Original Article

Electroencephalogram Sonification with Hybrid Intelligent System Design Based on Deep Network

Abstract

Background: The electroencephalogram (EEG) sonification is an audio portrayal of EEG signals to provide a better understanding of events and brain activity thereupon. This portrayal can be applied to better diagnosis and treatment of some diseases. **Methods:** In this study, a new method for EEG sonification is proposed based on extracting musical parameters and note sequences from the dominant frequency ratios and variations in the EEG signal. The ability of different classification structures in extracting musical scales and note sequences is evaluated. A music database has been created to train deep structures which, after extracting the frequency sequence of each piece of music as input, determines the scale label and note sequence in the output. A new algorithm is developed to combine the outputs of the deep structures and create a playable music repertoire. **Results:** The findings indicate that the convolutional neural network (CNN) classifier has an accuracy of 93.2% for the classification scales of musical pieces played in different octaves and 92.8% for pieces played in asymmetrical pieces. The convergence of EEG segments with musical scales is also reported for single channel, multi-channel of one person, different individuals, and different databases. The long short-term memory (LSTM) structure selected with an accuracy of 89.6% determines the sequence of notes. **Conclusion:** The results show that the proposed CNN determines the appropriate and convergent musical scales with each EEG signal fragment and the LSTM network has a promising performance in converting the dominant frequency variations of EEG signals into note sequences. This demonstrates the good performance of the proposed sonification method.

Keywords: *Deep neural network, electroencephalogram sonification, music scales, note sequences, time-frequency space*

Submitted: 17-Dec-2024 **Revised:** 23-Mar-2025 **Accepted:** 14-Apr-2025 **Published:** 01-Oct-2025

Hamidreza Jalali[1], Majid Pouladian[1], Ali Motie Nasrabadi[2], Azin Movahed[3]

[1]Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran, [2]Department of Biomedical Engineering, Faculty of Engineering, Shahed University, Tehran, Iran, [3]Department of Music, School of Performing Arts and Music, College of Fine Arts, University of Tehran, Tehran, Iran

Introduction

The central nervous system generates the electrical activity of neurons, enabling information processing. This continuous electrical activity, measured through electrodes embedded in the skull, is recorded as electroencephalogram (EEG) signals. These signals represent the sum of postsynaptic potentials from a large number of neurons.

Music, as an auditory stimulus with specific characteristics, has been shown to influence brain activity significantly. It has been the focus of various scientific studies due to its profound effect on cognitive and emotional processes. Understanding EEG signals in response to music is crucial, as it provides valuable insights into how the brain processes auditory stimuli. Many studies have explored this interaction, demonstrating the effects of music on EEG features such as the alpha rhythm in relation to tempo[1] and its role in emotional modulation through changes in alpha, theta, and gamma bands. The importance of music on brain activity has been further demonstrated through analyses of flash music,[2] the intensity of music practice across expertise levels,[3] and the influence of genre and speed.[4]

Relevance of prior studies

The reviewed studies provide significant insights into the connection between music and brain activity. For example, research on tempo and genre[5] highlights the influence of music on alpha and beta rhythms, while Banerjee *et al*.[6] showed that emotion-based brain activities are enhanced by listening to music. Studies such as Jenni *et al*.[7]

How to cite this article: Jalali H, Pouladian M, Nasrabadi AM, Movahed A. Electroencephalogram sonification with hybrid intelligent system design based on deep network. J Med Signals Sens 2025;15:29.

Address for correspondence:
Dr. Majid Pouladian, Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
E-mail: pouladian@srbiau.ac.ir

Access this article online

Website: www.jmssjournal.net

DOI: 10.4103/jmss.jmss_85_24

Quick Response Code:

explored how musical expertise impacts EEG signals, demonstrating distinct differences among nonmusicians, amateurs, and professionals. The application of independent component analysis to EEG signals recorded during music playback has revealed how modes and speeds influence brainwave patterns.[8]

Sonification techniques have also been extensively applied for clinical and therapeutic purposes.[9] For instance, EEG signal analysis has been used to diagnose congenital amusia,[10] evaluate anesthesia during surgery,[11] and detect epilepsy.[12] The effect of music on brain processes related to flash prediction,[13] as well as its role in controlling stress levels[14] and aiding in emotional regulation,[15] has been explored. Research has also focused on the real-time adjustment of musical parameters to optimize the therapeutic benefits of music.[16]

Despite these advances, gaps remain in the field.[17-19] Most studies employ simplified approaches, such as single-parameter mappings,[19-23] or focus on specific aspects such as amplitude and intensity. In addition, although some studies combine EEG with other imaging modalities such as functional magnetic resonance imaging,[24] these approaches are not yet widespread. Comprehensive frameworks for EEG-to-music conversion that incorporate temporal and frequency features are lacking, leaving room for significant improvement in this area.[25-31]

Research aim

The primary aim of this study is to address these limitations by proposing a novel framework for EEG-to-music conversion. This framework will utilize advanced, multi-parameter mappings that incorporate frequency, temporal, and statistical features to create richer and more meaningful musical outputs. By leveraging these features, the study seeks to bridge the gap between neuroscience and musicology, offering enhanced applications in diagnosis, treatment, and therapeutic interventions.

In this study, a deep neural intelligent system is presented to make music from the EEG signal, where, first, a database containing note sequences and different pitches are formed; next, according to this time–frequency changes, a deep neural convolutional neural network (CNN) network is trained to determine the scale music and an RNN is trained to determine the sequences notes. The trained neural networks extract scale and sequence notes from the EEG signal, and then, a multi-step algorithm converts the output labels of the networks into a playable music repertoire. A sequence of 7 consecutive notes, each at a specific frequency distance from the main note (tonic), is called a musical scale. The scale in music expresses the mood and notes used in it and is the most important adjustable parameter in creating a piece of music. Furthermore, each note has a specific spectral space that has the most power at a specific frequency, which is why a note is tuned to that frequency (for example, A is 440 Hz). The sequence of notes represents the sound waves with specific frequencies that create a piece of music and frequency sequence refers to the frequencies of different notes in succession.

The general structure of this article is as follows:

In the method section, the music databases-making procedure and EEG databases, preprocessing of each, the method of extracting time–frequency sequences and processing, the applied and designed classifications, and the method of making music repertoire from output signals are described. In the results section, the detection percentages of tonal pitch and note sequence of music signals are assessed in different conditions and the degree of EEG signals convergence into music pitch for time, channel, people, and different databases is expressed subject to different situations and at the end, the results are discussed and concluded.

