International Journal on Advanced Science Engineering Information Technology

Vol.12 (2022) No. 4
ISSN: 2088-5334

Heart Response to Harmonic Music Interval Stimuli Via Deep Learning Structures

Ennio Idrobo-Ávila [a,*], Humberto Loaiza-Correa [a], Flavio Muñoz-Bolaños [b], Leon van Noorden [c], Rubiel Vargas-Cañas [d]

[a] PSI - Intelligent Systems and Perception, Universidad del Valle, Calle 13 No 100 – 00, building 354, office 1001, Cali, 76001, Colombia
[b] CIFIEX - Experimental Physiological Sciences, Universidad del Cauca, Carrera 5 # 13N – 36, Popayán, 190002, Colombia
[c] IPEM - Institute for Systematic Musicology, Ghent University, Miriam Makebaplein 1 B-9000, Ghent, Belgium
[d] SIDICO – Dynamic Systems Instrumentation and Control, Universidad del Cauca, Carrera 3 Nº 3N-100, Popayán, 190002, Colombia

Corresponding author: [*]ennio.idrobo@correounivalle.edu.co

***Abstract*— The effect of music on the heart is reflected in variables such as heart rate and electrocardiographic (ECG) signals. ECG is a record of heart electrical activity and is a useful tool in diagnosing various cardiopathies. Artificial intelligence techniques have recently been implemented to analyze ECG and RR-interval data and are used thus in the present study to examine the influence on the heart of harmonic musical intervals and colored noise. Harmonic intervals were chosen because of their emotional response, while noise has been linked to positive responses such as improved sleep quality. A deep learning system was implemented, employing the ResNet-18 and GoogLeNet pre-trained networks to discriminate 31 different classes of ECG and RR-interval responses to the sound stimuli. Following an exploratory approach, deep learning was selected as an alternative to traditional analysis with the expectation that it could be incorporated into future music perception research. Classification revealed the ability of the implemented system to demonstrate heart response to the stimuli. ECG signals performed best, with 97% accuracy and Matthew's coefficient of 0.97, while RR-interval achieved a 93% accuracy and Matthews coefficient of 0.93, suggesting that the considered stimuli of harmonic musical intervals and noise produced different responses in the heart. Moreover, the Matthews coefficient values above 0.7 and close to 1 imply a correlation between the two types of stimuli and the heart response, as measured by ECG and RR-interval signals.**

***Keywords*— Electrocardiographic signals; GoogLeNet; music; noise; ResNet-18; transfer learning.**

Manuscript received 13 Aug. 2021; revised 24 Sep. 2021; accepted 10 Oct. 2021. Date of publication 31 Aug. 2022.

I. INTRODUCTION

Heart health and the heart's response to medicines or stimuli are often evaluated using the electrocardiogram (ECG) [1]. A clinical record of the heart's electrical activity, ECG is usually recorded as an analogue signal [2] and employed in diagnosing cardiopathies. ECG has been used to discover the effect of music on the heart, using measurements of heart rate (HR) and heart rate variability (HRV) [2] [3]. Responses to music can be both psychological (e.g., subjective perception) and physiological (HRV parameters) [4]–[6]. Changes in HRV parameters have shown an increase in parasympathetic tone [7] and a reduction in sympathetic tone [8]. HRV has revealed that music can be used to support oncological therapy [8], [9]. Both health [10] and engineering [11] areas have researched responses of the heart and HRV to music. Research on music and its effect on the human body indicates multiple benefits, including the influence of music on infant brain activity [12] and the effects of music on HRV and pain relief in elderly patients with total knee replacement [13]. More recent studies indicate the influence of music on HRV parameters [14]. HRV has played a role in classifying emotions evoked by music [15], differentiating between distinct types of music [5], [16], and states of exposure to music and non-music [17], [18].

Concerning the use of artificial intelligence techniques, machine learning algorithms have proved useful tools for ECG data, e.g., neural networks [16] and support vector machines [19]. Meanwhile, deep learning deals better with large amounts of data [20] and has been used to analyze physiological signals such as ECG data. Thereby avoiding the feature extraction process of traditional techniques [20], and advancing the detection of arrhythmias [21] and classification of heart diseases [22]. Meanwhile, as a training technique, transfer learning is able to take advantage of knowledge embedded in algorithms pre-trained with deep learning techniques and reuse it in a particular application different from that of the training. Transfer learning has been used to implement deep learning algorithms (overcoming the need for a large volume of data in the training stage [23], [24], classifying normal and abnormal ECG signals [25], and ECG arrhythmias [23]. Few applications of deep learning techniques on HRV have been found to date. Elsewhere, spectral representation of HRV has been used to predict mental stress [26] and intervals between R-peaks (RR-intervals - RRI) to detect congestive heart failure using multi-input deep learning networks [27], LSTM-based deep networks [28], and convolutional neural networks [29].

Where do the problems lie? Most research classifies emotions elicited by music through recording physiological variables [30], [31], considering music with a broad brush as a complete element, without examining parts [5] [32], such as harmonic intervals. Pentatonic music, for example, produced observable changes in some HRV parameters [33], producing a parasympathetic response in infants. Studies have also pointed to the effects of monochord music on HRV [34]; the influence of tempo on signals of EEG [35], HRV [36], [37], and baroreflex changes [38]; the production of changes in psychophysiological variables [39] and changes in specific elements of physiological signals such as ECG [40] by rhythm, frequency characteristics, and music-excerpts; and the influence of sound quality on HRV [41]. When research considers music from a broad viewpoint, it is only possible to observe general mechanisms of how music affects the variables of interest [18], [32]. The door to new research on this topic is very much open.

