

Recognition of valence judgments in music perception using electrocardiographic signals and machine learning

Ennio Idrobo-Ávila¹, Humberto Loaiza-Correa¹, Flavio Muñoz-Bolaños², Leon van Noorden³, Rubiel Vargas-Cañas⁴

¹ PSI - Percepción y sistemas inteligentes, Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia

{ennio.idrobo,humberto.loaiza}@correounivalle.edu.co

² CIFIEX - Ciencias Fisiológicas Experimentales, Departamento de Ciencias Fisiológicas, Universidad del Cauca, Popayán, Colombia

fgmunoz@unicauca.edu.co

³ IPEM - Institute for Psycho-acoustics and Electronic Music, Department of Art, Music and Theatre Sciences, Ghent University, Ghent, Belgium

leonvannoorden@mac.com

⁴ SIDICO – Sistemas Dinámicos de Instrumentación y Control, Departamento de Física, Universidad del Cauca, Popayán, Colombia

rubiel@unicauca.edu.co

Abstract. Music can interact with humans through sentiment and emotions and aid in relieving stress. While human - computer interaction (HCI) has begun to use ECG as a means of recognizing music-elicited emotions, the need remains for research related to implicit interaction interfaces in HCI systems, to determine how they might enhance human bandwidth in control or interaction. A system is proposed to recognize or classify emotional responses to sounds - from music and nature - based on ECG signals and a single scale of perception. It was possible to recognize the valence judgment of each subject using binary evaluation and ECG signals. Outcomes revealed a performance of 0.91 AUC and 0.86 accuracy in recognizing emotional valence elicited by the music offered. Subjects were found to give more uniform responses to nature sounds than to music.

Keywords: Emotion, Music, Nature Sounds, Perception

1 Introduction

Stress may affect human beings in beneficial ways, preserving cell homeostasis [1] and enhancing the manner in which people learn and remember [2]. Chronic exposure to severe conditions of stress, however, can lead to a number of obstacles to human health. Stress disorders affect the central nervous system, immune system function, and the condition of the gastrointestinal, endocrine, and cardiovascular systems [1]. Stress can furthermore produce changes in heart-brain interaction [3]. Due to the role of stress in human health, physicians require a wide range of therapeutic tools for

treatment. Pharmacological medication and nutraceuticals and non-pharmacological interventions are all relevant [1]. Non-pharmacological music interventions have shown a reduction in anxiety [4] and stress in psychophysiological variables [5] [6].

Research has examined how music may help with stress management. Music interacts with humans through their sentiments and emotions. The use of electrocardiographic signals (ECG) and heart rate variability (HRV) can help to assess states of stress [7] [8] [9] as well as emotions elicited by music [3] [10]. HRV revealed the varying music capacity for reducing stress [9] and how music could produce arousal responses and ease stress [11]. Pleasant and unpleasant music produced changes in periodic repolarization dynamics [12] [13]. ECG and galvanic skin response (GSR) were also considered in assessing the influence of sleep on emotional perception of music, revealing an association between emotion and sleep duration [14]. Another study showed that an increase in heart and respiratory rates occurred while subjects listened to happy music, whereas only heart rate increased with sad music; valence modified heart rate whereas arousal produced changes in respiratory rate [15].

Others used ECG and developed systems to classify different states or responses to music. Systems were developed for instance to recognize conditions before and after listening to music using: two musical genres [16]; responses to music with four different emotional characters [17]; and four emotional responses to music able to induce emotions related to four quadrants in the model of valence-arousal: joy, tension, sadness, and peacefulness [18]. Most research used classic statistical analysis tools [9] [11] [12] [13] [14] [15], while others used machine learning tools such as k-means, naïve Bayes and artificial neural networks [16], probabilistic neural network [17], and least squares support vector machine [18].