Methods

The general research method consisting of segmentation of EEG signal, preprocessing method, and extracting time–frequency sequences for pitch and note sequence detection and classification. The stages are shown in Figure 1. After determining the final note sequence, the proposed method of making a musical repertoire from the outputs of the classification structure is described.

Databases

The process of creating a music database for training and testing classifiers is explained below, followed by a description of the EEG databases utilized.

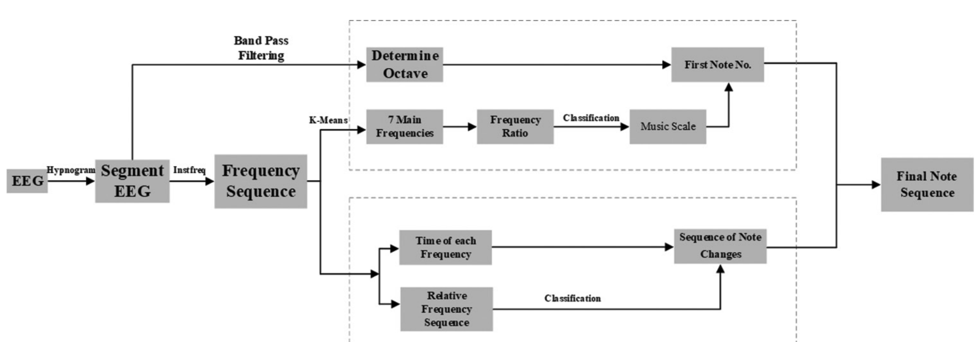

Figure 1: Diagram of the methods used to electroencephalogram sonification. EEG: Electroencephalogram

Music database

The YAMAHA CLP electric piano, featuring precise tuning, was used to generate the music database. The output of the music pieces was recorded and transferred through a software interface at a sampling frequency of 44,000 Hz, in monochannel and 8-bit format. Based on the structure of tonal music, the pieces were composed in the 12 major scales, each consisting of seven consecutive notes played for 250 ms, as follows:

A. Playing Seven Different Notes in One Scale (250 ms Each):
   - All possible sequences of the seven notes were played, resulting in 5040 unique audio pieces for each scale
   - These audio pieces were then performed across three octaves: 440 Hz, 220 Hz, and 880 Hz, all cantered on the note "La"
   - In total, the music database contained 181,440 audio pieces, encompassing all combinations of seven notes in each of the 12 major scales across the three octaves. Each piece was 1.75 s long
   - A unique number was assigned to each piano key corresponding to the notes, and the sequence of the played notes in each music piece was recorded.

B. Playing Fewer than Seven Notes with Combinations of Homonyms and Non-Homonyms in One Scale (Variable Durations):
   - Numerous combinations were possible when playing fewer than seven notes; however, to reduce errors in neural network training and classification, irregular pieces were avoided
   - To create these pieces, the three most important notes of each scale were identified, considering the tonal structure and sequence of sharps and flats
   - The following rules were applied to construct the sequences:
     - Each piece must include two or three notes from importance level 1
     - Each piece must include one or two notes from importance level 2
     - Each piece must include at least one note from importance level 3
     - Remaining notes in the sequence were filled with other notes from the scale.
   - Using this method, 19,320 pieces were recorded for the second mode
   - Similar to the first mode, a number was assigned to each piano key corresponding to the notes, and the sequence of the played notes was recorded.

This database provides a comprehensive and structured set of music pieces for use in training and testing classifiers.

The scale and its label, the notes in each scale, the sequence of their labels, and the important notes of each scale are tabulated in Table 1.

The electroencephalogram signal database

In this study, three databases are utilized to access EEG signals and validate the results at different database

Table 1: Identification of labels applied for major scales-note sequences

					Grade 1 note		Grade 2 note	Grade 3 note
Major Scale-note sequence	C	D	E	F	G	A	B	C
Labels	1	3	5	6	8	10	11	1
Major Scale-note sequence	G	A	Si	C	D	E	# F	G
Labels	2	10	12	1	3	5	7	8
Major Scale-note sequence	D	E	# F	G	A	B	# C	D
Labels	3	5	7	8	10	12	2	3
Major Scale-note sequence	A	B	# C	D	E	# F	# G	A
Labels	4	12	2	3	5	8	9	10
Major Scale-note sequence	E	# F	# G	A	B	# C	# D	E
Labels	5	8	9	10	12	2	4	5
Major Scale-note sequence	B	# C	# D	E	# F	# G	# A	B
Labels	6	2	4	5	8	9	11	12
Major Scale-note sequence	F	G	A	b B	C	D	E	F
Labels	7	8	10	11	1	3	5	6
Major Scale-note sequence	b B	C	D	b E	F	G	A	b B
Labels	8	1	3	4	6	8	10	11
Major Scale-note sequence	b E	F	G	b A	b B	C	D	b E
Labels	9	6	8	9	11	1	3	4
Major Scale-note sequence	b A	b B	C	b D	b E	F	G	b A
Labels	10	11	1	2	4	6	8	9
Major Scale-note sequence	b D	b E	F	b G	b A	b B	C	b D
Labels	11	4	6	7	9	11	1	2
Major Scale-note sequence	b G	b A	b B	b C	b D	b E	F	b G
Labels	12	9	11	12	2	4	6	7

levels. The selection criteria for these databases include their reliability, data recording, and presentation in accordance with R and K[32] and/or AASM[33] standards. The databases are the DREAMS Subjects (DRMS) Database,[34] The Sleep EDF (S-EDF) Database,[35] and ISRUC Database (Subgroup 3, ISRUC3).[36] The details of the databases are tabulated in Table 2.

Preprocessing

Music signal

To ensure accurate classification, the length of each music piece is standardized to 1.75 s. Due to hardware limitations, the duration of each note is set to 250 ms. To address resonance from other frequencies, nonessential frequencies are removed using frequency filtering methods, with spectral preprocessing applied to the samples. The frequency variations of the preprocessed signal are illustrated in Figure 2.

Electroencephalogram signal

Frequency domain filtering

A 10^{th}-order Butterworth filter is applied to the EEG signal within the meaningful frequency range (e.g., 0.5–45 Hz for healthy individuals and similarly for pathological recordings).

Ensemble empirical mode decomposition with canonical correlation analysis method

Motion artifacts are removed using the Ensemble Empirical Mode Decomposition with Canonical Correlation Analysis (CCA) method[37] and provided that the approach is valid for a single channel or at least one reliable channel. This method first decomposes the signal into a set of sub-bands (from high to low frequencies). After the initial wavelet-based denoizing, the CCA algorithm is applied to calculate correlations between each sub-band. Highly correlated components are retained, while less correlated ones are discarded.