Considering that previous research has observed the effects of music - and some particular elements it comprises - on the heart's activity, the aim of this study is, through an exploratory approach, to show changes in the cardiac response to sound stimuli using deep learning techniques. Making use of such a deep learning system, the goal of this research is to classify, via ECG and RRI (tachogram), the heart response to stimuli of harmonic musical intervals (HMI) and colored noise, harmonic intervals being two musical sounds played at the same time [42]. Harmonic intervals were chosen as they form a fundamental part of music and evoke an emotional response [43], while the use of noise has been related to positive responses, such as sleep improvement [44] and attention-deficit/hyperactivity disorder [45]. ECG and RRI, categorized according to sound stimuli, provide relevant information and could be used to determine whether or not the heart is influenced by the harmonic musical interval's size and the noise's frequency content.

Following a pre-processing procedure, the signals were transformed into a bidimensional spectrum using continuous wavelet transform [46] for presentation to the deep learning system, developed to discriminate 31 and later 30 different ECG and RRI classes as responses to different types of HMI and noise. The system used the ResNet-18 [47] and GoogLeNet [48] pre-trained convolutional neural networks; deep learning allows for dealing with several dependencies at once, such as all classes of stimuli in the study. These pretrained networks were selected due to the high accuracy performance, a low number of operations, and density of parameters to be configured compared to other models [49]. The system was trained following a transfer learning approach and an interpatient scheme [50], with 70% of the data for training and 30% for testing.

System performance was evaluated using accuracy, the Cohen's kappa coefficient (κ), and the Matthews correlation coefficient (MCC). However, a further metric of classification cost was also proposed to take account of the nature of the classes (harmonic intervals and noise), the quality of consonance and dissonance (harmonic intervals), and the frequency content (high and low octaves in harmonic intervals, and high, low, or all frequencies in noise). The cost was determined assuming that each stimulus type would influence heart response (ECG and RRI) differently. By similar stimuli producing similar heart responses; incorrect classifications of responses to stimuli that are similar in consonance or frequency content will have a lower cost than those responses to completely different stimuli – harmonic intervals and noise, consonant and dissonant intervals, or noise with high and low frequencies. The research hypothesis is that the stimuli will produce different responses in the heart, revealed using deep learning, contributing to understanding the hearing process.

The rest of this manuscript is organized as follows: Section 2 presents the materials and methods, which include the experimental procedure, auditory stimuli, data processing, classification system, and evaluation. The results and their discussion are reported in Section 3. Finally, Section 4 concludes this document.

II. Materials and Method

A. *Experimental procedure*

During the experimental stage, 17 males and 9 females, with a mean age of 25.3 years (SD=7.1), were voluntarily enlisted to participate in the experimental phase. After an explanation of the experiment, participants signed the consent form. The Internal Ethical Committee of Universidad del Cauca approved the experimental procedures and the research was carried out in accordance with the approved protocol and Declaration of Helsinki. The subjects, while remaining in a supine position, listened to 30 sound stimuli of 10 seconds each: 6 different types of noises and 24 harmonic musical interval sounds in octaves 2 and 4 (Fig. 1 and Table 1).

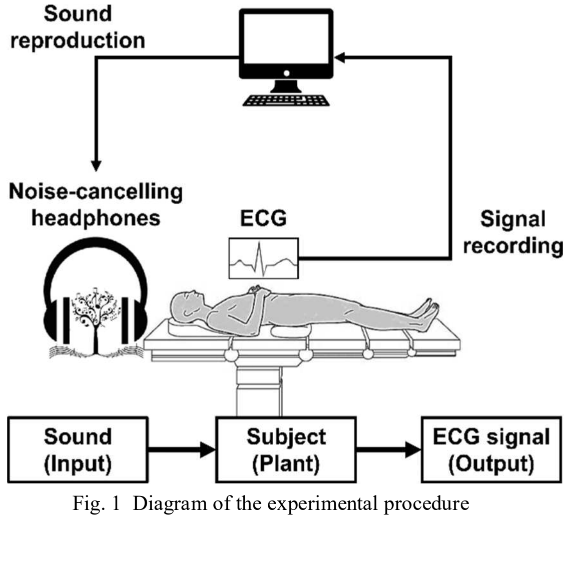

Fig. 1 Diagram of the experimental procedure

TABLE I
STIMULI IN THE EXPERIMENTAL PHASE

Type of stimulus	Description
Silence	Noise-cancelling headphone output
Noise	Blue
	Brown
	Grey
	Pink
	Purple
	White
Harmonic interval	Minor second (2m)
(Octaves 2 and 4)	Major second (2M)
	Minor third (3m)
	Major third (3M)
	Perfect fourth (4)
	Augmented fourth (4aug)
	Perfect fifth (5)
	Minor sixth (6m)
	Major sixth (6M)
	Minor seventh (7m)
	Major seventh (7M)
	Octave (8)

In order to reduce carryover effects [51], the sounds were separated by a period of 15 seconds of silence, and they were presented in random order using noise-canceling headphones. During the experiment, ECG signals from lead II were recorded. Lead II is an electrode configuration that reveals the electrical activity from the perspective of the inferior wall of the heart. In this configuration, the right arm (negative pole) is connected with the left leg (positive pole) [52].

The color of noise is assigned as analogous to colors in the visible spectrum, in which each color has a different power spectrum: the frequency content of noise is expressed as follows [53]:

f^2 for purple noise
f for blue noise
1 for white noise
1/f for pink noise
$1/f^2$ for brown noise

The frequency content of grey noise is similar to that of pink noise, but the loudness in the whole spectrum is normalized [54].

B. Data Processing

The signal processing step comprised six stages: pre-processing, dataset augmentation, signal cutting, signal windowing, continuous wavelet transform representation (CWT), and classification. In the pre-processing stage, the signal trend was removed by subtracting the output of a third-order one-dimensional median filter from the original signal. In this case, the baseline is removed to reduce the effect of issues with electrodes, movement of the subject, and breathing [55]. In the dataset augmentation stage, the wavelet-based shrinkage filtering method was implemented [56], in which two mother wavelets were included for filtering the signals: Daubechies 4 (db4) and Daubechies 6 (db6); for each signal, two new signals were generated. Data augmentation is carried out to improve performance and reduce overfitting in machine/deep learning algorithms [57]. Following signal augmentation, the ECG signals were cut to a duration of 10 seconds to maintain the order of presentation of the stimuli.