The field of human-computer interaction (HCI) has not yet fully explored ECG as a measure for recognizing conditions of stress and emotions elicited by music. But examples in the literature include the following: A system for reducing human effort and improving the practicality of HCI was developed, comprising a music player that selects songs depending on the emotion of the user. The system recognized four emotions (happy, angry, sad and neutral) by means of facial expression, using convolutional neural networks [19]. Another system for playing notes on a musical score was developed, controlled by eye-tracker and head movements on fixing the gaze on a point on the score [20]. Elsewhere, a structure was created for multimodal control of musical performance and sonic interaction, through explicit interaction with tangible objects (pucks), and implicit interaction by measuring brain activity (electroencephalography-EEG) and ECG [21]. In other studies, physiological measurements were used to perform sound synthesis and control of tempo in a digital musical instrument, while a system was developed to detect emotion using EEG, pulse, and blood pressure, recommending color and music according to emotional states of users of a self-driving vehicle, classifying emotions by support vector machines [22].

However, since implicit interaction interfaces (e.g. that use physiological measurements) are still emerging, research is needed on such interfaces in HCI systems, to determine to what extent they might enhance human bandwidth in control or interaction [21]. The use of physiological signals to generate music automatically was thus proposed in the field of physiological computing [23]. Although physiological re-

sponse elicited by music has been assessed using such signals as ECG and HRV [11] [16], most systems do not incorporate each subject's own perception. Most considered a two-dimensional emotion space [17]. Research carrying out the classification of emotions has generally used music previously categorized according to the emotion it is able to produce (emotional music). Emotions associated with sadness, peacefulness, and happiness have commonly been considered, together with emotions related to tension such as threat, scary, or fear [14] [17] [18] [24]. There is room for improvement, as classification rates oscillate between 60 and 90% [18] [22] [25] [26] in the best cases.

1.1 Our contribution

Given the above history, recognizing that ECG signals are seldom used for emotion recognition and that very little research has associated heart and emotions using machine learning [27], we have constructed a system to recognize or classify emotional responses to music and nature sounds, from ECG signals and using a single perception scale.

Our hypothesis is that emotional responses to such sounds, measured by a single scale, can be recognized using only ECG signals classified by machine learning techniques. For future applications to be able to take decisions aimed at reducing stress in subjects based on their emotional response, it is clearly important that machines can recognize states of stress. Further studies are also required regarding pleasant and unpleasant emotions elicited by music, an area of research "still in its infancy" [3] by which sounds and their influence could enhance the quality of human life. Similarly, physiology-based interfaces in HCI have also been suggested as a means of improving single- and multi-user experiences [21].

2 Methodology

2.1 Data acquisition and experimental procedure

Data acquisition required the participation of 23 healthy adult volunteers - 9 females and 14 males, with a mean age of 25.5 yrs (SD=6.8) - in an experimental auditive procedure. All gave informed consent. The ethical committee of the Universidad del Cauca approved the informed consent and the experimental procedures in their entirety. Subjects were assessed through an audiometry exam, and were fully instructed about the procedures of the experimental phase. For the presentation of auditive stimuli, subjects were in a supine position, equipped with noise-cancelling headphones (**Fig. 1.**) and instructed to keep their eyes closed. Their heart electrical activity was then recorded using an ECG acquisition system (OpenBCI - Cyton Biosensing Board (8-channels) [28]); the electrode position in the ECG corresponded to that of the lead II. ECG was recorded with a sample rate of 250 Hz. Each subject was tested individually.

Once the subjects were prepared, music and nature sounds were played randomly to them. The 11 music tracks comprised instrumental song sections; ten nature tracks made up the test. At the end of each track, the subjects opened their eyes and assessed their perception (by means of a graphical user interface projected on the ceiling above them) on a scale of 1 to 9, with low figures representing a negative response and higher figures reflecting a more positive reaction (these responses were then grouped in binary form – as “Positive” or “Negative” to carry out the subsequent classification). After subjects recorded their perception, a 15-second period of silence was incorporated before the next track. Volume was normalized across all tracks and each track comprised 12 seconds. Music included Colombian music from the Pacific coast, Latin Jazz, Merengue, Rock in Spanish, Classical, Bachata, Popular music, Vallenato, Reggaeton, and Ballads in Spanish. These genres were chosen given the experiment cultural context in Southwestern Colombia. The 21 tracks were played a single time, in a random manner, to each subject.