Extracting time–frequency sequences

Frequency variations in music, such as changes in treble, bass, or the distance between two notes, determine the pitch and sequence of notes. The process of creating a music database includes recording note durations, note labels, and step labels for each piece of music. Similarly, EEG signals are labeled and segmented based on the dominant frequency band or hypnogram.

Step recognition of sequential frequencies

The notes played in a piece of music define its scale type, with an emphasis on tonal music. The arrangement of different notes determines the context of a 12-major scale, where the scale detection classifier outputs 12 distinct labels. Although each step involves a sequence of 7 notes (including repeated and nonrepeated notes during performance), the classifier input comprises 7 entries, as detailed below:

1. Filtering Frequency Resonance and Noise: To remove artifacts caused by frequency resonance in music or momentary EEG signal interference, segmented signals corresponding to the sleep stage or music are filtered based on the relevant frequency octave
2. Estimating Instantaneous Frequency: The instfreq function is used to calculate the instantaneous dominant sequential frequency of the filtered signal.

Table 2: Details of electroencephalogram databases used

Dataset name	Criteria	Epoch length (s)	Number of subjects	Recoding files	Age	Sampling frequency (Hz)	EEG channel	Total number of epochs
S-EDF	R and K	30	26	34	25–96	100	Pz-Oz	104,643
DRMS		20	20	20	20–65	200	Cz-A1	30,401
DRMS	AASM	30	20	20	20–65	200	Cz-A1	20,265
ISRUC3		20	10	10	30–58	200	C3-A2	8889

EEG – Electroencephalogram; DRMS – DREAMS subjects; S-EDF – Sleep European Data Format; AASM – American academy of sleep medicine

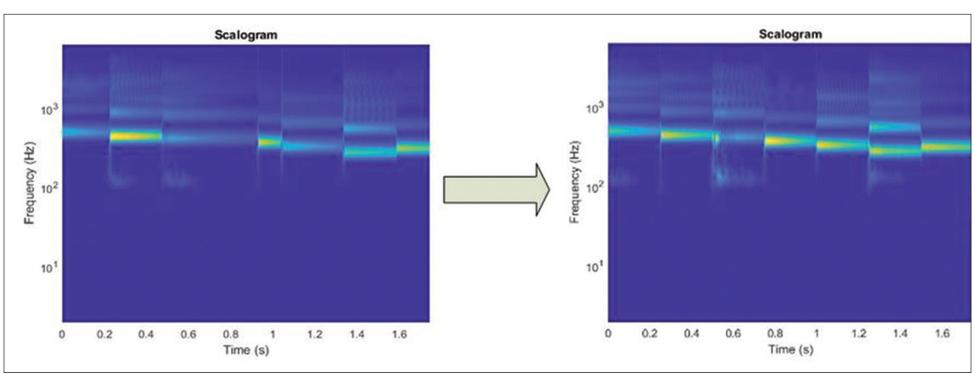

Figure 2: Scalogram of the music signal before and after preprocessing

This algorithm also outputs the time status of each momentary frequency, which contributes to constructing the note sequence and final repertoire
3. Clustering Frequencies: The estimated instantaneous frequencies are input to the K-means clustering algorithm,[38] which groups the frequencies into 7 clusters. The cluster centers and their respective weights (frequency counts) are determined as outputs
4. Frequency Normalization: Each cluster's center frequency is divided by the smallest cluster center frequency. Musically, this represents the frequency ratio of each note to the first note of the octave, providing the intervals between notes and defining the scale.

These 7 normalized values represent the dominant frequencies of the music or EEG signal segments and serve as the classifier input, while the music scale label is the output.

Changes in note sequences

To analyze note sequences, recurrent classifiers are employed. The input to these classifiers is the ratio of each instantaneous frequency to the preceding one, which reflects the rate of frequency change.

In music, such changes define the pitch and indicate subsequent notes. The classifier output captures the number of note changes and the duration of each note (or the constant state of instantaneous frequency). The duration of each note is derived from the instfreq function. The instantaneous frequency of a nonstationary signal is a time-varying parameter that represents the signal's average frequency content over time. The instfreq function estimates this frequency as the first conditional spectral moment of the signal's time-frequency distribution. Specifically, the function performs the following steps:[39]

- Computes the spectrogram power spectrum P (t, f) of the input signal using the pspectrum function, treating it as the time–frequency distribution
- Estimates the instantaneous frequency based on the resulting time–frequency representation as defined in Eq. 1.

$$f_{\text{inst}}\left(t\right)=\frac{\int_0^\infty fP(t,f)\,df}{\int_0^\infty P(t,f)\,df} \tag{1}$$

The following steps outline the process:
1. Extract Instantaneous Frequencies: Retrieve the sequence of instantaneous frequencies from the music signal
2. Determine Frequency Timing: Record the occurrence time of each frequency
3. Calculate Frequency Ratios: Create a sequence of ratios between consecutive instantaneous frequencies
4. Numerical Note Differences: Quantify the difference between consecutive notes. For example, if the first note is Re (3), the second note Fa (6), and the third note Mi (5), the sequence would be: 5,-1. If no frequency change occurs, the entry is 0
5. Classifier Input and Output Formation: The input consists of sequences of frequency ratios extracted through instfreq, while the output contains sequences of consecutive note changes as classifier labels
Note: After training the classifiers on music sequences, momentary frequency ratios from EEG signals are used to determine the sequence of changes or note numbers at the output.
6. Data Segmentation for Classifiers: To define the input size for classification systems, the extracted sequences are segmented into 7-bit chunks.

This structured approach facilitates the recognition and classification of musical scales and note sequences, enabling applications in both music analysis and EEG signal interpretation.

Classification

In classification, the issues of determining a signal piece scale and the sequence of note changes of a signal piece are essences:

The music scale

To determine the 12 scales of the music signal and EEG signal, different classification structures and the designed CNN are applied and their results are analyzed. The input of these classification structures is a 7-bit sequence extracted through the K-means algorithm. The K-means input is the instantaneous frequency sequences obtained from the music or EEG signal. To train, the music database containing 181,440 audio pieces including all the states of 7 different notes placement of one step out of 12 major scales, in three octaves in the first mode and 19,320 pieces of music with different scales and repeated notes in the second mode. The output of the classifier structures is proportional to the 12-scale label input signal. The details of the proposed structures of concern together with the regulatory parameters are tabulated in Table 3.