After the signal cutting, a Hamming windowing process of five seconds was performed on the ECG signals. This window duration was chosen since, at a "normal" heart rate, it is possible to find between 4 and 8 heartbeats [58]. Each signal was split into segments, and the cuts were made from 0 to 5, 2.5 to 7.5, and 5 to 10 seconds. This process was performed to inspect the signals in more detail; thus, they were analyzed in segments instead of taking the signals as unique elements. After signal windowing, the R-peaks were segmented; commonly, the R-peak is the dominant peak in ECG signals, and its duration is about 30 ms in lead II [1]. The Pan-Tompkins algorithm was utilized to segment the R-peaks [59]; undetected peaks were marked manually. The RRI signal was extracted by computing the time difference between R-peaks [60] to measure R-distances in milliseconds [61] [62].

Finally, CWT was obtained from each signal segment in both the ECG and RRI signals. CWT was computed to transform time-domain signals into time-frequency domain signals [63]. At the end of the signal processing stage, there were a total of 7254 instances.

C. Classification

After the computation of CWT, a deep learning classification system was developed using the pre-trained networks ResNet-18 [64] and GoogLeNet [65]. These convolutional neural networks have been trained with more than a million images from the ImageNet database [66]. ResNet-18 has 18, and GoogLeNet has 22 deep layers, and each can classify up to 1000 different image classes [64] [65]. In this application, the CWT output served as input to both pre-trained networks. This CWT process produced a set of images that were adjusted to input from the networks. Thus, images were resized to 224x224x3. After image resizing, the last fully connected layer and the final classification layer of the networks were replaced. The fully connected layer was substituted with a new fully-connected layer with 31 outputs, and the classification layer was superseded with a new one with different class labels. The system was trained following an interpatient scheme [50], with 70% of the data for training and 30% for testing. In this classification task, 31 classes with 234 instances each were considered. Classes were associated with each stimulus type, and instances were related to the number of elements in each class. The first 30 classes corresponded to stimuli described in the experimental procedure, but a class corresponding to silence or without stimulus was also added.

D. Evaluation

Finally, to evaluate the classification system, accuracy, the Cohen's kappa coefficient (κ), and Matthews correlation coefficient (MCC) were computed; MCC because it allows assessment of the classification performance even with unbalanced datasets [67]. MCC represents a correlation between observation and prediction in classification tasks and can take values between -1 and 1 [68]. A value of 1 indicates a wholly correct prediction, -1 represents a wholly incorrect prediction, while 0 suggests an almost random prediction.

In addition to the abovementioned metrics, a cost metric was developed. The cost metric can take values between 0 and 1, where 1 represents the worst-case classification, i.e., the highest cost to pay because of an extremely poor classification, and 0 represents the lowest cost due to an ideal classification. Thus, values near to 0 are desired. The cost metric is based on a cost matrix defined by a set of rules. The first establishes that the maximum cost, equal to 1, should be divided between the complete range of classification possibilities in each instance. The second defines that minimum cost, equal to 0, should be assigned to well-classified instances. The third rule divides the total cost to pay between general classes; thus, a cost of 0.5 is assigned to all harmonic interval classes and all noise classes. As a result, the highest cost will be paid if a harmonic interval class element is classified as noise, or a noise class element is classified as a harmonic interval. Here, it is crucial to bear in mind that the Silence class is considered within the harmonic interval classes because the Silence and 8_L classes behave similarly in the RRI classification.

The cost of general classes is divided according to the Gompertz function [69] (Equation 1).

$$SC_i = \frac{e^{-\beta e^{-\gamma i}} - e^{-\beta e^{-\gamma}}}{\#GeneralClasses * \sum_{j=1}^{n} e^{-\beta e^{-\gamma i}} - e^{-\beta e^{-\gamma}}},\ i = [1, n] \quad (1)$$

In this application β and γ took the values 4 and 0.1, respectively. The cost is incremented from minimum to maximum value and assigned to each class of the general class. For instance, in the case of harmonic interval classes, costs take values between 0.000124 and 0.057102 and are attributed according to the degree of consonance. Ordering of the stimuli by consonance was based on the classification of harmonic intervals regarding tonal aspects, following proportional theory [70], and is defined (Table 2) from most to least consonant.

TABLE II
ORDER OF CONSONANCE IN THE HARMONIC INTERVALS

Order of consonance	**Interval**	**Order of consonance**	**Interval**
1	Silence	14	3m_H
2	8_L	15	6m_H
3	5_L	16	7m_L
4	4_L	17	2M_L
5	6M_L	18	7M_L
6	3M_L	19	2m_L
7	3m_L	20	4aug_L
8	6m_L	21	7m_H
9	8_H	22	2M_H
10	5_H	23	7M_H
11	4_H	24	2m_H
12	6M_H	25	4aug_H
13	3M_H		

Stimuli with similar characteristics will have the lowest costs, and different characteristics will have the highest costs. Thus, in a consonant class such as 4_L, other consonant classes such as 8_L, 5_L, 6M_L, and 3M_L will have the lowest cost, and dissonant classes such as 2M_H, 7M_H, 2m_H, and 4aug_H the highest costs. Once the cost matrix is established, it is multiplied by the confusion matrix to determine the classification cost matrix. All classification cost matrix values are then summed to obtain the total cost. Finally, to standardize the value between 0 and 1, the total cost is divided by the maximum possible cost. Maximum possible cost is determined considering the case in each instance where classes are classified as the classes with the highest cost in the cost matrix.