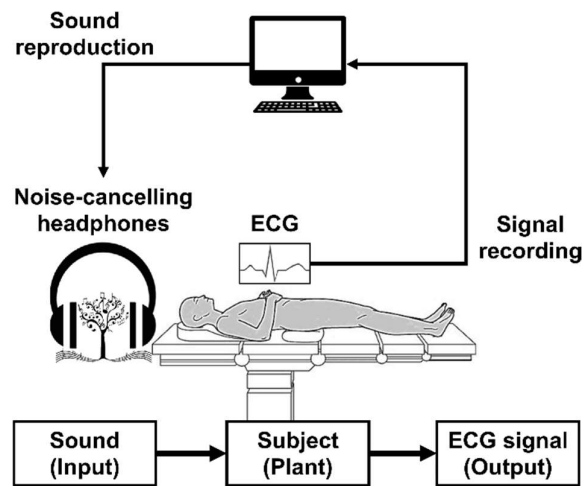


Fig. 1. Diagram of experimental procedure

2.2 Signal processing

Captured signals ($n=483$) were conditioned by applying a third-order one-dimensional median filter and after that, filtered signals were subtracted from the original signals to perform baseline wander correction [29]. A data augmentation process was then carried out using wavelet-based shrinkage filtering [30]. Following the method implemented in [30], three mother wavelets were used to filter the signals - Daubechies 4 (db4), Daubechies 6 (db6), and Symlets 8 (sym8). The outcome of this process was 1932 signals. Signals were then segmented by means of a Hamming Window; each signal was divided into three segments of six seconds each, with 50% overlapping [31]. After that, feature extraction was conducted, where maximum, mean, variance, skewness, kurtosis, energy, entropy, and Katz and Higuchi ($k=3$) fractal dimensions

of the signals were extracted; these features were computed directly from the signals. Feature extraction was carried out in Matlab software [32].

2.3 Classification

Having extracted the features, classification processes were implemented using six artificial intelligence algorithms: k-nearest neighbors (kNN), lightGBM, neural network, random forest, and XGBoost (Table 1). These were designed to discriminate between the binary perception responses: Positive and Negative.

Table 1. Configuration parameters of the implemented artificial intelligence algorithms

Classification algorithms	Configuration
k-nearest neighbors	Number of neighbors: 20, metric: Manhattan, weight: distance
LightGBM	Max bin=600, learning rate=0.001, number of leaves=12, boosting=gbd, number of iterations=7000
Neural network	Multi-layer perceptron with backpropagation. Neurons in hidden layers: 90, activation: ReLu, regularization: alpha=0.05, solver: Adam
Random forest	Number of trees: 45
XGBoost	Learning rate=0.3, max depth=5, min child weight=1, gamma=0, alpha=0.1, subsample=1, colsample by tree=1, objective=multi:softprob

2.4 Data, training, test and evaluation of models

Regarding the dataset generated, the rate of positive and negative sentiments were 53% and 47% respectively for music, and 70% and 30% respectively for nature sounds. The ten folds cross-validation method was utilized for training the algorithms. Performance was evaluated using the area under the ROC curve (AUC), accuracy, precision, recall, and F1-score [33]. These metrics allow evaluation of classifier performance working on imbalanced datasets.

3 Results

After training the selected algorithms, performance was evaluated through the AUC, accuracy, precision, recall, and F1-score metrics. The classification evaluations are shown according to the type of stimuli considered - music (**Fig. 2.**) and nature sounds (**Fig. 3.**). The outcomes showed accuracy up to 0.86 and AUC 0.91 for music stimuli (**Fig. 2.**) and both accuracy and AUC up to 0.92 for nature sounds (**Fig. 3.**).