Designed convolutional neural network structure

A CNN is a type of deep learning model primarily used for processing structured grid data, such as images. It is

Table 3: Description of control parameter of classifier used

Classifier	Control parameter	Range
KNN[40]	Centre number	3–7
	Distance measure	Euclidean-Manhattan-Minkowski
SVM[41]	Kernel	Linear, Gaussian, polynomial order 2 and 3
Decision tree[42]	Structure type	Subscale - simple scale - middle scale
Bagging[43]	Voting method	Majority vote
Boosting[43]	Voting method	Adaboost

SVM – Support vector machine; KNN – k-nearest neighbor designed to automatically and adaptively learn spatial hierarchies of features through convolutional layers, pooling layers, and fully connected layers. The convolutional layer applies a set of learnable filters (kernels) to the input to extract features such as edges or textures. The convolution operation is according to Eq. 2.

$$\left(f * g\right)\left(x, y\right) = \sum_i \sum_j f(i, j).g\ (x - i, y - j) \qquad (2)$$

Where, $f(i, j)$ is filter (kernel) values, $g(x–i, y–j)$ is input pixel values at position $(x–i, y–j)$. The result is passed through an activation function, such as ReLU ($ReLU\ (x)$ = max (0, x)).

Pooling reduces the spatial dimensions of feature maps, retaining important information while reducing computational cost. There are two kinds of pooling operators, maximum and average. The fully connected layer maps the learned features to the output classes using softmax function as Eq. 3.

$$y = \text{softmax}\left(Wx + b\right) \qquad (3)$$

Where, W, x, and b denote the weight matrix, input feature vector, and bias vector, respectively.

In this study, a lightweight 1D CNN is designed to improve the classification results. The input of this network is a 7-bit sequence extracted from the signal and the output is 12-scale labels, Figure 3.

Modifying CNN by applying cluster weight: the input of the classifiers is a 7-bit sequence extracted from the K-mean algorithm, while it might be one or more noise data. When the number of types of notes played in the music sequence is <7 different notes, the input of the classifier becomes sensitive to the number of notes played. To improve the decision-making performance, each cluster weight is applied in the classification. To add the cluster weight, after entering 7 features and applying different convolutional filters, the weighted averaging is applied in the pooling layer, which corresponds to the cluster weights for each input feature. To check the ability of applied classifiers in step detection after applying pooling with weighted averaging, different classifiers are applied in the softmax layer and their results are reported.

Note sequence determination

Long short-term memory (LSTM) is a type of RNN capable of learning long-term dependencies. It uses *gates* to control the flow of information and a cell state to maintain long-term memory.

Forget Gate (f_t): Decides which information to forget from the cell state according to Eq. 4.

$$f_t = \sigma\ (W_f\ [h_{t-1,} x_t] + b_f) \qquad (4)$$

Input Gate (i_t) and Candidate Memory (C_t): Decides which information to add to the cell state according to Eq. 5 and Eq. 6, respectively.

$$i_t = \sigma\ (W_i\ [h_{t-1,}\ x_t] + b_i) \qquad (5)$$

$$C_t = \tanh\ (W_c\ [h_{t-1}, x_t]\ b_c) \qquad (6)$$

Update Cell State (C_t): Updates the cell state with the forget and input gates according to Eq. 7.

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \qquad (7)$$

Output Gate (o_t) and Hidden State (h_t): Decides the output of the LSTM according to Eq. 8 and Eq. 9, respectively.

$$o_t = W_o\ [h_{t-1}, x_t] + b_o \qquad (8)$$

$$h_t = o_t \odot \tanh(C_t) \qquad (9)$$

x_t is the input at time t, h_{t-1} is the previous hidden state, C_t is the cell state at time t, W_f, W_i, W_c, W_o are the weight matrices, bi, b_f, b_c, b_o are the bias terms, σ is the sigmoid activation function, *tanh* is the hyperbolic tangent activation function, and ⊙ denotes element-wise multiplication.

Considering the short and long-time dependence of the note sequence in a piece of music, applying an LSTM structure is influential in classification.[44,45] The input of this classifier is a 7-bit sequence, where each bit is the instantaneous frequency extracted from the music signal to its previous instantaneous frequency ratio in the training and a similar 7-bit sequence extracted from the EEG signal in the test. The output of this classifier is a 7-bit sequence, where the number assigned of each bit represents the existing note's being pitch compared to the previous note. This sequence

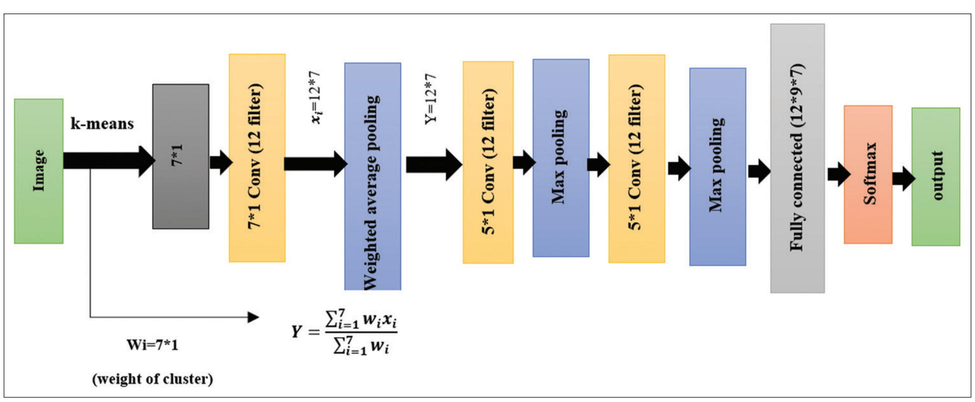

Figure 3: Overall structure of the designed convolutional neural network model for music scale recognition

indicates the distance difference between the existing note or piano key and the previous note.

LSTM network is of RNN type[46,47] with the main components, a sequence input layer and a short-term long-term memory layer able to learn short or long-term dependencies among successive data time steps. The input layer enters the sequence of time series data into the network and the LSTM layer learns the dependencies between the time steps of the successive data.

The architecture of an applicable LSTM network for music classification and production, where the size of the input sequence layer is 7 and the size of the fully connected layer is set to 7 numbers is shown in Figure 4.

The control parameters of the LSTM layer count, the neuron count, and the forgetting parameter are considered and reported after being assessed.

Creating a music repertoire

Determining the final scale label (the key signature of the piece of music)

A scale label is selected for each signal piece, while during a performance of a repertoire, the scale of the whole piece remains unchanged. To select the dominant scale in the music repertoire, all the labels of the music pieces are applied, and the dominant scale and the key signature are determined through the averaging method.