III. Result and Discussion

A. Continuous Wavelet Transform

As a result of CWT, a 2D representation was obtained for each segment of ECG and RRI signal captured during the presentation of the stimuli to the subjects (Fig. 2 and 3). Each representation corresponds to five seconds (x-axis), 135 scales (y-axis), and CWT coefficient absolute values changes. Time progresses from left to right in the images, while frequency increases from bottom to top. Changes, or differences, in CWT representations as a response to stimuli are difficult to observe with ECG (Fig. 2), but with RRI (Fig. 3), these changes can be more readily seen, the clearest case being the RRI response to Minor Second, in which a prominent peak is observed.

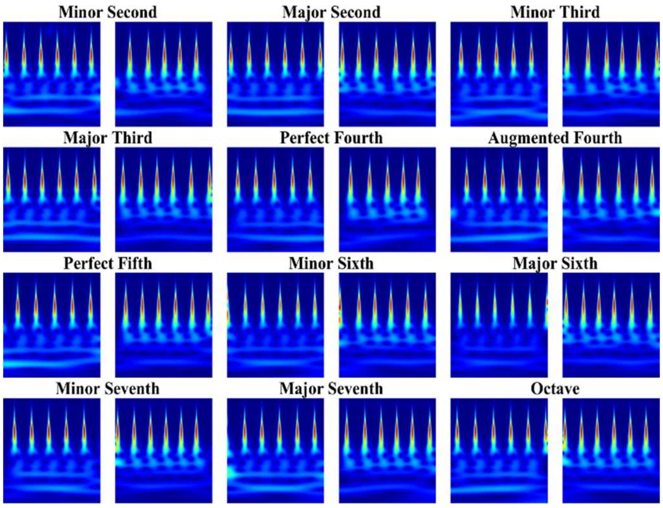

Fig. 2 Continuous wavelet transform of ECG signals for the lowest octave in two different subjects (blue = minimum value; red = maximum value)

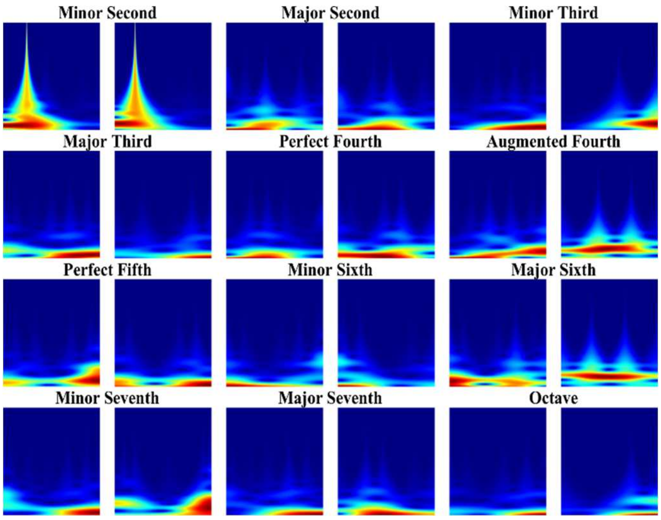

Fig. 3 Continuous wavelet transform of RRI signals for the lowest octave in two different subjects (blue = minimum value; red = maximum value)

B. Training Process

Following training on the ECG signals, the ResNet-18 network produced a testing accuracy of 0.89 and κ and MCC of 0.89, while the cost was 0.03; with GoogLeNet, accuracy and κ MCC were 0.97 and 0.97, with a cost of 0.01 (Table 3). Classification of RRI was also carried out. With the ResNet-18 network a testing accuracy of 0.73 and, κ and MCC of 0.72 was obtained, with cost equal to 0.07; while GoogLeNet gave an accuracy of 0.71, and κ and MCC of 0.70, with a cost of 0.07 (Table 3).

On observing the RRI classification, it is remarkable that the system is unable to differentiate between the silence and 8_L classes. Then, bearing in mind that the best classification outcomes were achieved using GoogLeNet, this network was used to perform a new RRI classification. This time, however, the Silence class was removed (i.e., 30 classes remained). In this case, a testing accuracy of 0.93 and κ and MCC of 0.93 were obtained, with a cost of 0.03 (Table 3).

TABLE III
PERFORMANCE OF THE CLASSIFICATION PROCESS OF ECG AND RRI SIGNALS WITH THE RESNET-18 AND GOOGLENET NETWORKS

Network	ECG			RRI		
	Accuracy	Kappa/Matthews coefficient	Cost	Accuracy	Kappa/Matthews coefficient	Cost
31 classes (including silence)						
ResNet-18	0.89	0.89	0.03	0.73	0.72	0.07
GoogLeNet	0.97	0.97	0.01	0.71	0.70	0.07
30 classes (excluding silence)						
GoogLeNet	---	---	---	0.93	0.93	0.03

C. Discussion

In this manuscript, the ECG and RRI signals of the study subjects were classified according to the presented stimuli of harmonic intervals and noise. The best classifier performance with the test set provided an accuracy of 97%, a κ and Matthew's coefficient of 0.97, and a cost of 0.01. These outcomes suggest a correlation between the selected stimuli and the ECG and RRI measured. In light of the experimental design, each stimulus might be considered to produce a different response in the heart, manifested through RRI and electrical activity registered in the ECG signal. In any case, initial indications exist of a direct effect of selected sounds on heart behavior; at this early stage, it is neither clear what the impact of this is nor its significance.

The deep learning system implemented could differentiate all the classes provided. Furthermore, the pre-trained networks selected, ResNet-18 and GoogLeNet, were also essential in classifying the CWT representations. The general performance was better with GoogLeNet than with ResNet-18; GoogLeNet was similarly reported to obtain better accuracy than ResNet-18 [49]. The cost metric, meanwhile, was seen to provide a coherent description of the performance of the systems and behaved consistently concerning the accuracy, κ, and Matthew's coefficient.

