the activation function, we opt for a Rectified Linear Unit (ReLU) as an activation function [55]. First of all, our system has as an input the extracted features vector with a length of 104 features (e.g. SDNN,…). This vector is sent to the first layer C1, which applies a kernel size of 5 (i.e. filter) and 64 feature maps. So, the output of C1 layer is calculated as follows: input_size - (kernel_size - 1) = 104 - (5 −1) = 100 features. Note that the relationship (feature maps-reached features) are written (feature maps@reached features) such as (64@100). The found 100 features are then presented to the next layer C2 with convolution kernel size of 5, which in its turn (i.e. C2) gives rise to 64@96 (as output). Next, max-pooling S1 is used to reduce the number of computations with a kernel size of 2, so the output is 64@48.

The subsampling layers (also called Pooling layers) (i.e. S1 and S2) are used to minimize the input size of the next layer and to extract more useful features. We suggest applying the Max pooling layers to keep only the effective features from the Conv1D layer. The Equation 2 calculates the output of the k-th neuron of the subsampling layers.

$$x_k^l = subsample(x_{k_{cluster}}^{l-1})(2)$$

Where, x_k^l is the output of the k-th neuron of layer l, which is calculated by the down sampled operation performed on the output of k-th cluster of layer l – 1.

Similarly, C3 layer is performed after the max pooling process where the output is 128@46 (here, the filter size is 3). Next layer is C4, having as output 128@44. For the next step, S2 is applied and the output pooling features of S2 (128@22) are rearranged as a feature vector and input to the fully connected layer F1 for classification.

The vector resulting from ECG features extraction state is dispatched to the input neurons of the fully connected network layers (FCN) F1, this might help to perform the training and testing process of the model. The phase of prediction is performed by the final FCN. The Equation 3 is used to calculate the output of the neuron in fully connected layer which is equal to 256.

$$x_k^l = f(\sum_{i=1}^{N} x_i^{l-1} \ * \ w_{ik} + b_k)(3)$$

Where, x_k^l is the output of the k-th neuron in layer l, w_{ik} is the weight vector between the k-th neurons in layer l and the i-th neuron in layer $l-1, b_k$ represents the bias of the k-th neuron in layer l. However, the total number of neurons in layer $l-1$ is defined by N.

In the output layer, softmax activation function is used for the final classification to obtain the five levels of the arousal, valence, and dominance.

5. Experiments and results

5.1. Experiments

In this section, we describe the realized tests and obtained results of a series of experiments including the phase of features extraction and classification. In order to evaluate the effectiveness of the proposed method for ECG data augmentation approach to detect the emotion, comparative experiments are conducted in this work using different classifiers; this is done with and without data augmentation. The data

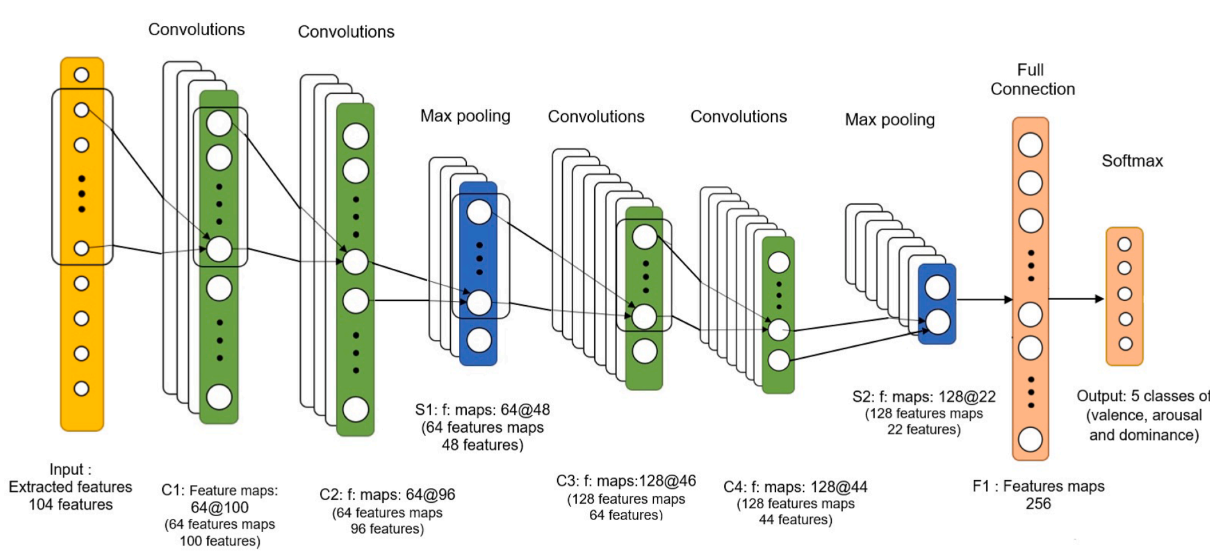


Fig. 9. Architecture of seven-layer CNN model for ECG emotion recognition system.

generated from the proposed method will be used by the classifier in the learning phase. For our contribution, an seven-layer CNN classifier is adopted. The considered classifiers for comparisons are: NN, and SVM. Additionally, we also study the influence of every augmentation of samples on the classifiers by the use of 5000 and more than 10000 samples. We have used R studio platform to train the developed networks for emotion detection with and without data augmentation. We used in this study, the public database DREAMER containing 23 individuals' ECG data (14 males and 9 females), this data was elicited by audio-visual stimuli and using low-cost devices. Moreover, 18 film clips with a duration of 65 – 393s are proposed to tested persons in order to elicit different emotion. Note that the ECG signals were recorded at 256 Hz. Next, the proposed method is run to increase the number and the diversity of ECG dataset. For the phase of extracting features, the HRV data from the ECG signal based on the extracted peaks are measured using HRVAS Software (HRV Analysis) to extract 104 features. After extracting and merging features, CNN is launched to classify and detect emotions.

5.2. Results obtained and discussion

In order to evaluate any classifier in machine learning field, it is common to divide the dataset into two separate sets: a training set and a testing set [69]. To this end, there are different partition schemes that could be applied on the dataset. In our study, we apply both 70–30 technique and 10-fold cross-validation technique. 70–30 technique is performed to illustrate the importance of training data and how vital training information is in the proposed classifier, in this technique, the dataset is split into 70% train data and 30% test data [70]. The 10-fold cross validation technique is used to estimate the skill of a machine learning model on unseen data [71]. Specifically, the 10-fold cross validation involves randomly dividing the dataset into 10 folds or subsets of approximately equal size. Of the 10 folds, a single fold is retained as the validation data for testing the model, and the remaining 9 folds are used as training data. This process is then repeated 10 times, with each of the 10 folds used for testing and the rest for training.

This section describes the results obtained from the experiments that were conducted in this research. Results of the classifiers to recognize emotions in terms of valence, arousal, and dominance levels are listed in the tables below.