In conceptual context, the convergence to a specific sound–distance ratio indicates the frequency of the different notes embodied in the step and ultimately in the main key.

Location of the notes on the five carrier lines

The output of the label note sequence is the changes of each note concerning the previous note. If the location and the main number of the first note are determined on the musical lines, the location of the remaining notes will be determined as well. To determine the first note, the following steps are mandatory:

The frequency of the homonym notes changes at a 2:1 ratio in different octaves. On most musical instruments, only 5–7 octaves can be performed. In this study, the dominant bands and each sleep stage, according to the frequency band of the brain signal and the hypnogram curve correspond to one octave, consequently, the octave of the first note in the repertoire is determined from the sleep stage label extracted from the hypnogram of the EEG signal. According to the rules of music, the first note in a repertoire is a homonym with the step label of that piece of music indicating that the position and number of the first note of the repertoire are determined and the other notes are determined according to the sequence of note changes.

Note merging

The duration of output note time stretch from NN is about 250 ms, resembling one harp note (the tempo in the range of 120 black notes per/min). The consecutive similar notes are changed through the merging process by considering their count and note marks in the repertoire. For example: If the consecutive notes of two similar labels are determined, it would convert the two harp notes into one black note, Figure 5.

Results

Figure 6 shows an example of instantons frequency for various sleep stages and corresponding notes. The duration of the EEG signal is 60 s (horizontal axis of left images). As it can be observed, different stages represent a specific instantons frequency and notes.

The results of this study are of two categories:

The scale of a piece of music is an expression of the dominant frequency ratios in a piece of music signal. Considering the importance of determining the pitch in the construction of a piece of music, the ability of different classification structures and the NN designed to extract the musical scale should be assessed and reported. The classification structure should be applied in determining the music scale of the EEG signal, if better results are sought in the classification of the music signals' scale formed in different modes in the music database. Selecting an effective structure to determine the note sequence running tests on recursive deep structures is of major concern. In this context, first, the ability of the structure with different parameters in determining the note sequence from the frequency sequences of music signals is assessed and next, the best structure is selected and applied to determine the note sequence from the frequency changes in the EEG signal. After assessing and selecting the appropriate

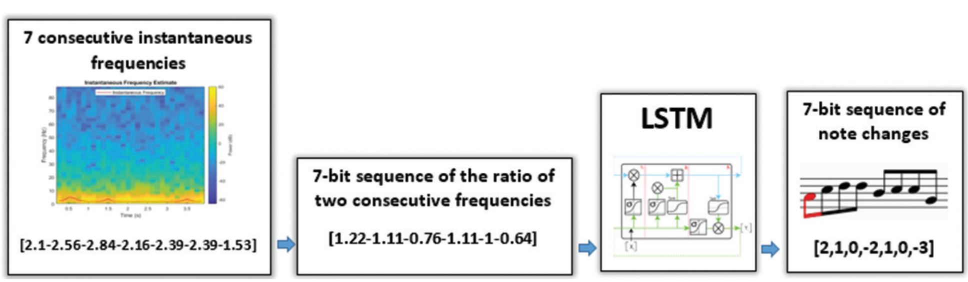

Figure 4: Flowchart to represent the note sequence generation using the long short-term memory model. LSTM: Long short-term memory

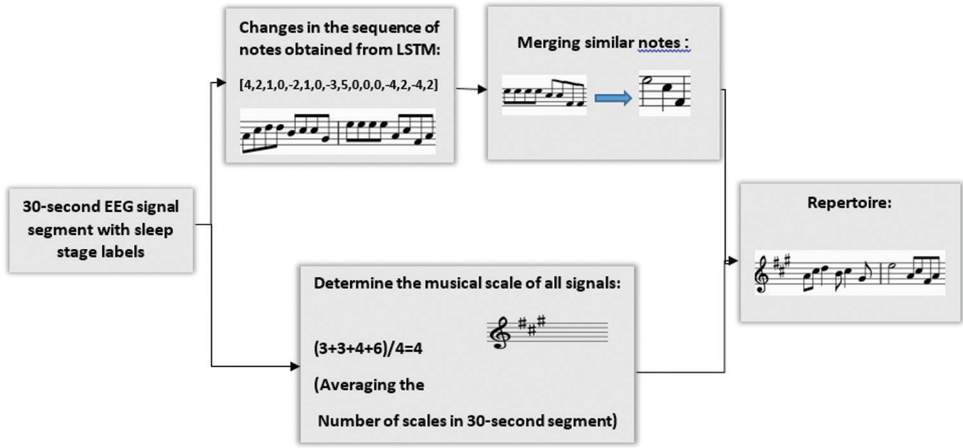

Figure 5: Diagram of the methods used to generate music repertoire from electroencephalogram signal. LSTM: Long short-term memory, EEG: Electroencephalogram

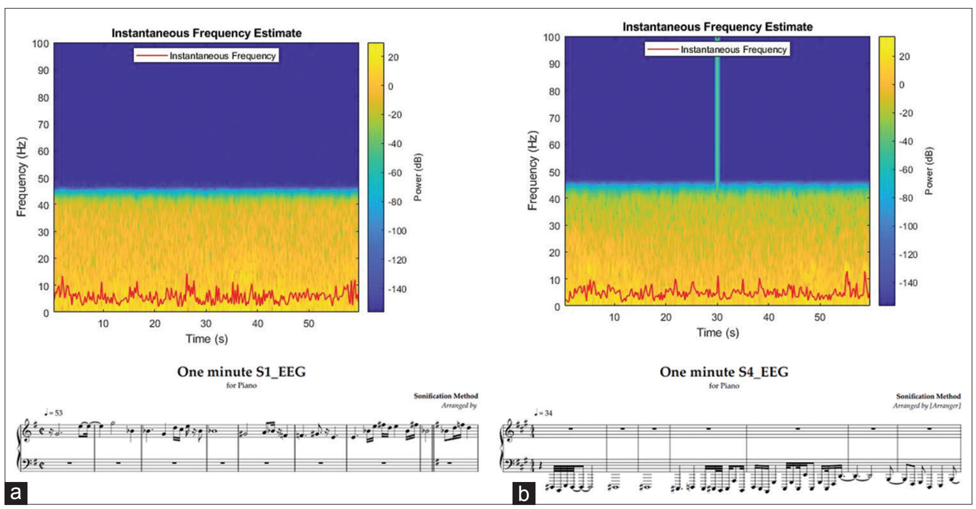

Figure 6: (a) Left: instfreq conversion image for S1 stage. Right: music repertoire determined for 1-min S1 stage. (b) Left: Instfreq conversion image for S4 stage. Right: music repertoire determined for 1-min S4 stage. EEG: Electroencephalogram

classification structure from the pitch and note sequence of the frequency sequences of music signals, the tonal pitch and note sequence are determined from the EEG signal by applying sequence to frequency EEG signal ratios. The criterion for reporting performance in this stage is the degree of convergence of the determined tonal pitch for each part of the EEG signal in different sleep stages, different standards, different channel counts, and different database counts to a given tonal scale.