CWT was employed as a time-frequency representation of the ECG and RRI signals. These signals are measurements of the electrical activity of the heart and time variability between heartbeats, while CWT represents the general behavior of the heart. Given that it was possible to infer that the presented stimuli affected the behavior or state of the heart, this may well affirm that variations can be registered in ultra-short RRI and ECG recordings of only five seconds. Such a claim arises from the segmenting of signals in the study into periods of five seconds. Looking at the results, it is further conceivable to conclude that the stimuli used might be affecting subjects at a more central level in addition to affecting the peripheral ear. As such, the outcomes observed might be explained best by theory relating to brain-heart interactions [71], [72].

In the case of RRI signals, the classification system could not differentiate between the silence and 8_L classes. RRI in these conditions may be similar. In other words, more generally, the RRI response to perfect consonance in the lower registers might be similar to that produced under conditions of silence. RRI response to perfect consonance in the higher frequencies, however, was able to be differentiated from the other classes. This was true even on including the classes of silence and 8_L (perfect consonance in the lower register). Considering the results detailed above, heart behavior was discovered to change as the different stimuli were presented to the subjects. However, it was impossible to determine exactly the nature of those changes. For this reason, in terms of future work, it is essential to conduct new analyses to determine what elements in ECG and RRI signals are affected and also how they are changing, and it is hoped to build up an understanding of the perception of more complex parts of music, such as harmonic progressions, from the ground up [73]. By understanding the response of the heart to stimuli, these stimuli could go on to be used to produce a desired and controlled response in the hearts of listeners.

IV. CONCLUSION

The first element to observe from the outcomes is that despite CWT not being commonly used to analyze RRI signals, it proved to be an efficient tool to represent the signals treated in this experiment, ECG and RRI signals. Nevertheless, further research would be useful to study other ways to represent these types of signals. Secondly, the results confirm the hypothesis that the considered stimuli (harmonic musical intervals and noise) produce different responses in the heart. As regards the evaluation of the system, the proposed cost metric presented coherent information about the system performance relative to accuracy, Cohen's kappa coefficient, and Matthew's coefficient. The MCC values above 0.7 and near to 1 suggest that a correlation exists between the stimuli and the heart response measured by ECG and RRI signals. Moreover, bearing in mind the experimental design, a cause-effect relationship might be inferred between the stimuli and the observed effects on the heart. The high performance in the classification system particularly means that the heart presented different responses to particular stimuli such as harmonic musical intervals (two different octaves) and colored noise (six different types). This result presents quantitative evidence that the stimuli used produced changes in heart behavior assessed through ECG and RRI signals. It is necessary to explore other data analysis methods to figure out in detail what these effects are and practically how their advantages may be taken. Although the implemented method does not provide information about the nature of the changes produced in the heart, the study outcomes open doors to new studies to explore the possibility of manipulating or massaging heart behavior using sounds.

Acknowledgment

This work was supported by Colciencias (Call No. 727 of 2015), Colombia. The funders (who did not have direct participation) had no role in the study design, data collection, and analysis, neither deciding to publish nor preparing the manuscript. The authors acknowledge Universidad del Cauca and Universidad del Valle for their support in this research. We are especially grateful to Colin McLachlan for suggestions relating to the English text.

References

[1] A. Capucci, *New Concepts in ECG Interpretation*. 2019.

[2] S. Koelsch and L. Jäncke, "Music and the heart," *Eur. Heart J.*, p. ehv430, 2015.

[3] H.-J. Trappe, "Music and heart: What is verified, what is not, what's new? ," *Kardiologe*, vol. 11, no. 6, pp. 486–496, 2017, doi: 10.1007/s12181-017-0192-7.

[4] K. Kantono, N. Hamid, D. Shepherd, Y. H. T. Lin, S. Skiredj, and B. T. Carr, "Emotional and electrophysiological measures correlate to flavour perception in the presence of music," *Physiol. Behav.*, vol. 199, pp. 154–164, 2019, doi: 10.1016/j.physbeh.2018.11.012.

[5] J. K. Jain and R. Maheshwari, "Effect of indian classical music and pop music on heart rate variability: A comparative study," *Indian J. Community Heal.*, vol. 31, no. 4, pp. 556–560, 2019.

[6] M. Erfanian, A. J. Mitchell, J. Kang, and F. Aletta, "The psychophysiological implications of soundscape: A systematic review of empirical literature and a research agenda," *Int. J. Environ. Res. Public Health*, vol. 16, no. 19, Sep. 2019, doi: 10.3390/ijerph16193533.

[7] A. W. C. Yuen and J. W. Sander, "Can natural ways to stimulate the vagus nerve improve seizure control?," *Epilepsy Behav.*, vol. 67, pp. 105–110, 2017, doi: 10.1016/j.yebeh.2016.10.039.

[8] M. Warth, J. Kessler, T. K. Hillecke, and H. J. Bardenheuer, "Trajectories of Terminally Ill Patients' Cardiovascular Response to Receptive Music Therapy in Palliative Care," *J. Pain Symptom Manage.*, vol. 52, no. 2, pp. 196–204, 2016, doi: 10.1016/j.jpainsymman.2016.01.008.

[9] S. Palma, M. Keilani, T. Hasenoehrl, and R. Crevenna, "Impact of supportive therapy modalities on heart rate variability in cancer patients–a systematic review," *Disabil. Rehabil.*, vol. 42, no. 1, pp. 36–43, Jan. 2020, doi: 10.1080/09638288.2018.1514664.

[10] X.-Y. Zhang, J.-J. Li, H.-T. Lu, W.-J. Teng, and S.-H. Liu, "Positive effects of music therapist's selected auditory stimulation on the autonomic nervous system of patients with disorder of consciousness: a randomized controlled trial," *Neural Regen. Res.*, vol. 16, no. 7, pp. 1266–1272, 2021, doi: 10.4103/1673-5374.301021.