5.2.1. Accuracy of arousal detection

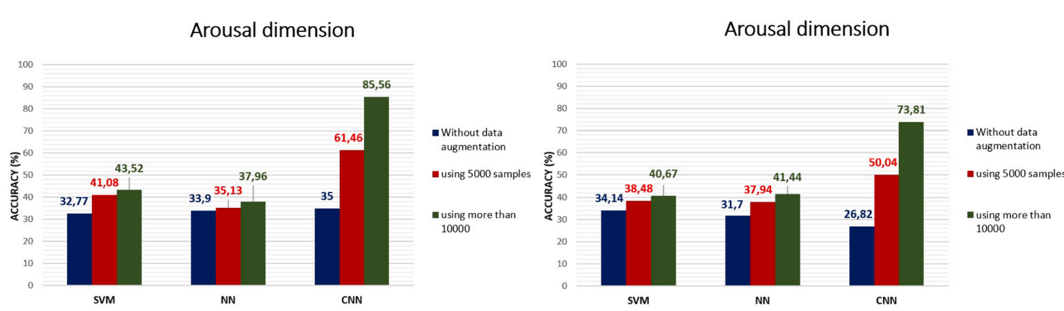
Fig. 10 depicts the accuracy of arousal detected by various machine learning approaches (i.e. SVM, NN, and CNN) in both cases with and without data augmentation. As shown in Fig. 10(a) and for CNN classifier, we have fixed 75.84% samples for training and 24.16% for testing, as user parameters selected in an experimental manner, leading to the best found results. Therefore, for the case when is no data augmentation, CNN records an accuracy of 35% to detect arousal, outperforming NN and SVM, which have given an accuracy of 33.90% and 32.77%, respectively. However, in the Fig. 10(b), it was observed that using 10-fold cross validation, the accuracy without data augmentation is 32.16% for SVM, 31.7% for NN and 26.19% for CNN. It is clear that these values are very moderate, it is due to the limited number of recorders used to train each Machine Learning (ML) approaches in addition to restricted diversity of these processed samples. On the other side, when 5000 samples are considered to augment data to enhance machine learning performance, Fig. 10(a) showed that CNN obtained the best accuracy to detect the arousal with 61.46% compared to NN with 35.13%. Moreover, if we consider more than 10000 samples, we remark that CNN (with 75,84% of samples for training and 24,16% for testing) reached also the best accuracy with 85.56% against the other ML approaches that have given 37.96% for NN, and 43.52% for SVM. Even when we used 10-fold cross validation, the accuracy presented in Fig. 10(b) with data augmentation is increased with 40.31%, 41.44% and 73.81%, for SVM, NN, and CNN, respectively.

5.2.2. Accuracy of valence detection

Fig. 11(a) presents the accuracy comparison of valence dimension between NN, SVM and CNN. For the CNN classifier, we used 63,76% of samples for training and 36,24% for testing, as user parameters experimentally selected given the best found results in this case. Besides, the Fig. 11(b) compares these classifiers using 10-fold cross validation with and without data augmentation. The results without data augmentation showed low classification accuracy with all classifiers with 24.36% for SVM classifier, 22.88% for NN and 22.00% for CNN. The reason for these poor results is the limited number of samples as well as the unbalanced number of each sample in the dataset. Also, Fig. 11(a) showed that with an augmentation of a number of 5000 samples, the CNN obtained the highest accuracy to detect valence dimension with 59.69% compared to the other classifiers such as SVM with 34.76% and NN with 27.53%. Moreover, the results with an augmentation of a number of more than 10000 samples showed that CNN (with 75,75% of samples for training and 24,25% for testing) performs better than the other classifiers, with an accuracy of 95.16% compared to SVM with 36.01% and NN with 31.02%. From Fig. 11(b) and for the case of data augmentation, it can be clearly seen that the CNN performs better than the other classifiers, with an accuracy of 75.1% compared to SVM with 33.6% and NN with 36.32%.

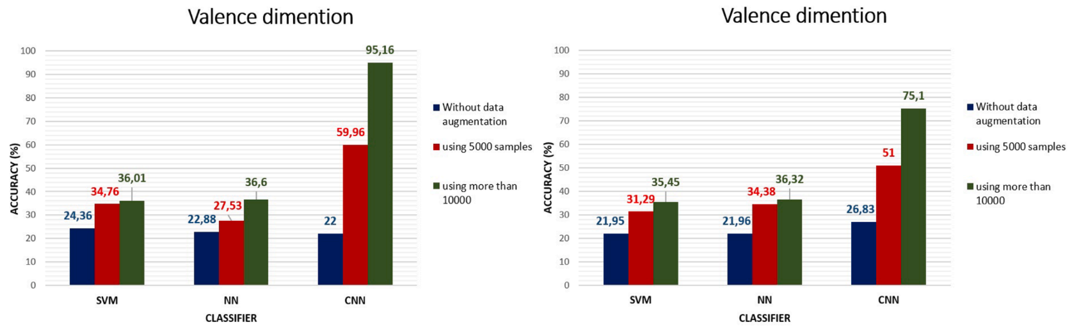
5.2.3. Accuracy of dominance detection

Concerning the dominance dimension in both cases without and with data augmentation, we used 84,30% of samples for training and 15,70% for testing, as user parameters. Fig. 12(a) depicts the accuracy obtained for each model (i.e. SVM, NN and CNN), where Fig. 12(b) presents the accuracy obtained using 10-fold cross validation. The results presented in this figure using the original dataset (i.e without data augmentation) showed that the accuracy of various classifiers is low due to the use of a



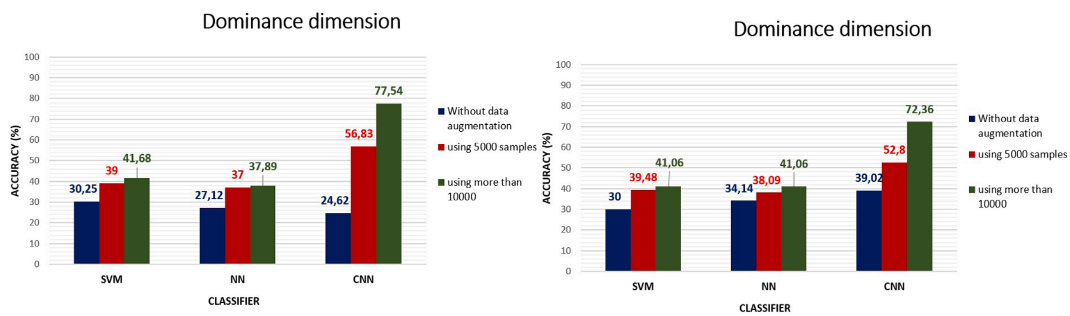
(a) (b)

Fig. 10. Accuracy comparison between CNN and other machine learning with/without data augmentation (emotion expressed in terms of arousal). (a: 70% for training, and 30% for test, b: the 10-fold cross validation).



(a) (b)

Fig. 11. Accuracy comparison between CNN and other machine learning with/without data augmentation (emotion expressed in terms of valence). (a: 70% for training, and 30% for test, b: the 10-fold cross validation).