Selecting the appropriate classifier structure for music signal

Assessing the ability of classifier structures in determining the music signals' scale recorded in different octaves.

A step is a sequence of 7 consecutive notes, the type and name of which are determined according to the frequency ratios therein. Different classifiers are applied to determine the 12 music scales, and the results are assessed. The input of each network consists of the frequency ratios extracted from the recorded different octaves, centered on La 440HZ, 220HZ, and 880HZ. The output of these networks consists of one of the 12-scale labels. The criterion for selecting the best structure for determining the scale at this stage is the accurate classification of music pieces in the abovementioned 3 octaves. Classes and the classification of the multi-octave database. This database consists of a random selection of 100,000 pieces of music signals in different octaves, and different structures are expected to have accurate performance in determining each music piece scale according to different frequency ranges, Table 4.

The ability of different structures in determining musical pieces' pitch where the note weight or time stretch is not equal

Another criterion for selecting an effective classification structure in pitch classification is its ability in determining different pieces of music be able to determine whether all or part of the 7 notes of the scale are played. To set up a database by playing part of the notes of the step based on music theories in the sharps and flats sequence, many notes with a higher degree of importance for each step are selected, thus, forming different parts with different scales. Each input is a sequence of frequency ratios extracted from pieces of music with different scales. The output of the classifiers is the 12-scale labels. The efficiencies of each classifier with the correct detection percentage of each step in each mode are tabulated in Table 5.

Optimal parameters of classifier in determining note sequence

An LSTM network is applied to determine the net sequence. The input of this classifier is a 7-bit sequence, where each bit is the outcome of an instantaneous frequency music signal to its previous instantaneous frequency ratio, and the output is a 7-bit sequence, where the number assigned to each bit represents the level of the existing note's being pitch than the previous note, Table 6.

Applying the trained classification structure for the electroencephalogram signal

In this step, the selected structures are applied to determine and report the EEG signals' note sequence and scale. The selected signals from different databases and the convergence degree of the EEG signal parts to the tonal music scale are assessed:

Single-channel electroencephalogram electroencephalogram of one person

Each one of the 12 tonal scales is an expression of the dominant frequencies' ratio in a signal. At this stage, the convergence of each sleep stage of one channel EEG signal of a selected person from the DRMS database, registered with ASSM and R and K standards, to one of the 12 tonal scales of music is assessed. The criterion for selecting one person from the database signals is the signal with less noise. Concerning the sleep stage labels of the database, the highest step label is determined as the main step label for that person, and the limits of convergence of all the EEG signal fragments with the main tonal scales are given as a percentage tabulated in Table 7.

Multi-channels electroencephalogram of one person

In this step, the similarity degree of the selected tonal scales for different recording channels in each part of the EEG signal of the selected person is assessed in the DRMS database, and the similarity percentage of different tonal scales of channels recorded in each part of the recorded signal of the person are tabulated in Table 8.

Different people

In this step, different scales of sleep convergence degree in all people of the DRMS database concerning tonal scales are assessed. The main scale is determined for each individual and the convergence average of all sleep stages in the recorded signal of each individual to the subject tonal pitch is calculated and the findings are tabulated in Table 9.

Table 4: Accuracy percentages of various classifiers in identifying music scales across octaves

Classifiers	Octave with A220 center	Octave with A440 center	Octave with A880 center	Multi octave
Best SVM	86.1	84.9	85.6	79.8
Best KNN	90.6	92.5	91.8	89.4
Best decision tree	76.4	77.6	75.9	71.6
Bagging	89.9	88.4	89.1	86.6
Boosting	97.6	96.5	96.9	95.8
Neural network	95.8	94.6	93.9	93.2

SV – Support vector machine; KNN – k-nearest neighbor

Table 5: Accuracy percentages for weighted and nonweighted classifiers to determining of music scales

Classifiers	A piece of music with 7 notes per scale (%)	A piece of music with several notes of higher importance (%)
SVM	80.1	68.6
Weighted SVM	78.8	72.4
KNN	88.5	73.5
Weighted KNN	86.7	76.7
Decision tree	71.8	59.4
Weighted decision tree	66.4	61.7
Bagging	87.1	7103
Weighted bagging	84.8	78.1
Boosting	94.8	84.7
Weighted boosting	91.2	86.7
CNN	94.6	89.9
Weighted CNN	93.7	92.8

CNN – Convolutional neural network; SVM – Support vector machine; KNN – k-nearest neighbor

Table 6: Accuracy percentages for different long short-term memory models to determining of the note sequence

Number of block	Number of hidden unit (%)		
	32	64	128
LSTM (2)	71.4	72.8	72.4
LSTM (3)	79.4	82.4	81.4
LSTM (4)	86.7	89.6	87.1
LSTM (5)	83.4	85.7	82.9

LSTM – Long short-term memory

Different databases

To measure the ability of this proposed method for different databases, the convergence of the sleep stages of all people in the DRMS, S-EDF, and ISRUC databases is calculated into 12 tonal scales and the findings are tabulated in Table 10.

Discussion

The objective of this study is to establish an EEG sonification framework by mapping EEG signals to musical scales using

Table 7: The percentage of convergence of all electroencephalogram signals in each sleep stage with main music scales

Sleep stage	Main Scale	Data with main label		Data with a difference of 2 from the main label		Data with more difference	
		R and K	AASM	R and K	AASM	R and K	AASM
Awa	4	57	59	26	29	17	12
REM	2	59	63	30	27	11	10
S1 (N1)	8	64	69	31	23	5	8
S2 (N2)	7	66	73	24	25	10	2
S3 (N4)	9	70	72	28	22	2	6
S4	6	69	-	20	-	11	-

AASM – American academy of sleep medicine; REM – Rapid eye movement

Table 8: The percentage of convergence of different channels of electroencephalogram in each sleep stage with main music scales

Sleep stage	Three channels to converge one scale		Two channels to converge one scale		Each channel to converge different scales	
	R and K	AASM	R and K	AASM	R and K	AASM
Awa	61	63	33	29	6	8
REM	64	62	28	35	8	1
S1 (N1)	67	66	26	27	7	7
S2 (N2)	65	60	26	26	9	14
S3 (N4)	59	58	35	24	6	18
S4	60	-	36	-	4	-