[11] S. Sieciński and P. Kostka, "Influence of Music on HRV Indices Derived from ECG and SCG," *International Scientific Conference Advances in Applied Biomechanics, AAB 2020*, vol. 1223. Springer Science and Business Media Deutschland GmbH, Faculty of Biomedical Engineering, Department of Biosensors and Processing of Biomedical Signals, Silesian University of Technology, Roosevelta 40, Zabrze, 41-800, Poland, pp. 381–389, 2021, doi: 10.1007/978-3-030-52180-6_39.

[12] J. Sa de Almeida *et al.*, "Music enhances structural maturation of emotional processing neural pathways in very preterm infants," *Neuroimage*, vol. 207, 2020, doi: 10.1016/j.neuroimage.2019.116391.

[13] C.-C. Hsu, S.-R. Chen, P.-H. Lee, and P.-C. Lin, "The Effect of Music Listening on Pain, Heart Rate Variability, and Range of Motion in Older Adults After Total Knee Replacement," *Clin. Nurs. Res.*, vol. 28, no. 5, pp. 529–547, 2019, doi: 10.1177/1054773817749108.

[14] T. McPherson, D. Berger, S. Alagapan, and F. Fröhlich, "Active and Passive Rhythmic Music Therapy Interventions Differentially Modulate Sympathetic Autonomic Nervous System Activity," *J. Music Ther.*, vol. 56, no. 3, pp. 240–264, 2019, doi: 10.1093/jmt/thz007.

[15] V. Moaiyed, M. Firoozabadi, and M. Khezri, "Recognition of Music-Induced Emotions Based on Heart-Brain Connectivity," 2018, doi: 10.1109/ICBME.2017.8430259.

[16] D. Najumnissa, P. Alagumariappan, A. Bakiya, and M. S. Ali, "Analysis on the effect of ECG signals while listening to different genres of music," 2019, doi: 10.1109/ICACCP.2019.8882925.

[17] G. Yadu, D. Panigrahi, S. K. Nayak, A. Dey, and K. Pal, "Wavelet Packet Analysis of ECG signals to Understand the Effect of a Motivating Song on Heart of Indian Male Volunteers," in *Expert System Techniques in Biomedical Science Practice*, National Institute of Technology, Rourkela, India: IGI Global, 2018, pp. 168–192.

[18] S. Paul, G. Yadu, S. K. Nayak, A. Dey, and K. Pal, "Recurrence quantification analysis of electrocardiogram signals to recognize the effect of a motivational song on the cardiac electrophysiology," vol. 575. Springer Verlag, Department of Biotechnology and Medical Engineering, National Institute of Technology, Rourkela, 769008, India, pp. 165–172, 2020, doi: 10.1007/978-981-13-8687-9_16.

[19] K. Wang, W. Wen, and G.-Y. Liu, "The autonomic nervous mechanism of music therapy for dental anxiety," 2017, pp. 289–292, doi: 10.1109/ICCWAMTIP.2016.8079858.

[20] A. Mincholé, J. Camps, A. Lyon, and B. Rodríguez, "Machine learning in the electrocardiogram," *J. Electrocardiol.*, 2019, doi: 10.1016/j.jelectrocard.2019.08.008.

[21] S. Parvaneh, J. Rubin, S. Babaeizadeh, and M. Xu-Wilson, "Cardiac arrhythmia detection using deep learning: A review," *J. Electrocardiol.*, 2019, doi: 10.1016/j.jelectrocard.2019.08.004.

[22] T.-H. Nguyen, T.-N. Nguyen, and T.-T. Nguyen, "A deep learning framework for heart disease classification in an IoTs-based system," *Intelligent Systems Reference Library*, vol. 165. Springer Science and Business Media Deutschland GmbH, Department of Industrial Electronic-Biomedical Engineering, Ho Chi Minh City University of Technology and Education, Ho Chi Minh City, Viet Nam, pp. 217–244, 2020, doi: 10.1007/978-3-030-23983-1_9.

[23] M. Salem, S. Taheri, and J. Yuan, "ECG arrhythmia classification using transfer learning from 2-dimensional deep CNN features," in *2018 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2018, pp. 1–4.

[24] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," *Comput. Methods Programs Biomed.*, vol. 161, pp. 1–13, 2018, doi: 10.1016/j.cmpb.2018.04.005.

[25] L. Chen, G. Xu, S. Zhang, J. Kuang, and L. Hao, "Transfer Learning for Electrocardiogram Classification Under Small Dataset BT - Machine Learning and Medical Engineering for Cardiovascular Health and Intravascular Imaging and Computer Assisted Stenting," 2019, pp. 45–54.

[26] J. L. Hagad, K. ichi Fukui, and M. Numao, "Modelling Naturalistic Work Stress Using Spectral HRV Representations and Deep Learning," in *Advances in Intelligent Systems and Computing*, 2020, vol. 1128 AISC, pp. 267–277, doi: 10.1007/978-3-030-39878-1_24.

[27] S.-H. Huang, B.-L. Chuang, Y.-H. Lin, C.-S. Hung, and H.-P. Ma, "A congestive heart failure detection system via multi-input deep learning networks," 2019, doi: 10.1109/GLOBECOM38437.2019.9013460.

[28] L. Wang and X. Zhou, "Detection of congestive heart failure based on LSTM-based deep network via short-term RR intervals," *Sensors (Switzerland)*, vol. 19, no. 7, 2019, doi: 10.3390/s19071502.

[29] Y. Li *et al.*, "Combining Convolutional Neural Network and Distance Distribution Matrix for Identification of Congestive Heart Failure," *IEEE Access*, vol. 6, pp. 39734–39744, 2018, doi: 10.1109/ACCESS.2018.2855420.