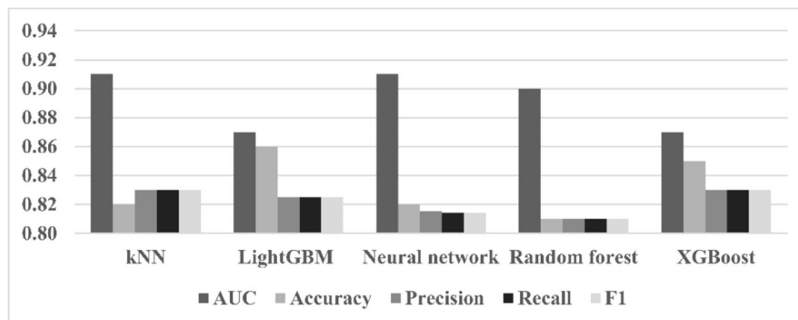


Fig. 2. AUC, accuracy, precision, recall, and F1-score for music stimuli

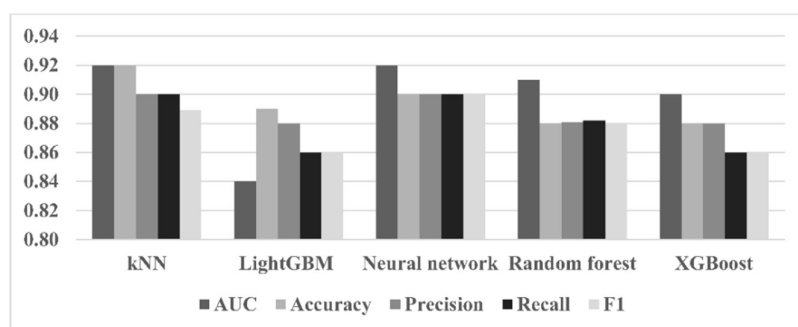


Fig. 3. AUC, accuracy, precision, recall, and F1-score for nature sound stimuli

4 Discussion

From the outcomes it was observed that a single scale of emotion evaluation with a particular ECG signal lead allowed the selected machine learning algorithms to discriminate valence judgments elicited by music and nature sounds. With music stimuli, the best AUC performance was achieved with the kNN and neural network algorithms, both attaining 0.91, and accuracy was greatest using lightGBM, with an accuracy of 0.86. The best performance with nature sound stimuli was with kNN, with 0.92 AUC and accuracy, while neural network also achieved an AUC of 0.92. In general, classification of emotions elicited by nature sounds was better than with music (Fig. 2. and 3), although this could be influenced by the unbalanced responses to nature sounds. All considered evaluation metrics confirm this higher performance. This suggests that subjects responded more uniformly to nature sounds than to music. This behavior may arise from the fact that people generally are exposed to the same common nature sounds, provoking more or less the same responses. People are exposed to music however in many different ways in their daily lives. Moreover, personal prefer-

ences and tastes have a great influence, producing more variability in the responses between subjects.

Discriminating between relaxed and excited states elicited by a multimedia exposure that incorporated relaxing and exciting music, 90% accuracy was achieved using GSR, ECG, electrooculogram, EEG, and photoplethysmography [25]. Another study, classifying three responses to Persian music - happy, peaceful, sad - attained an emotion classification accuracy of 90% [34]. While the above studies mainly classified responses to music associated with pre-anticipated emotions or emotions defined previously, in the present work, the emotional responses of subjects to a range of stimuli was considered, rendering remarkable the high performance in accuracy obtained. Even more remarkable given that the perception of each subject plays an important role in the classification.

Another study in the field of HCI classified four emotional states (stability, relaxation, tension, and excitement) from EEG data with a performance of 86% [22]. Similarly, other work classified music-evoked emotions from EEG data, in which 67% for valence was the best performance [26]. Meanwhile, in [18], for positive and negative valence, high and low arousal, and four types of emotion (joy, tension, sadness, and peacefulness) correct classification rates of 83, 73, and 62% were achieved, respectively. Our results correspond closely to the 83% valence classification. Differences were seen as subjects presented their perception report for each stimulus individually while in [18] subjects reported their perception after listening to longer sessions of several minutes in which several stimuli were presented. Additionally, our results suggest that with only 12 seconds it is possible to produce an emotional response in subjects. The comparative studies above considered EEG signals while in our experimentation, emotion recognition was performed using ECG signals.