(a) (b)

Fig. 12. Accuracy comparison between CNN and other machine learning with/without data augmentation (emotion expressed in terms of dominance). (a: 70% for training, and 30% for test, b: the 10-fold cross validation).

small and unbalanced dataset, except for SVM, giving a classification accuracy equal to 30.25% to detect dominance, however, CNN has an accuracy of only 24.62% and NN achieved 27.12%. On the other hand, when the number of samples increases, as presented in Fig. 12(a), the accuracy of the classifiers increases gradually. It can be seen that with an augmentation with 5000 samples, the CNN achieved the best accuracy with 56.83% compared with SVM with 39% and NN with 37%. Additionally, the accuracy of the network is also improved with an increase of more than 10000 samples, where the accuracy of the CNN using 83,10% of samples for training and 16,90% for testing is much higher

Table 3
Confusion matrix for CNN without data augmentation.

a: 70% for training and 30% for test

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**7**	9	8	6	11	**0**	0	0	0	0	**0**	0	0	0	0
	2	7	**6**	6	5	3	6	**12**	4	4	4	4	**5**	1	1	1
	3	4	5	**7**	8	3	8	4	**12**	6	6	6	6	**5**	5	7
	4	9	6	6	**8**	8	4	3	3	**9**	8	2	2	5	**4**	3
	5	3	4	3	3	**5**	2	1	1	1	**2**	1	0	2	3	**2**
Accuracy		**22.00%**					**35.00%**					**24.62%**				

b: 10-fold cross validation

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**1**	1	3	3	2	**0**	1	1	2	0	**0**	0	1	0	0
	2	3	**4**	1	2	0	0	**2**	2	4	1	1	**3**	1	4	3
	3	1	2	**1**	3	0	0	4	**5**	4	2	0	0	**6**	3	1
	4	2	1	1	**3**	1	0	1	1	**3**	1	0	3	2	**5**	2
	5	1	1	1	1	**2**	1	2	1	2	**1**	0	2	2	0	**2**
Accuracy		**26.83%**					**26.82%**					**39.02%**				

than that of the other classifiers with 77.54% compared to SVM with 41.68% and NN with 37.89%. Moreover, the accuracy presented in Fig. 12 (b) using 10-fold cross validation showed that with data augmentation, CNN achieved highest accuracy with 72.36%, compared to SVM with 40.77% and NN with 41.06%.

5.2.4. The confusion matrices: classification correctness

To illustrate the quality of the proposed classifier against SVM and NN for valence, arousal and dominance, Tables 3–8 present the confusion matrices for both cases without/with data augmentation. It is worth noting that the correct classifications (i.e. true positives) are shown in the diagonals of each dimension (i.e. valence, arousal and dominance). Each column of the each matrix represents the instances in a predicted class (i.e. valence, arousal, and dominance), while each row represents the instances in the actual class. In this study, we mention that the total number of the original dataset is 414 and we note that the choice of training or test samples percentages was done in an experimental fashion, with approximately to 70%/ 30% splits between training and test data, leading to the best found results.

- **Without data augmentation case:**

  The Table 3:a shows the confusion matrix of the training results on the test set for CNN. For this, we have used a roughly equal number of samples for each class of valence, arousal and dominance in training; this number is equal to 30 samples per class. So, in terms of valence, a total of 264 samples is used for training with 63.76% and 150 samples for the test with 36.24%. About the arousal, the total of 314 samples is selected for training with 75.84% and 100 samples for test with 24.16%, however, for dominance, we have opted for 349 samples for training with 84.30% and 65 samples for test with 15.70%.

  When we used the SVM classifier, we have applied the same percentages of dataset partition (training/test) for arousal, valance or dominance as follows: 71% of dataset is reserved for training and 29% for the test. Consequently, 295 samples are devoted for training and 119 samples for testing, as depicted in the Table 5:a.

  Also for the NN classifier (see Table 7:a), the same percentages of dataset partition is considered; 71.50% of data are used for training and 28.50% for the test, which means that 296 samples are devoted for training and 118 samples for testing.

  When we analyze Table 3:a, we can see that the CNN is confused when there are no data augmentation with an accuracy of actual/ predicted valence equal to 22.00%, 35.00% for arousal, and 24.62% for dominance. We can also see in Table 5:a that the SVM records only an accuracy of 24.36% for valence, 32.77% for arousal, and 30.25% for dominance. A close outcome is also seen in Table 7:a, which is given by the NN with only 22.88%, 33.90% and 27.12% for valence, arousal and dominance, respectively.

  In addition, we have performed 10-fold cross validation on a training data, the results reached are presented in Table 3:b, 5:b and 7:b, as confusion matrices of optimal model from the 10 generated models, according to CNN, SVM, and NN classifiers, respectively. We can clearly see that all these classifiers records low accuracy, we cite for instance CNN that gives an accuracy of 26.83% for valence, 26.82% for arousal and 39.02% for dominance.

- **With data augmentation case:**

  Considering the total number of original dataset which is 414, the proposed data augmentation increases the number to be 10350, which is calculated as follows: (Total number of original dataset * x) + (Total number of original dataset). Where, x is the number of copies generated from each ECG signal to be added to the original dataset. In our study, x is fixed at 24 (as an experimental parameter), then the new dataset (with data augmentation) will contain: (414*24) + (414) = 10350 samples.

  To evaluate the classification accuracy of the proposed CNN for the case of valence, we have chosen the total of 7840 samples for training with 75.75% and 2510 samples for test with 24.25%. About the arousal, a total of 7850 samples is selected for training with 75.84% and 2500 samples for test with 24.16%, however, for dominance, we have chosen the total of 8600 samples for training with 83.10% and 1750 samples for test with 16.90%. About the SVM, We have selected for all emotion levels, the same sampling as follows: 7301 samples (i.e 70.54%) devoted for training and 3049 samples (i.e 29.46%) for testing, as depicted in Table 6:a.

  Concerning the NN and for valence, we have used 70.04% for training which represents 7249 samples and 29.96% of dataset for test with 3101 samples, however, for arousal and dominance, we consider 7460 samples (i.e. 72.06%) for training and 2890 samples

Table 4
Confusion matrix for CNN with data augmentation.

a: 70% for training and 30% for test

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Actual	1	**482**	5	3	1	1	**195**	0	0	0	1	**0**	0	0	0	0
	2	5	**480**	3	3	16	140	**487**	7	5	4	45	**336**	0	3	2
	3	1	3	**482**	9	14	114	6	**490**	6	9	228	2	**341**	8	1
	4	10	11	11	**484**	28	50	5	3	**489**	8	52	11	9	**339**	6
	5	2	1	1	3	**451**	1	2	0	0	**478**	25	1	0	0	**341**
Accuracy				**95.16%**					**85.56%**					**77.54%**		

b: 10-fold cross validation

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Actual	1	**160**	7	8	23	2	**39**	9	7	9	1	**22**	4	6	3	4
	2	13	**161**	14	16	6	2	**159**	34	44	11	1	**149**	24	34	3
	3	11	19	**151**	22	11	4	23	**245**	41	4	8	28	**203**	27	24
	4	23	22	27	**208**	14	3	18	21	**235**	4	6	29	33	**260**	13
	5	4	4	7	5	**97**	0	10	15	11	**86**	2	12	16	9	**115**
Accuracy				**75.1%**					**73.81%**					**72.36%**		

Table 5
Confusion matrix for SVM without data augmentation.

a: 70% for training and 30% for test

		Valence dimension					**Arousal dimension**					**Dominance dimension**				
		Predicted					**Predicted**					**Predicted**				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**0**	1	0	0	0	**0**	0	0	0	0	**0**	0	0	0	0
	2	0	**0**	0	1	0	0	**1**	2	0	0	1	**1**	2	2	0
	3	7	6	**8**	7	4	1	5	**8**	7	3	2	8	**9**	13	5
	4	15	16	22	**21**	11	3	17	30	**30**	12	2	15	21	**26**	12
	5	0	0	0	0	**0**	0	0	0	0	**0**	0	0	0	0	**0**
Accuracy				**24.36%**					**32.77%**					**30.25%**		

b: 10-fold cross validation

		Valence dimension					**Arousal dimension**					**Dominance dimension**				
		Predicted					**Predicted**					**Predicted**				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**0**	0	0	0	0	**0**	0	0	0	0	**0**	0	0	0	0
	2	0	**0**	0	0	0	0	**1**	0	2	0	0	**2**	1	4	0
	3	2	5	**3**	1	3	1	1	**5**	4	3	1	3	**2**	5	3
	4	4	6	7	**5**	2	0	7	7	**6**	2	1	6	3	**8**	2
	5	1	0	1	0	**1**	0	0	0	0	**2**	0	0	0	0	**0**
Accuracy				**21.95%**					**34.14%**					**30.0%**		