[30] M. Merino-Monge, I. M. Gómez-González, J. A. Castro-García, A. J. Molina-Cantero, and R. Quesada-Tabares, "A preliminary study about the music influence on EEG and ECG signals," 2018, pp. 100–106.

[31] Y.-L. Hsu, J.-S. Wang, W.-C. Chiang, and C.-H. Hung, "Automatic ECG-Based Emotion Recognition in Music Listening," *IEEE Trans. Affect. Comput.*, vol. 11, no. 1, pp. 85–99, 2020, doi: 10.1109/TAFFC.2017.2781732.

[32] M. A. Alves, D. M. Garner, J. A. T. do Amaral, F. R. Oliveira, and V. E. Valenti, "The effects of musical auditory stimulation on heart rate autonomic responses to driving: A prospective randomized case-control pilot study," *Complement. Ther. Med.*, vol. 46, pp. 158–164, 2019, doi: 10.1016/j.ctim.2019.08.006.

[33] A. Ranger *et al.*, "Physiological and emotional effects of pentatonic live music played for preterm neonates and their mothers in the Newborn Intensive Care Unit: A randomized controlled trial,"

Complement. Ther. Med., vol. 41, pp. 240–246, 2018, doi: 10.1016/j.ctim.2018.07.009.

[34] C. Gäbel, N. Garrido, J. Koenig, T. K. Hillecke, and M. Warth, "Effects of Monochord Music on Heart Rate Variability and Self-Reports of Relaxation in Healthy Adults," *Complement. Med. Res.*, vol. 24, no. 2, pp. 97–103, 2017, doi: 10.1159/000455133.

[35] M. J. Mollakazemi, D. Biswal, J. Evans, and A. Patwardhan, "Eigen Decomposition of Cardiac Synchronous EEGs for Investigation of Neural Effects of Tempo and Cognition of Songs," 2018, vol. 2018-July, pp. 2402–2405, doi: 10.1109/EMBC.2018.8512806.

[36] J. Kim, C. A. Strohbach, and D. H. Wedell, "Effects of manipulating the tempo of popular songs on behavioral and physiological responses," *Psychol. Music*, vol. 47, no. 3, pp. 392–406, 2019, doi: 10.1177/0305735618754688.

[37] B. Bretherton, J. Deuchars, and W. L. Windsor, "The Effects of Controlled Tempo Manipulations on Cardiovascular Autonomic Function," *Music Sci.*, vol. 2, p. 205920431985828, Jan. 2019, doi: 10.1177/2059204319858281.

[38] D. Biswal, M. J. Mollakazemi, S. Thyagarajan, J. Evans, and A. Patwardhan, "Baroreflex Sensitivity during Listening to Music Computed from Time Domain Sequences and Frequency Domain Transfer Function," 2018, vol. 2018-July, pp. 2776–2779, doi: 10.1109/EMBC.2018.8512861.

[39] S. Sharma *et al.*, "Indian classical music with incremental variation in tempo and octave promotes better anxiety reduction and controlled mind wandering—A randomised controlled EEG study," *Explore*, 2020, doi: 10.1016/j.explore.2020.02.013.

[40] M. Orini, F. Al-Amodi, S. Koelsch, and R. Bailón, "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.*, vol. 10, 2019, doi: 10.3389/fphys.2019.01465.

[41] S. Ishimitsu, K. Oue, A. Yamamoto, and Y. Date, "Sound quality evaluation using heart rate variability analysis," 2017.

[42] B. Benward and M. Saker, *Music in theory and practice, Volume I*, 8th ed. New York: McGraw-Hill, 2009.

[43] C. J. Murrock and P. A. Higgins, "The theory of music, mood and movement to improve health outcomes: Discussion paper," *J. Adv. Nurs.*, 2009, doi: 10.1111/j.1365-2648.2009.05108.x.

[44] M. R. Ebben, P. Yan, and A. C. Krieger, "The effects of white noise on sleep and duration in individuals living in a high noise environment in New York City," *Sleep Med.*, vol. 83, pp. 256–259, 2021, doi: 10.1016/j.sleep.2021.03.031.

[45] T. A. Pickens, S. P. Khan, and D. J. Berlau, "White noise as a possible therapeutic option for children with ADHD," *Complement. Ther. Med.*, vol. 42, pp. 151–155, 2019, doi: 10.1016/j.ctim.2018.11.012.

[46] Y.-A. Chiou *et al.*, "Electrocardiogram lead selection for intelligent screening of patients with systolic heart failure," *Sci. Rep.*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-81374-6.

[47] M. Marei, S. E. Zaatari, and W. Li, "Transfer learning enabled convolutional neural networks for estimating health state of cutting tools," *Robot. Comput. Integr. Manuf.*, vol. 71, 2021, doi: 10.1016/j.rcim.2021.102145.

[48] A. Yildiz, H. Zan, and S. Said, "Classification and analysis of epileptic EEG recordings using convolutional neural network and class activation mapping," *Biomed. Signal Process. Control*, vol. 68, 2021, doi: 10.1016/j.bspc.2021.102720.

[49] A. Canziani, A. Paszke, and E. Culurciello, "An analysis of deep neural network models for practical applications," *arXiv Prepr. arXiv1605.07678*, 2016.

[50] T. Li and M. Zhou, "ECG Classification Using Wavelet Packet Entropy and Random Forests," *Entropy* , vol. 18, no. 8. 2016, doi: 10.3390/e18080285.

[51] M. Takahashi and Y. Takai, "Low-temperature growth of YBa2Cu3Ox by pulsed laser deposition," *Supercond. Sci. Technol.*, vol. 11, no. 3, pp. 265–269, 1998, doi: 10.1088/0953-2048/11/3/002.