New HCI proposals in physiological computing suggest the use of such physiological signals as EEG, ECG or electrodermal activity to automatically generate music using machine learning techniques. Our system represents an approximation to this aim [23]. Moreover, this type of system could be applied in the field of education where background music is expected to improve learning processes [10]. Applications in music therapy, stress management, the practice of yoga or mindfulness might also be implemented. New studies may even address the support of therapies to manage depression using implicit interaction (measuring ECG signals) in HCI systems based on music stimuli. A great advantage in systems that use ECG is its simplicity for measurements in comparison with physiological signals such as EEG. For very lite applications ECG could be measured using only two electrodes [35] [36].

5 Conclusions

A system based on machine learning techniques and ECG signals was designed, to recognize emotional responses to music and nature sounds, measured by a single scale of perception. Subjects were found to give more uniform responses to nature sounds than to music. The kNN algorithm gave the best mean performance. It was possible to recognize the distinctive perception of subjects using binary evaluation and ECG

signals. These outcomes encourage new research and applications that take account of the individual perception of each subject in such a way that an HCI system based on music perception could operate in a more personalized way with more accuracy.

This study showed that with a short period of exposure to a stimulus it is possible to elicit emotions in listeners. A recommendation for future work however concerns the exploration of longer stimuli; another recommendation involves the emotion evaluation scale, initially suggesting the inclusion of a neutral possibility in addition to negative and positive perception; finally, studies using more subjects are recommended, leading to more robust systems with a better performance.

Acknowledgments. This work was supported by Universidad del Valle, Universidad del Cauca, and Colciencias (Funding call No. 727 of 2015), Colombia. The sponsors (did not have direct participation) had no role in the study design, data collection and analysis, neither decision to publish, nor preparation of the manuscript. We are especially grateful to Colin McLachlan for suggestions relating to the English text.

Conflict of interest statement. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be interpreted as a potential conflict of interest.

References

1. Yaribeygi, H., Panahi, Y., Sahraei, H., Johnston, T.P., Sahebkar, A.: The impact of stress on body function: A review. *EXCLI J.* 16, 1057–1072 (2017). <https://doi.org/10.17179/excli2017-480>.
2. Goldfarb, E. V: Enhancing memory with stress: Progress, challenges, and opportunities. *Brain Cogn.* 133, 94–105 (2019). <https://doi.org/10.1016/j.bandc.2018.11.009>.
3. Ramasamy, M., Varadan, V.: Study of emotion-based neurocardiology through wearable systems. Presented at the (2016). <https://doi.org/10.1117/12.2219525>.
4. Umbrello, M., Sorrenti, T., Mistraletti, G., Formenti, P., Chiumello, D., Terzoni, S.: Music therapy reduces stress and anxiety in critically ill patients: a systematic review of randomized clinical trials. *Minerva Anesthesiol.* 85, 886–898 (2019). <https://doi.org/10.23736/S0375-9393.19.13526-2>.
5. de Witte, M., Spruit, A., van Hooren, S., Moonen, X., Stams, G.-J.: Effects of music interventions on stress-related outcomes: a systematic review and two meta-analyses. *Health Psychol. Rev.* 14, 294–324 (2020). <https://doi.org/10.1080/17437199.2019.1627897>.
6. Kacem, I., Kahloul, M., El Arem, S., Ayachi, S., Hafsia, M., Maoua, M., Ben Othmane, M., El Maalel, O., Hmida, W., Bouallague, O., Ben Abdessalem, K., Naija, W., Mrizek, N.: Effects of music therapy on occupational stress and burn-out risk of operating room staff. *Libyan J. Med.* 15, 1768024 (2020). <https://doi.org/10.1080/19932820.2020.1768024>.
7. Chanwimalueang, T., Aufegger, L., Adjei, T., Wasley, D., Cruder, C., Mandic, D.P., Williamon, A.: Stage call: Cardiovascular reactivity to audition stress in musicians. *PLoS One.* 12, (2017). <https://doi.org/10.1371/journal.pone.0176023>.