Table 6
Confusion matrix for SVM with data augmentation.

a: 70% for training and 30% for test

		Valence dimension					**Arousal dimension**					**Dominance dimension**				
		Predicted					**Predicted**					**Predicted**				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**75**	28	9	17	26	**1**	0	0	0	0	**1**	0	0	0	0
	2	92	**191**	63	95	21	39	**188**	77	92	64	12	**29**	9	5	8
	3	167	143	**250**	183	82	25	96	**339**	126	73	54	289	**536**	274	204
	4	239	225	339	**497**	163	65	414	479	**759**	170	42	281	330	**675**	269
	5	27	5	12	15	**85**	0	2	0	0	**40**	0	0	1	0	**30**
Accuracy				**36.01%**					**43.52%**					**41.68%**		

b: 10-fold cross validation

		Valence dimension					**Arousal dimension**					**Dominance dimension**				
		Predicted					**Predicted**					**Predicted**				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**42**	9	10	15	14	**1**	0	0	0	0	**1**	0	0	0	0
	2	15	**56**	11	34	11	24	**58**	52	31	19	7	**10**	2	1	13
	3	49	50	**92**	55	39	9	47	**115**	45	38	13	90	**194**	88	68
	4	69	95	97	**147**	63	21	117	156	**233**	53	10	105	105	**213**	108
	5	10	6	7	9	**30**	0	1	1	0	**14**	0	0	0	0	**7**
Accuracy				**35.45%**					**40.67%**					**41.06%**		

(i.e. 27.94%) for test, see Table 8:a.

When analyzing different confusion matrices (case of data augmentation), we can see that CNN classifier records the best classification precision with very low confusion. For instance, the results given by Table 4:a showed that the CNN is able to label and classify correctly each class with accuracy of actual/predicted valence equal to 95.16%, 85.56% for arousal, and 77.54% for dominance.

Nevertheless, the accuracy achieved by SVM presented in Table 6:a is only 36.01% for valence, 43.52% for arousal, and 41.68% for dominance. Also, the Table 8:a expressed low accuracy values of NN classifier equal to 31.02%, 37.96% and 37.89% for valence, arousal and dominance, respectively.

As depicted in Table 4:b, 6:b and 8:b that found from the 10 cross validation, it has been clearly demonstrated that the accuracy is obviously improved especially for CNN, which performs better than the other classifiers (SVM and NN), with an accuracy of 75.1% for valence, 73.81% for arousal and 72.36% for dominance.

5.2.5. *Precision, recall, and F1-Score*

As seen in Table 10, and for all the three dimensions valence, arousal and dominance the CNN reaches high precision with a high recall which

Table 7
Confusion matrix for NN without data augmentation.

a: 70% for training and 30% for test

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**3**	3	3	4	1	**0**	0	2	1	2	**0**	0	0	0	0
	2	0	**0**	3	1	1	0	**2**	0	0	1	1	**0**	2	5	4
	3	9	17	**14**	20	8	2	7	**9**	15	0	4	7	**19**	21	6
	4	3	2	6	**3**	1	3	14	21	**27**	6	0	10	8	**9**	9
	5	5	3	0	4	**4**	0	2	1	1	**2**	0	2	3	4	**4**
Accuracy				**22.88%**					**33.90%**					**27.12%**		

b: 10-fold cross validation

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**1**	0	1	4	0	**1**	0	0	0	0	**2**	0	0	0	0
	2	0	**1**	0	1	0	0	**0**	0	0	0	1	**3**	1	2	0
	3	2	2	**2**	3	3	1	4	**1**	2	0	1	1	**2**	2	1
	4	4	2	6	**4**	1	3	5	8	**10**	5	1	3	10	**6**	4
	5	0	2	0	1	**1**	0	0	0	0	**1**	0	0	0	0	**1**
Accuracy				**21.96%**					**31.70%**					**34.14%**		

Table 8
Confusion matrix for NN with data augmentation.

a: 70% for training and 30% for test

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**161**	95	75	150	51	**1**	0	2	0	0	**5**	0	0	0	0
	2	47	**160**	109	122	53	28	**99**	58	39	20	4	**46**	30	43	8
	3	149	149	**269**	201	83	21	102	**185**	120	26	52	172	**386**	243	165
	4	193	219	230	**319**	145	101	432	564	**778**	247	43	331	390	**602**	234
	5	25	15	12	16	**53**	0	9	13	11	**34**	5	23	28	24	**56**
Accuracy				**31.02%**					**37.96%**					**37.89%**		

b: 10-fold cross validation

		Valence dimension					Arousal dimension					Dominance dimension				
		Predicted					Predicted					Predicted				
		1	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**	**1**	**2**	**3**	**4**	**5**
Actual	**1**	**51**	31	26	34	17	**5**	0	1	1	0	**16**	0	1	1	0
	2	21	**53**	30	36	16	7	**62**	25	25	10	5	**59**	28	27	10
	3	39	61	**115**	69	39	11	46	**105**	71	37	14	50	**106**	71	30
	4	68	53	63	**113**	26	19	83	112	**213**	43	22	98	97	**205**	43
	5	10	7	5	8	**44**	12	21	48	34	**44**	13	21	49	30	**39**
Accuracy				**36.32%**					**41.44%**					**41.06%**		

means that each result restored by a search was relevant. Additionally, a high recall corresponds that the search retrieved all the true positives. Indeed, the F1-score is also calculated using both precision and recall, confirming the highest CNN precision and recall. In contrary, the results obtained without data augmentation in Table 9, the CNN has low precision with low recall (and low F1-score) which means that CNN returns a lot of false positives and returns so few results. Table 11.

5.2.6. K-fold cross validation

In this subsection, we tested the proposed model to check the effectiveness of the CNN classifier on a dataset (with and without data augmentation) with the K-Cross-validation technique, using three statistical metrics which are:

- **RMSE:** Root means square error, that is, how far apart the predicted values are from the observed values in a dataset.
- **MAE:** The mean absolute error, it corresponds to the average absolute error between the model prediction and the actual observed data.
- **R-squared:** The measure of the correlation between the predictions made by the model and the actual observations.

Table 10
Precision, recall and F1-score of CNN with data augmentation.

	Valence					Arousal					Dominance				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Precision	0.979	0.946	0.965	0.889	0.984	0.994	0.757	0.784	0.881	0.993	/	0.870	0.587	0.812	0.929
Recall	0.964	0.960	0.964	0.968	0.902	0.390	0.974	0.980	0.978	0.956	0.00	0.960	0.974	0.968	0.974
F1-score	0.971	0.953	0.965	0.927	0.941	0.560	0.852	0.871	0.927	0.974	/	0.913	0.733	0.884	0.951
Accuracy			**95.16%**					**85.56%**					**77.54%**		

Table 9
Precision, recall and F1-score of CNN without data augmentation.