[52] H. Moeinzadeh *et al.*, "Wctecgdb: A 12-lead electrocardiography dataset recorded simultaneously with raw exploring electrodes' potential directly referred to the right leg," *Sensors (Switzerland)*, vol. 20, no. 11, pp. 1–12, 2020, doi: 10.3390/s20113275.

[53] A. Azizi and P. Ghafoorpoor Yazdi, "Introduction to Noise and its Applications BT - Computer-Based Analysis of the Stochastic Stability of Mechanical Structures Driven by White and Colored Noise," A. Azizi and P. Ghafoorpoor Yazdi, Eds. Singapore: Springer Singapore, 2019, pp. 13–23.

[54] D. Self, "Small signal audio design." Focal Press, Oxford, U.K.; Burlington, Mass., 2010.

[55] R. Kher, "Signal Processing Techniques for Removing Noise from ECG Signals," *J. Biomed. Eng. Res.*, vol. 3, 2019.

[56] C. Wang, S. Yang, X. Tang, and B. Li, "A 12-Lead ECG Arrhythmia Classification Method Based on 1D Densely Connected CNN," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11794 LNCS. Springer, Chengdu Spaceon Electronics Co., Ltd., Chengdu, China, pp. 72–79, 2019, doi: 10.1007/978-3-030-33327-0_9.

[57] M. Chala, B. Nsiri, M. H. El yousfi Alaoui, A. Soulaymani, A. Mokhtari, and B. Benaji, "An automatic retinal vessel segmentation approach based on Convolutional Neural Networks," *Expert Syst. Appl.*, vol. 184, 2021, doi: 10.1016/j.eswa.2021.115459.

[58] O. Archangelidi, M. Pujades-Rodriguez, A. Timmis, X. Jouven, S. Denaxas, and H. Hemingway, "Clinically recorded heart rate and incidence of 12 coronary, cardiac, cerebrovascular and peripheral arterial diseases in 233,970 men and women: A linked electronic health record study," *Eur. J. Prev. Cardiol.*, vol. 25, no. 14, pp. 1485–1495, 2018, doi: 10.1177/2047487318785228.

[59] S. Kuila, N. Dhanda, and S. Joardar, "ECG signal classification for arrhythmia detection using DEA and ELM," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 14, pp. 3403–3416, 2021.

[60] F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Front. Public Heal.*, vol. 5, p. 258, Sep. 2017, doi: 10.3389/fpubh.2017.00258.

[61] G. Zimatore *et al.*, "Recurrence quantification analysis of heart rate variability during continuous incremental exercise test in obese subjects," *Chaos*, vol. 30, no. 3, 2020, doi: 10.1063/1.5140455.

[62] W. Guo, C. Xu, J. Tan, and Y. Li, "Review and implementation of driving fatigue evaluation methods based on RR interval," vol. 503. Springer Verlag, Beijing Key Lab of Urban Intelligent Traffic Control Technology, North China University of Technology, Shijingshan, Beijing 100144, China, pp. 833–843, 2019, doi: 10.1007/978-981-13-0302-9_81.

[63] R. He *et al.*, "Automatic Detection of Atrial Fibrillation Based on Continuous Wavelet Transform and 2D Convolutional Neural Networks ," *Frontiers in Physiology* , vol. 9. p. 1206, 2018.

[64] M. Guo and Y. Du, "Classification of Thyroid Ultrasound Standard Plane Images using ResNet-18 Networks," in *13th IEEE International Conference on Anti-Counterfeiting, Security, and Identification, ASID 2019*, 2019, vol. 2019-Octob, pp. 324–328, doi: 10.1109/ICASID.2019.8925267.

[65] J. M. Rathod and H. S. Salehi, "Vision system with deep learning classifiers for automatic quality inspection," in *Pattern Recognition and Tracking XXXI 2020*, 2020, vol. 11400, doi: 10.1117/12.2555032.

[66] Olga Russakovsky*, Jia Deng*, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg and Li Fei-Fei. (* = equal contribution) ImageNet Large Scale Visual Recognition Challenge. IJCV, 2015.

[67] Ü. Çavuşoğlu, "A new hybrid approach for intrusion detection using machine learning methods," *Appl. Intell.*, vol. 49, no. 7, pp. 2735–2761, 2019, doi: 10.1007/s10489-018-01408-x.

[68] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, 2020, doi: 10.1186/s12864-019-6413-7.

[69] L. Frunzo, R. Garra, A. Giusti, and V. Luongo, "Modeling biological systems with an improved fractional Gompertz law," *Commun. Nonlinear Sci. Numer. Simul.*, vol. 74, pp. 260–267, 2019, doi: 10.1016/j.cnsns.2019.03.024.

[70] N. W. Alt and S. Jochum, "Sound Design Under the Aspects of Musical Harmonic Theory." SAE International , 2003, doi: 10.4271/2003-01-1508.

[71] J. Yang, Y. Pan, and Y. Luo, "Investigation of brain-heart network during sleep," in *42nd Annual International Conferences of the IEEE Engineering in Medicine and Biology Society, EMBC 2020*, 2020, vol. 2020-July, pp. 3343–3346, doi: 10.1109/EMBC44109.2020.9175305.

[72] K. Schiecke, A. Schumann, F. Benninger, M. Feucht, K.-J. Baer, and P. Schlattmann, "Brain-heart interactions considering complex physiological data: Processing schemes for time-variant, frequency-dependent, topographical and statistical examination of directed interactions by convergent cross mapping," *Physiol. Meas.*, vol. 40, no. 11, 2019, doi: 10.1088/1361-6579/ab5050.

[73] E. Bigand, R. Parncutt, and F. Lerdahl, "Perception of musical tension in short chord sequences: The influence of harmonic function, sensory dissonance, horizontal motion, and musical training," *Percept. Psychophys.*, 1996, doi: 10.3758/BF03205482.