AASM – American academy of sleep medicine; REM – Rapid eye movement

Table 9: The percentage of convergence of different channels of electroencephalogram in each sleep stage with main music scales

Sleep stage	Awa	REM	S1 (N1)	S2 (N2)	S3 (N4)	S4
Average convergence of all persons						
R and K	61	59	69	71	68	65
AASM	63	62	72	73	70	-

AASM: American academy of sleep medicine; REM; Rapid eye movement

Table 10: The average percentage of convergence of all person's electroencephalogram signal in each sleep stage with main music scales in various dataset

Database	Average convergence to scales					
	Awa	REM	S1 (N1)	S2 (N2)	S3 (N4)	S4
DRMS	62	60	70	72	69	65
S-EDF	56	52	59	54	61	58
ISRUC3	51	49	57	54	59	-

DRMS – DREAMS subjects; S-EDF – Sleep European Data Format; REM – Rapid eye movement

classifiers. A database of musical pieces was created to evaluate classifiers' accuracy in identifying musical pitches across octaves. The boosting classifier achieved the highest accuracy (95.8%), followed by a neural network (93.2%). These models effectively distinguished musical scales under varying conditions, including asymmetrical sequences, by incorporating cluster weights and weighted average pooling, as shown in Table 5.

The CNN structure was tested for its ability to map EEG signals to musical scales. Analysis of EEG signal convergence, defined as the dominant frequency ratios' similarity across time and channels, demonstrated a 67% match for scales [Table 8] and a 73% match for individual pitch labels [Table 7]. This indicates the feasibility of using dominant frequency ratios from EEG signals to generate music aligned with individuals' unique neural patterns.

To refine this process, an LSTM-based model was developed to determine musical note sequences. This architecture, LSTM (4, 64), achieved 89.6% accuracy in estimating sequences from frequency changes. The model enables constructing a musical repertoire by aligning note sequences with EEG-derived dominant frequencies, addressing prior methods' limitations, such as disregarding signal intensity and musical rules. Researchers in Moradi *et al.*[31] addressed similar challenges by decomposing EEG signals into frequency sub-bands and mapping them to musical notes using an RNN.

Unlike earlier approaches relying on ad hoc rules, this study's AI-based method enhances EEG sonification by aligning frequency changes between EEG and music signals into a comparable range and integrating musical scales as critical parameters.

The findings of this study are summarized as follows:

1. Development of a method to analyze frequency changes in both music signals and EEG signals, ensuring that their rate of change falls within the same range

2. Evaluation of the ability to distinguish musical scales across different classes of music pieces
3. Design of a deep learning system with a high accuracy for identifying musical scales
4. Determination of musical note sequences using an LSTM-based model
5. Assessment of EEG signal convergence through musical scales, including single-channel fragments from one individual, multi-channel signals from one individual, signals from different individuals, and signals from various databases
6. Creation of a functional musical repertoire generated from EEG signals.

Conclusion

A method for converting EEG signals to music (EEG sonification) is proposed, based on extracting musical scales, and note sequences from the ratios and dominant frequency changes in the EEG signals. To train intelligent classifiers for determining musical scales and note sequences, music databases recorded in various octaves and modes are utilized, alongside extensive EEG signal databases for evaluation.

The results demonstrate that the proposed method, combined with the designed CNN, effectively identifies the musical scale of music signals and the convergence of an individual's EEG signal to a musical scale. Furthermore, the adoption of an LSTM network, selected based on the findings of this study, shows promising performance in converting dominant frequency changes in music or EEG signals into accurate note sequences.

It is also recommended to use neural network structures capable of extracting time-based information by optimizing filter parameters. The process of converting EEG signals to music could be further enhanced by incorporating additional features such as rhythm, tempo, and genre extracted from the EEG signals.

Authors' contributions

Hamidreza Jalali designed and coordinated the study, participated in most of the experiments, and drafted the manuscript. Majid Poladian helped design the study and supervised the study. Ali Moti Nasrabadi helped with all experiments and did the final editing of the text. Azin Movahed contributed to all experiments and edited the text. All authors have read and approved the content of the manuscript.

Acknowledgements

We highly appreciate the people who helped us in recording music signal and processing biological signals.

Financial support and sponsorship

Nil.

Conflicts of interest

There are no conflicts of interest.

References

1. Yuan Y, Lai YX, Wu D, Yao DZ. A study on melody tempo with EEG. J Electron Sci Technol 2009;7:88-91.
2. Brattico E, Tervaniemi M, Näätänen R, Peretz I. Musical scale properties are automatically processed in the human auditory cortex. Brain Res 2006;1117:162-74.
3. James CE, Oechslin MS, Michel CM, De Pretto M. Electrical neuroimaging of music processing reveals mid-latency changes with level of musical expertise. Front Neurosci 2017;11:613.
4. Lin YP, Duann JR, Feng W, Chen JH, Jung TP. Revealing spatio-spectral electroencephalographic dynamics of musical mode and tempo perception by independent component analysis. J Neuroeng Rehabil 2014;11:18.
5. Hurless N, Mekic A, Pena S, Humphries E, Gentry H, Nichols DF. Music genre preference and tempo alter alpha and beta waves in human non-musicians. Impulse 2013;22:1-11.
6. Banerjee A, Sanyal S, Patranabis A, Banerjee K, Guhathakurta T, Sengupta R, *et al*. Study on brain dynamics by non linear analysis of music induced EEG signals. Phys A Stat Mech Appl 2016;444:110-20.
7. Jenni R, Oechslin MS, James CE. Impact of major and minor mode on EEG frequency range activities of music processing as a function of expertise. Neurosci Lett 2017;647:159-64.
8. Omigie D, Pearce M, Lehongre K, Hasboun D, Navarro V, Adam C, *et al*. Intracranial recordings and computational modeling of music reveal the time course of prediction error signaling in frontal and temporal cortices. J Cogn Neurosci 2019;31:855-73.
9. Rogenmoser L, Zollinger N, Elmer S, Jäncke L. Independent component processes underlying emotions during natural music listening. Soc Cogn Affect Neurosci 2016;11:1428-39.
10. Norman-Haignere SV, Albouy P, Caclin A, McDermott JH, Kanwisher NG, Tillmann B. Pitch-responsive cortical regions in congenital amusia. J Neurosci 2016;36:2986-94.
11. Glen J. Use of audio signals derived from electroencephalographic recordings as a novel 'depth of anaesthesia' monitor. Med Hypotheses 2010;75:547-9.
12. Khamis H, Mohamed A, Simpson S, McEwan A. Detection of temporal lobe seizures and identification of lateralisation from audified EEG. Clin Neurophysiol 2012;123:1714-20.
13. Arslan B, Brouse A, Castet J, Filatriau JJ, Lehembre R, Noirhomme Q, *et al.* "Biologically-driven musical instrument." In Proceedings of the Summer Workshop on Multimodal Interfaces (eNTERFACE'05), Mons, BL: Faculté Polytechnique de Mons, 2005:18.
14. Miranda ER, editor. Brain-Computer Music Interface for Generative Music. Proceedings of International Conference on Disability, Virtual Reality and Associated Technologies. Esbjerg, Denmark: Citeseer; 2006.
15. Frassineti L, Barba C, Melani F, Piras F, Guerrini R, Manfredi C. Automatic detection and sonification of nonmotor generalized onset epileptic seizures: Preliminary results. Brain Res 2019;1721:146341.
16. González-Castañeda EF, Torres-García AA, Reyes-García CA, Villaseñor-Pineda L. Sonification and textification: Proposing methods for classifying unspoken words from EEG signals. Biomed Signal Process Control 2017;37:82-91.
17. Höller Y, Thomschewski A, Schmid EV, Höller P, Crone JS, Trinka E. Individual brain-frequency responses to self-selected music. Int J Psychophysiol 2012;86:206-13.