8. Delmastro, F., Di Martino, F., Dolciotti, C.: Physiological Impact of Vibro-Acoustic Therapy on Stress and Emotions through Wearable Sensors. Presented at the (2018). <https://doi.org/10.1109/PERCOMW.2018.8480170>.
9. Jain, J.K., Maheshwari, R.: Effect of indian classical music and pop music on heart rate variability: A comparative study. *Indian J. Community Heal.* 31, 556–560 (2019).
10. Artifice, A., Ferreira, F., Marcelino-Jesus, E., Sarraipa, J., Jardim-Gonçalves, R.: Student's attention improvement supported by physiological measurements analysis, https://www.scopus.com/inward/record.uri?eid=2-s2.0-85018160066&doi=10.1007%2F978-3-319-56077-9_8&partnerID=40&md5=6af68be745d8e86f78af8f3140e72c87, (2017). https://doi.org/10.1007/978-3-319-56077-9_8.
11. Das, N., Chakraborty, M.: Effects of Indian Classical Music on Heart Rate Variability. Presented at the (2019). <https://doi.org/10.1109/I2CT45611.2019.9033756>.
12. Cerruto, G., Mainardi, L., Koelsch, S., Orini, M.: The periodic repolarization dynamics index identifies changes in ventricular repolarization oscillations associated with music-induced emotions. Presented at the (2017). <https://doi.org/10.22489/CinC.2017.259-372>.
13. Orini, M., Al-Amodi, F., Koelsch, S., Bailón, R.: The Effect of Emotional Valence on Ventricular Repolarization Dynamics Is Mediated by Heart Rate Variability: A Study of QT Variability and Music-Induced Emotions. *Front. Physiol.* 10, (2019). <https://doi.org/10.3389/fphys.2019.01465>.
14. Goshvarpour, A., Abbasi, A., Goshvarpour, A.: Evaluating autonomic parameters: The role of sleep duration in emotional responses to music. *Iran. J. Psychiatry.* 11, 59–63 (2016).
15. Anuharshini, K., Sivaranjani, M., Sowmiya, M., Mahesh, V., Geethanjali, B.: Analyzing the Music Perception Based on Physiological Signals. Presented at the (2019). <https://doi.org/10.1109/ICACCS.2019.8728546>.
16. Najumnissa, D., Alagumariappan, P., Bakiya, A., Ali, M.S.: Analysis on the effect of ECG signals while listening to different genres of music. Presented at the (2019). <https://doi.org/10.1109/ICACCP.2019.8882925>.
17. Goshvarpour, A., Abbasi, A., Goshvarpour, A., Daneshvar, S.: Fusion framework for emotional electrocardiogram and galvanic skin response recognition: Applying wavelet transform. *Iran. J. Med. Phys.* 13, 163–173 (2016). <https://doi.org/10.22038/ijmp.2016.7960>.
18. Hsu, Y.-L., Wang, J.-S., Chiang, W.-C., Hung, C.-H.: Automatic ECG-Based Emotion Recognition in Music Listening. *IEEE Trans. Affect. Comput.* 11, 85–99 (2020). <https://doi.org/10.1109/TAFFC.2017.2781732>.
19. Abhirami, P.H., Saba, E., Mathew, R., Jacob Sebastian, E., Abraham, C.M.: Implementation of convolutional neural network to realize a real time emotion based music player. *Int. J. Recent Technol. Eng.* 8, 164–171 (2019). <https://doi.org/10.35940/ijrte.B1027.0982S1119>.
20. Ju, Q., Chalon, R., Derrode, S.: Assisted music score reading using fixed-gaze head movement: Empirical experiment and design implications. *Proc. ACM Human-Computer Interact.* 3, (2019).
21. Cincuegrani, S.M., Jordà, S., Våljamäe, A.: Physiopucks: Increasing user motivation by combining tangible and implicit physiological interaction. *ACM Trans. Comput. Interact.* 23, (2016). <https://doi.org/10.1145/2838732>.
22. Kim, T.-Y., Ko, H., Kim, S.-H.: Data Analysis for Emotion Classification Based on Bio-Information in Self-Driving Vehicles. *J. Adv. Transp.* 2020, (2020). <https://doi.org/10.1155/2020/8167295>.