	Valence					Arousal					Dominance				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Precision	0.170	0.222	0.259	0.216	0.277	/	0.400	0.333	0.333	0.285	/	0.416	0.172	0.250	0.250
Recall	0.233	0.200	0.233	0.266	0.166	0.000	0.600	0.600	0.450	0.100	0.000	0.384	0.384	0.307	0.153
F1-score	0.197	0.210	0.245	0.238	0.208	/	0.480	0.428	0.383	0.148	/	0.400	0.238	0.275	0.190
Accuracy			**22.00%**					**35.00%**					**24.62%**		

Table 11
The results of K-Fold Cross-Validation of CNN (With and without data augmentation).

	Without data augmentation			With data augmentation		
	Valence	Arousal	Dominance	Valence	Arousal	Dominance
MAE	1.439	1.261	1.119	0.443	0.413	0.449
RMSE	1.893	1.640	1.535	0.985	0.885	0.934
R-squared	0.001	0.003	0.001	0.519	0.439	0.392

The results obtained using 10-fold cross validation showed that the lower values of RMSE and MAE reflect the good ability of the model to accurately predict the data, conversely, the high value of R-squared means that the model can predict the actual observations and the model performance is good.

5.2.7. PR and ROC curves

We also evaluate Precision-Recall (PR) Curve and Receiver Operating Characteristic (ROC) curve. PR and ROC curves are considered as metric tools to evaluate the performance of the CNN classifier. The PR curve is a graph that represents Precision values against the Recall values, where ROC curve is a graph, representing true positives rate (TPR) values depending on false positive rate values (FPR). On the one hand, Figs. (ab)(ab)(ab)13–15 display PR and ROC curves, of valence, arousal, and dominance given by the CNN classifier without data augmentation, respectively. In these cases the CNN is not able to differentiate between the positive and negative classes.

On the other hand, the PR and ROC curves of Figs. 13(c,d), 14(c,d) and 15(c,d) given by the CNN classifier with data augmentation, yield an excellent performance in order to distinguish between positive and negative classes.

6. Conclusion

Detecting human emotion using ECG signal is an important research domain with promising application future. However, the existing methods to detect emotion using ECG are not able to identify accurately the emotion of a human. Consequently, it is very important to make this process automatic and to predict human emotion more precisely. Moreover, training CNN to be able to perform the task of detecting emotion using ECG signal needs a huge amount of ECG data. Furthermore, it is very difficult to provide this large number of data due to the sensitive nature of this medical data. To deal with this problem, we proposed, a novel ECG data augmentation strategy, this strategy is capable of generating artificial data in a way that resembles existing ones. Then, we have developed a predictive model using an CNN as a classifier compared to other machine learning models.

Our proposal extracts RR intervals and randomly concatenates them from ECG episode intervals to form new samples of ECG signals. Using the DREAMER database, the emotions are expressed in terms of valence, arousal and dominance. Additionally, the HRV features were extracted from RR interval series time. As shown in the experimental results, we came to the conclusion that the performances of all the classifiers used are affected and improved differently from each other by the proposed method of ECG data augmentation. The CNN classifier achieved the best performance in terms of valence, arousal and dominance with a recognition accuracy rates of 95.16%, 85.56% and 77.54%, respectively. In this research, we focus on the data augmentation strategy, whose goal is to improve the classification performance as well as to increase the size and diversity of the dataset. According to the results obtained, the

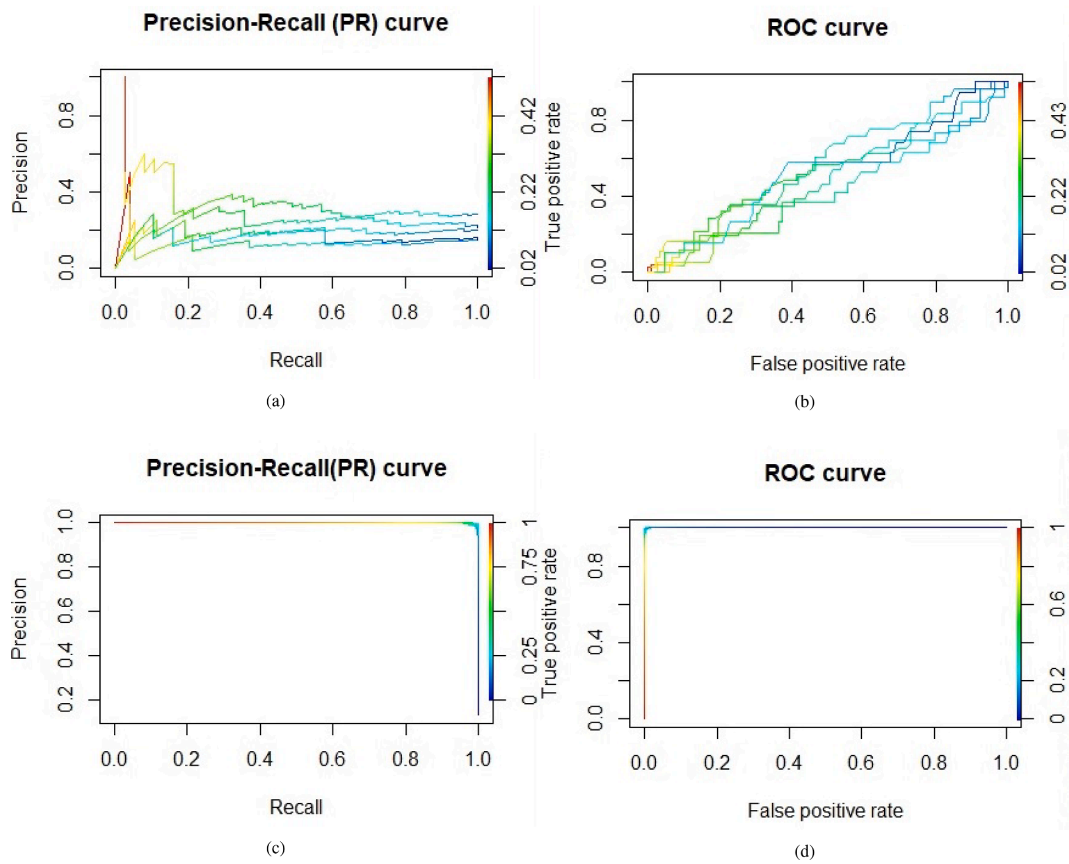


Fig. 13. Valence dimension: (a: PR curve, b: ROC curve [without data augmentation]), (c: PR curve, d: ROC curve [with data augmentation]).