18. Hinterberger T, Baier G. Parametric orchestral sonification of EEG in real time. IEEE Multimed 2005;12:70-9.
19. Adrian E. The bebgee ehythm: Potential changes feom the occipital lobes in man. Physiol Lab Camb 1934;57:357-85.
20. Straebel V, Thoben W. Alvin Lucier's music for solo performer: Experimental music beyond sonification. Organised Sound 2014;19:17-29.
21. Baier G, Hermann T, Stephani U. Event-based sonification of EEG rhythms in real time. Clin Neurophysiol 2007;118:1377-86.
22. Wu D, Li C, Yin Y, Zhou C, Yao D. Music composition from the brain signal: Representing the mental state by music. Comput Intell Neurosci 2010;2010:267671.
23. Wu D, Li CY, Yao DZ. Scale-free music of the brain. PLoS One 2009;4:e5915.
24. Lu J, Wu D, Yang H, Luo C, Li C, Yao D. Scale-free brain-wave music from simultaneously EEG and fMRI recordings. PLoS One 2012;7:e49773.
25. Wu D, Li C, Yao D. Scale-free brain quartet: Artistic filtering of multi-channel brainwave music. PLoS One 2013;8:e64046.
26. Lu J, Guo S, Chen M, Wang W, Yang H, Guo D, *et al.* Generate the scale-free brain music from BOLD signals. Medicine (Baltimore) 2018;97:e9628.
27. Fernandes CM, Migotina D, Rosa AC. Brain's night symphony (BraiNSy): A methodology for EEG sonification. IEEE Trans Affect Comput 2018;12:103-12.
28. Fernandes CM, Mora A, Merelo JJ, Rosa AC, editors. Pherogenic Drawings-Generating Colored 2-dimensional Abstract Representations of Sleep EEG with the KANTS Algorithm. International Conference on Evolutionary Computation Theory and Applications. SCITEPRESS; 2012.
29. Franco P, Värri A. Experiments of the sonification of the sleep electroencephalogram. Finn J eHealth eWelfare 2015;7:65-74.
30. Tulilaulu A, Paalasmaa J, Waris M, Toivonen H, editors. Sleep Musicalization: Automatic Music Composition from Sleep Measurements. Advances in Intelligent Data Analysis XI: 11th International Symposium, IDA 2012, Helsinki, Finland: 2012 Proceedings 11. Springer; 2012.
31. Moradi F, Mohammadi H, Rezaei M, Sariaslani P, Razazian N, Khazaie H, *et al*. A novel method for sleep-stage classification based on sonification of sleep electroencephalogram signals using wavelet transform and recurrent neural network. Eur Neurol 2020;83:468-86.
32. Hori T, Sugita Y, Koga E, Shirakawa S, Inoue K, Uchida S, *et al*. Proposed supplements and amendments to 'a manual of standardized terminology, techniques and scoring system for sleep stages of human subjects', the Rechtschaffen and Kales (1968) standard. Psychiatry Clin Neurosci 2001;55:305-10.
33. Berry RB, Brooks R, Gamaldo CE, Harding SM, Marcus C, Vaughn BV. The AASM Manual for the Scoring of Sleep and Associated Events. Rules, Terminology and Technical Specifications. Vol. 176. Darien, Illinois: American Academy of Sleep Medicine; 2012. p. 2012.
34. Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, *et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101:E215-20.
35. Devuyst S, Dutoit T, Stenuit P, Kerkhofs M, Stanus E. Cancelling ECG artifacts in EEG using a modified independent component analysis approach. EURASIP J Adv Signal Process 2008;2008:1-13.
36. Khalighi S, Sousa T, Santos JM, Nunes U. ISRUC-sleep: A comprehensive public dataset for sleep researchers. Comput Methods Programs Biomed 2016;124:180-92.
37. Sweeney KT, McLoone SF, Ward TE. The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique. IEEE Trans Biomed Eng 2013;60:97-105.
38. Arthur D, Vassilvitskii S, editors. K-Means++ the Advantages of Careful Seeding. Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms; 2007.
39. Boashash B. Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. Proc IEEE 1992;80:520-38.
40. Mitchell TM. Artificial neural networks. Mach Learn 1997;45:127.
41. Kecman V. Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. MIT press; 2001.
42. Breiman L, Friedman J, Olshen R, Stone C. Classification and regression trees–crc press. Boca Raton, Florida. 1984;685.
43. Kotsiantis SB. Bagging and boosting variants for handling classifications problems: A survey. Knowledge Eng Rev 2014;29:78-100.
44. Mangal S, Modak R, Joshi P. LSTM based music generation system. arXiv preprint arXiv:1908.01080. 2019.
45. Agarwal S, Saxena V, Singal V, Aggarwal S, editors. LSTM Based Music Generation with Dataset Preprocessing and Reconstruction Techniques. 2018 IEEE Symposium Series on Computational Intelligence (SSCI); IEEE; 2018.
46. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput 1997;9:1735-80.
47. Van Houdt G, Mosquera C, Nápoles G. A review on the long short-term memory model. Artif Intell Rev 2020;53:5929-55.