23. Kosunen, I., Väljamäe, A.: Designing symbiotic composing. *Acoust. Sci. Technol.* 41, 322–325 (2020). <https://doi.org/10.1250/ast.41.322>.
24. Goshvarpour, A., Abbasi, A., Goshvarpour, A.: An accurate emotion recognition system using ECG and GSR signals and matching pursuit method. *Biomed. J.* 40, 355–368 (2017). <https://doi.org/10.1016/j.bj.2017.11.001>.
25. Anderson, A., Hsiao, T., Metsis, V.: Classification of emotional arousal during multimedia exposure. Presented at the (2017). <https://doi.org/10.1145/3056540.3064956>.
26. Bo, H., Ma, L., Liu, Q., Xu, R., Li, H.: Music-evoked emotion recognition based on cognitive principles inspired EEG temporal and spectral features. *Int. J. Mach. Learn. Cybern.* 10, 2439–2448 (2019). <https://doi.org/10.1007/s13042-018-0880-z>.
27. Al-Galal, S.A.Y., Alshaikhli, I.F.T., Rahman, A.W.B.A.: Automatic emotion recognition based on EEG and ECG signals while listening to quranic recitation compared with listening to music. Presented at the (2017). <https://doi.org/10.1109/ICT4M.2016.55>.
28. OpenBCI: Open Source Biosensing Tools (EEG, EMG, EKG, and more).
29. Clifford, G.D., Azuaje, F., McSharry, P.: *Advanced Methods And Tools for ECG Data Analysis*. Artech House, Inc., Norwood, MA, USA (2006).
30. Wang, C., Yang, S., Tang, X., Li, B.: A 12-Lead ECG Arrhythmia Classification Method Based on 1D Densely Connected CNN, https://www.scopus.com/inward/record.uri?eid=2-s2.0-85075760761&doi=10.1007%2F978-3-030-33327-0_9&partnerID=40&md5=ae9cf40beff62d244788a1b7ea39513d, (2019). https://doi.org/10.1007/978-3-030-33327-0_9.
31. Patané, A., Kwiatkowska, M.: Calibrating the classifier: Siamese neural network architecture for end-to-end arousal recognition from ECG. In: *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2019). https://doi.org/10.1007/978-3-030-13709-0_1.
32. MathWorks: Makers of MATLAB and Simulink, <https://www.mathworks.com/>, last accessed 2019/04/15.
33. Williams, A.M., Liu, Y., Regner, K.R., Jotterand, F., Liu, P., Liang, M.: Artificial intelligence, physiological genomics, and precision medicine. *Physiol. Genomics.* 50, 237–243 (2018). <https://doi.org/10.1152/physiolgenomics.00119.2017>.
34. Khodabakhshian, B., Moharreri, S., Parvaneh, S.: Evaluating the Effects of Traditional Persian Music on Nonlinear Parameters of HRV. Presented at the (2019). <https://doi.org/10.23919/CinC49843.2019.9005806>.
35. Hsieh, H.-Y., Luo, C.-H., Ye, J.-W., Tai, C.-C.: Two-electrode-pair electrocardiogram with no common ground between two pairs. *Rev. Sci. Instrum.* 90, 114703 (2019). <https://doi.org/10.1063/1.5016939>.
36. Raj, P.S., Hatzinakos, D.: Feasibility of single-arm single-lead ECG biometrics. In: *2014 22nd European Signal Processing Conference (EUSIPCO)*. pp. 2525–2529 (2014).