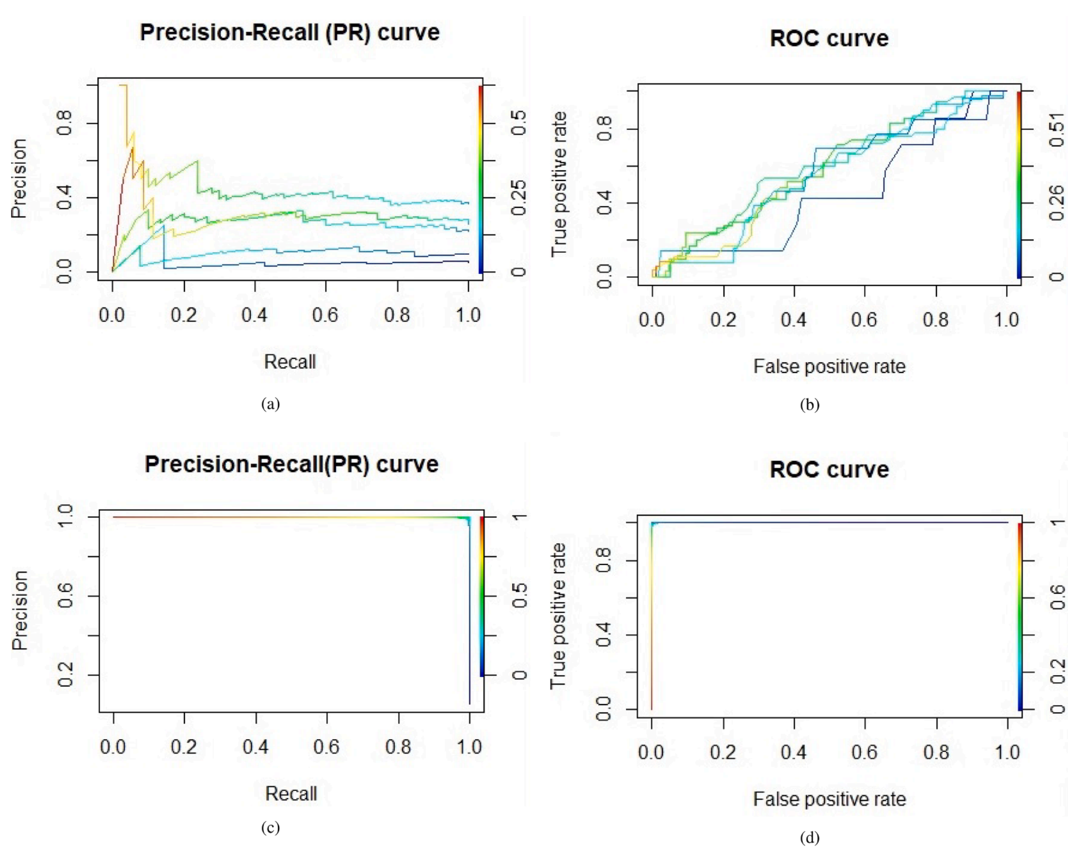


Fig. 14. Arousal dimension: (a: PR curve, b: ROC curve [without data augmentation]), (c: PR curve, d: ROC curve [with data augmentation]).

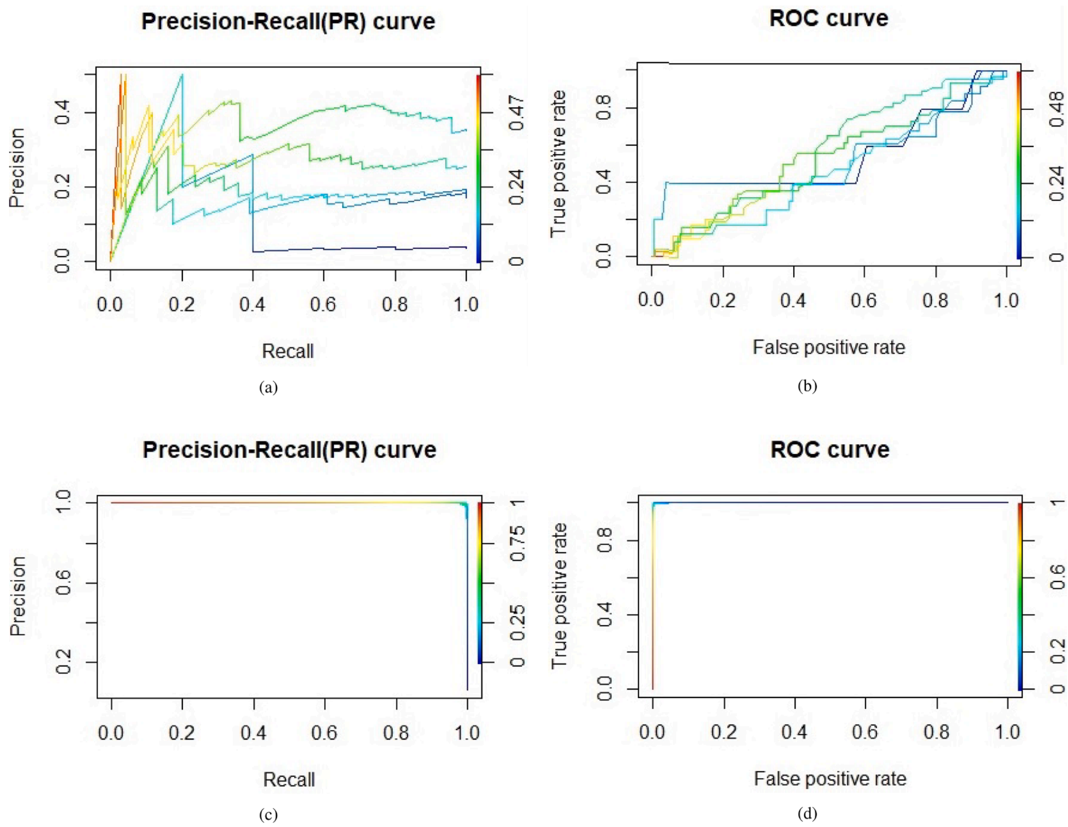


Fig. 15. Dominance dimension: (a: PR curve, b: ROC curve [without data augmentation]), (c: PR curve, d: ROC curve [with data augmentation]).

datasets with a great number of samples are the best and optimal to improve classification performance and therefore to the emotion prediction, it is obtained when the number of instances is increased.

As a future research direction, we suggest to conduct a new study aiming at conceiving an embedded system that could be integrated in different tiny devices and machines like smart watches, smartphones, on-board computer, etc. This could help to make concrete this proposal as a real-world application. To do that, some real-time and operating concerns could be tackled like respecting time constrain and task scheduling to ensure an integrated functioning of this system within a complicate computational device.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] M.S. Likitha, S.R.R. Gupta, K. Hasitha et al., " Speech based human emotion recognition using MFCC," in international conference on wireless communications, signal processing and networking (WiSPNET). IEEE, 2017, p. 2257-2260.
[2] S. Albanie, A. Nagrani, A. Vedaldi, et al., Emotion recognition in speech using cross-modal transfer in the wild, in: Proceedings of the 26th ACM international conference on Multimedia, 2018, pp. 292–301.
[3] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, et al., Emotion recognition in human-computer interaction, IEEE Signal Process. Mag. 18 (1) (2001) 32–80.
[4] S. Katsigiannis, N. Ramzan, DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost offthe- shelf devices, IEEE J. Biomed. Health Inform. 22 (1) (2017) 98–107.
[5] P. Sarkar and A. Etemad, " Self-supervised learning for ecg-based emotion recognition, " in ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE., 2020, p. 3217–3221.
[6] W. Tao, C. Li, R. Song, et al., EEG-based Emotion Recognition via Channel-wise Attention and Self Attention, IEEE Trans. Affective Comput. (2020).
[7] J. Posner, J.A. Russell, B.S. Peterson, The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology, Develop. Psychopathol. 17 (3) (2005) 715.
[8] Z. Li, C. Wang, A.F.T. Mak, et al., Effects of acupuncture on heart rate variability in normal subjects under fatigue and non-fatigue state, Eur. J. Appl. Physiol. 94 (5–6) (2005) 633–640.
[9] R.W. Picard, E. Vyzas, J. Healey, Toward machine emotional intelligence: analysis of affective physiological state, IEEE Trans. Pattern Anal. Mach. Intell. 23 (10) (2001) 1175–1191.
[10] E. Vyzas and R.W. Picard, " Affective pattern classification, " in Emotional and Intelligent: The Tangled Knot of Cognition., 1998, vol. 176182.
[11] J.A. Healey, R.W. Picard, Detecting stress during real-world driving tasks using physiological sensors, IEEE Trans. Intell. Transp. Syst. 6 (2) (2005) 156–166.
[12] M. Nardelli, G. Valenza, A. Greco, et al., Recognizing emotions induced by affective sounds through heart rate variability, IEEE Trans. Affective Comput. 6 (4) (2015) 385–394.
[13] P. Sarkar, K. Ross, A.J. Rodenbura et al., " Classification of cognitive load and expertise for adaptive simulation using deep multitask learning," In: 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII). IEEE, p. 1–7, 2019.
[14] J. Zhu, L. Ji and C. Liu et al., Heart rate variability monitoring for emotion and disorders of emotion, in Physiological measurement 40(6) (2019) 064004.
[15] A. Mikoajczyk, M. Grochowski, " Data augmentation for improving deep learning in image classification problem," in: 2018 international interdisciplinary PhD workshop (IIPhDW). IEEE, p. 117–122, 2018.
[16] S. Katsigiannis, N. Ramzan, " DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost offthe- shelf devices, " [Online]. Available:https://zenodo.org/record/546113#.XvfW3efgrIU. Accessed on: jun 9 2020.
[17] D.B. Lindsley, Emotion (1951).
[18] J.A. Russell, A circumplex model of affect, J. Personality Soc. Psychol. 39 (6) (1980) 1161.
[19] A.B. Warriner, V. Kuperman, M. Brysbaert, Norms of valence, arousal, and dominance for 13,915 English lemmas, Behav. Res. Methods 45 (4) (2013) 1191–1207.
[20] R. Kehrein, " The prosody of authentic emotions," in Speech Prosody 2002, International Conference, 2002.
[21] G.K. Verma, U.S. Tiwary, Affect representation and recognition in 3d continuous valence arousal dominance space, Multimedia Tools Appl. 76 (2) (2017) 2159–2183.
[22] F.H. Wilhelm, J.A. Rattel, M. Wegerer, et al., Attend or defend? Sex differences in behavioral, autonomic, and respiratory response patterns to emotion eliciting films, Biolog. Psychol. 130 (2017) 30–40.

[23] E. Rendon-Velez, P.M. Van Leeuwen, R. Happee, et al., The effects of time pressure on driver performance and physiological activity: a driving simulator study, Transp. Res. Part F 41 (2016) 150–169.

[24] A. Goshvarpour, A. Abbasi, A. Goshvarpour, An accurate emotion recognition system using ECG and GSR signals and matching pursuit method, Biomed. J. 40 (6) (2017) 355–368.

[25] M.P. Barnard, P. Chapman, The effects of instruction and environmental demand on state anxiety, driving performance and autonomic activity: Are ego-threatening manipulations effective? Transp. Res. Part F 55 (2018) 123–135.

[26] B. Scott-Parker, Emotions, behaviour, and the adolescent driver: A literature review, Transp. Res. Part F 50 (2017) 1–37.

[27] K. Rymarczyk, C. Biele, A. Grabowska, et al., EMG activity in response to static and dynamic facial expressions, Int. J. Psychophysiol. 79 (2) (2011) 330–333.

[28] B. Reimer, B. Mehler, J.F. Coughlin, et al., The impact of a naturalistic hands-free cellular phone task on heart rate and simulated driving performance in two age groups, Transp. Res. Part F 14 (1) (2011) 13–25.

[29] K.N. Minhad, S.H.M. Ali, M.B.I. Reaz,"Happy-anger emotions classifications from electrocardiogram signal for automobile driving safety and awareness," J. Transport Health, vol. 7, p. 75-89, 2017.

[30] S. Nita, S. Bitam and Mellouk.A, " A Body Area Network for Ubiquitous Driver Stress Monitoring based on ECG Signal," in 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob). IEEE., 2019. p. 1–6.

[31] S. Nita, S. Bitam and Mellouk.A, " An enhanced random forest for cardiac diseases identification based on ECG signal," in: 2018 14th International Wireless Communications and Mobile Computing Conference (IWCMC). IEEE., 2018. p. 1339-1344.

[32] J.T. Catalano, "Guide to ECG analysis," Lippincott Williams & Wilkins., 2002.

[33] M.S. Müller, A.L. Vyssotski, M. Yamamoto et al., " Heart rate variability reveals that a decrease in parasympathetic ('rest-and-digest') activity dominates autonomic stress responses in a free-living seabird", Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology., vol. 212, p. 117–126. 2017. 55.

[34] F. Agrafioti, D. Hatzinakos, A.K. Anderson, ECG pattern analysis for emotion detection, IEEE Trans. Affective Comput. vol. 3, no 1, p. 102–115. 2011. 54.

[35] J. Kim, E. André, Emotion recognition based on physiological changes in music listening, IEEE Trans. Pattern Anal. Mach. Intell. 30 (12) (2008) 2067–2083.

[36] P. Cao, X. Li, K. Mao, et al., novel data augmentation method to enhance deep neural networks for detection of atrial fibrillation, Biomed. Signal Process. Control 56 (2020), 101675.

[37] S. Haradal, H. Hayashi, S. Uchida, Biosignal data augmentation based on generative adversarial networks, in: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2018, pp. 368–371.

[38] B.K. Iwana and S. Uchida, " Time Series Data Augmentation for Neural Networks by Time Warping with a Discriminative," Teacher arXiv preprint arXiv: 2004.08780, 2020.

[39] Y.L. Hsu, J.S. Wang, W.C. Chiang, et al., Automatic ecg-based emotion recognition in music listening, IEEE Trans. Affective Comput. (2017).

[40] J.A.M. Correa, M.K. Abadi, N. Sebe, et al., Amigos: A dataset for affect, personality and mood research on individuals and groups, IEEE Trans. Affective Comput. (2018).

[41] P. Sarkar, A. Etemad," Self-supervised ecg representation learning for emotion recognition," IEEE Trans. Affective Computing, 2020.

[42] L. Santamaria-Granados, M. Munoz-Organero, G. Ramirez-Gonzalez, et al., Using deep convolutional neural network for emotion detection on a physiological signals dataset (AMIGOS), IEEE Access. 7 (2018) 57–67.

[43] G.B. Moody, R.G. Mark, The impact of the MIT-BIH arrhythmia database, IEEE Eng. Med. Biol. Mag. 20 (3) (2001) 45–50.

[44] A.R. Aslam, M.A.B. Altaf, " An 8 channel patient specific neuromorphic processor for the early screening of autistic children through emotion detection,"In: 2019 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE. p. 1–5, 2019.

[45] A.R. Aslam, M.A.B. Altaf, An on-chip processor for chronic neurological disorders assistance using negative affectivity classification, IEEE Trans. Biomed. Circuits Syst. 14 (4) (2020) 838–851.

[46] F. Wang, S.H. Zhong, J. Peng, et al., Data augmentation for eeg-based emotion recognition with deep convolutional neural networks, IEEE Trans. Biomed. Circuits Syst. 14 (4) (2020) 838–851.

[47] T.T. Um, F.M. Pichler, S. Endo, et al., Data augmentation of wearable sensor data for parkinsons disease monitoring using convolutional neural networks, in: Proceedings of the 19th ACM International Conference on Multimodal Interaction, 2017, pp. 216–220.

[48] N. Nonaka, J. Seita, " Data Augmentation for Electrocardiogram Classification with Deep Neural Network.,": arXiv preprint arXiv:2009.04398., 2020.

[49] Z. Yin, M. Zhao, Y. Yang, et al., Recognition of emotions using multimodal physiological signals and an ensemble deep learning model, Comput. Methods Programs Biomed. 140 (2017) 93–110.

[50] S.Z. Bong, K. Wan, M. Murugappan, et al., Implementation of wavelet packet transform and non linear analysis for emotion classification in stroke patient using brain signals, Biomed. Signal Processing Control. 36 (2017) 102–112.

[51] Y. Wei, Y. Wu, J. Tudo, A real-time wearable emotion detection headband based on EEG measureme, Sens. Actuators A: Phys. 263 (2017) 614–621.

[52] P. Zhong, D. Wang, C. Miao, EEG-based emotion recognition using regularized graph neural networks, IEEE Trans. Affective Comput. (2020).

[53] M.N. Dar, M.U. Akram, S.G. Khawaja, et al., CNN and LSTM-Based Emotion Charting Using Physiological Signals, Sensors 20 (16) (2020) 4551.

[54] H. Shi, L. Yang, L. Zhao, et al., Differences of heart rate variability between happiness and sadness emotion states: a pilot study, J. Med. Biol. Eng. 37 (4) (2017) 527–539.

[55] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th international conference on machine learning, 2010, p. (ICML-10)..

[56] G.E. Billman, Heart rate variability a historical perspective, Front. Physiol. 2 (2011) 86.

[57] S. Dong, F. Xu, B. Lingwood et al., "R-wave detection: A comparative analysis of four methods using newborn piglet ECG," In: 10th International Conference on Information Science, Signal Processing and their Applications (ISSPA 2010), IEEE, p. 320–323, 2010.

[58] Oura Team, February 10, 2020. [Online]. Available:https://ouraring.com/what-is-heart-rate-variability/. Accessed on: Jun 9, 2020.

[59] Marcelo Campos. MD, October 22, 2019. [Online]. Available:https://www.health.harvard.edu/blog/heart-rate-variability-new-waytrack-well-2017112212789. Accessed on: Jun 19, 2020.

[60] G. Valenza, A. Lanata, E.P. Scilingo, The role of nonlinear dynamics in affective valence and arousal recognition, IEEE Trans. Affective Comput. 3 (2) (2011) 237–249.

[61] H.G. Kim, E.J. Bai, D.S. Lee, et al., Stress and heart rate variability: A meta-analysis and review of the literature, Psychiatry Investig. 15 (3) (2018) 235.

[62] U.R. Acharya, K.P. Joseph, N. Kannathal, et al., Heart rate variability: a review, Med. Biolog. Eng. Comput. 44 (12) (2006) 1031–1051.

[63] L. Zhao, L. Yang, H. Shi et al., " Evaluation of consistency of HRV indices change among different emotions," in: 2017 Chinese Automation Congress (CAC). IEEE., 2017. p. 4783-4786.

[64] Mark Van Deusen, September 30, 2019. [Online]. Available:https://www.whoop.com/thelocker/heart-rate-variability-hrv/. Accessed on: May 6, 2020.

[65] [Online]. Available:https://help.elitehrv.com/article/159-hrv-frequency/. Accessed on: jun 9 2020.

[66] [Online]. Available:https://sourceforge.net/projects/hrvas/. Accessed on: jun 9 2020.

[67] J.T. Ramshur, " Design, evaluation, and application of heart rate variability analysis software (HRVAS)," Ph.D. Thesis. University of Memphis, 2010.

[68] Sztajzel, et al., Heart rate variability: a noninvasive electrocardiographic method to measure the autonomic nervous system, Swiss Med. Weekly 134 (35–36) (2004) 514–522.

[69] A. Bhardwaj, A. Tiwari, Breast cancer diagnosis using genetically optimized neural network model, Expert Syst. Appl. 42 (10) (2015) 4611–4620.

[70] D. Acharya, N. Varshney, A. Vedant, et al., An enhanced fitness function to recognize unbalanced human emotions data, Expert Syst. Appl. 166 (2021), 114011.

[71] C. Xu, A.S. Ding, S.S. Liao, A privacy-preserving recommendation method based on multi-objective optimisation for mobile users, Int. J. Bio-Inspired Comput. 16 (1) (2020) 23–32.

Sihem Nita (sihem.nita@univ-biskra.dz) is currently with Department of Computer Science, LESIA Laboratory, University of Biskra, Algeria. She received an Engineer degree in computer science from the University Biskra, Algeria in 2010, and the master's degree in computer science speciality 'Decision Support System and multimedia' from the same university in 2015.

She is currently pursuing the Ph.D. degree with the Computer Science, LESIA Laboratory, University of Biskra. Her main areas of research interest are E-health, Human emotion detection, Body sensor networks and machine learning.

Salim Bitam (s.bitam@univ-biskra.dz) is a full professor in the Computer Science Department at the University of Biskra, Algeria. He received an Engineer degree in computer science from the University of Constantine, Algeria, his Master's and Ph.D. in computer science from the University of Biskra, and a Doctorate of Sciences (Habilitation) diploma from the Higher School of Computer Science - ESI, Algiers, Algeria. His main research interests are vehicular ad hoc networks, cloud computing, and bio-inspired methods for routing and optimization. He has to his credit more than 35 publications in journals, books, and conferences, for which he has received two best paper awards. He has served as an editorial board member and a reviewer of several journals for IEEE, Elsevier, Wiley, and Springer, and on the technical program committees of several international conferences (IEEE GLOBECOM, IEEE ICC, IEEE/RSJ IROS, and others).

Matthieu Heidet, MD, PhD, is an emergency physician working at the emergency department and EMS (SAMU 94) of the Henri Mondor University Hospital in Créteil, France. He is member of the EA-3956 lab (Control in intelligent networks, CIR) of the University of Paris-Est Créteil (UPEC). His research focuses on the optimization of access to prehospital care, with the use of geographics, artificial intelligence, and technological innovations. He is particularly interested in the reduction of socioeconomic inequalities in access to prehospital care. He participated in the creation of the France-Canada registry for out-of-hospital cardiac arrest (RéACanROC).

Abdelhamid Mellouk (mellouk@u-pec.fr) is a Full Professor at University of Paris-Est Créteil (UPEC), IUT CV Networks & Telecommunications (N&T) Department, EPISEN IT4H Departement and LiSSi/TincNet Laboratory France. He graduated in computer network engineering from the Computer Science High Eng. School, University Oran1, Algeria, and the University of Paris Sud XI Orsay, received the Ph.D. in computer science from the same university, and a Doctorate of Sciences (Habilitation) diploma from UPEC. Founder of the Network Control Research activity in UPEC with extensive international academic and industrial collaborations, his general area of research is in adaptive realtime bio-inspired control for high-speed new generation dynamic wired/wireless networking in order to maintain acceptable quality of service/experience for added value services. He is an active member of the IEEE Communications Society and held several offices including leadership positions in IEEE Communications Society Technical Committees. He has published/coordinated 11 books and several refereed international publications in journals, conferences, and books, in addition to numerous keynotes and plenary talks in flagship venues. He serves on the Editorial Boards or as Associate Editor for several journals, and he is chairing or has chaired (or co-chaired) some of the top international IEEE conferences and symposia.