

# Deep learning Binary/Multi classification for music's brainwave entrainment beats

Rowayda A. Sadek<sup>1</sup>, Alaa A. Khalifa<sup>1</sup> and Marwa M.A. Elfattah<sup>2</sup>

<sup>1</sup>Department of Information Technology, Faculty of computers and Artificial Intelligence, Helwan University, Cairo, Egypt

<sup>2</sup>Department of Computer Science, Faculty of computers and Artificial Intelligence, Helwan University, Cairo, Egypt

## ABSTRACT

The last two decades have seen the emergence of a brand-new kind of music known as digital brain stimulant, also known as instrumental music or music without lyrics, which mostly comprises entrainment beats. While listening to it has the same ability to affect the brain as taking medication, it also has the risk of having a negative impact or encouraging unwanted behavior. This sparked the interest of a large number of studies in the psychological and physiological effects of music's brainwave entrainment beats on listeners. These studies started to categorize and examine how musical beats affected brainwave entrainment by looking at electroencephalogram (EEG) signals. Although this categorization represents a step forward for the early research efforts, it is constrained by the difficulty of having each musical track and conducting EEG tests on humans exposed to distortion due to noise in order to determine its influence. The work proposed in this article continues to explore this topic but in a novel, simple, accurate, and reliable categorization procedure based on the music signal elements themselves rather than dependent on EEG. VGGish and YAMNET based transfer deep learning models, are tuned to handle a straightforward, accurate real-time detector for the existence of the music beats inside music files with accuracy of 98.5 and 98.4, respectively. Despite the fact that they yield results that are equivalent, the YAMNET model is more suited for use with mobile devices due to its low power consumption and low latency. The article also proposes modified version of VGGish and YAMNET binary classifying models called BW-VGGish and BW-YAMNET respectively. The modification was to turn the binary classification into multi-classification. These multi-classifiers handle the classification of the influence of music beats (five different brain waves) on human brainwave entrainment with average accuracy of 94.5% and 94.5%, respectively. Since there was a lack of datasets addressing this kind of music, two datasets, the Brainwave Entrainment Beats (BWEB) dataset and the Brainwave Music Manipulation (BWMM) dataset, were generated for classification training and testing. The re-testing on a sample of music files that have their impact on brain waves (with their EEG) in an earlier study is done to strengthen the validity of the proposed work and to overcome the potential limitation of utilizing a music dataset that is not proved with its EEG. The success of the suggested models was demonstrated.

Submitted 14 April 2023

Accepted 18 September 2023

Published 3 November 2023

Corresponding author

Alaa A. Khalifa,  
alaa.abdelaziz@fci.helwan.edu.eg

Academic editor

Khalid Raza

Additional Information and  
Declarations can be found on  
page 17

DOI 10.7717/peerj-cs.1642

© Copyright

2023 A. Sadek et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

**Subjects** Bioinformatics, Artificial Intelligence, Brain-Computer Interface, Neural Networks

**Keywords** Music beats, VGGish, YAMNET, BW-VGGish, BW-YAMNET, EEG, Transfer learning, Binaural beats, Monaural beats, Entrainment beats

## INTRODUCTION

Music has become an integral part of our daily lives as we listen to it during our various daily activities. As a result, many types of music have appeared to suit all tastes, including music that incorporates entrainment beats that may affect the brain in an analogous way to how drugs do, while also having a negative impact or encouraging undesirable behaviors (*Jirakittayakorn & Wongsawat, 2017; Wahbeh, Calabrese & Zwickey, 2007*). These beats can aid with anxiety, concentration, memory, mood, creativity, pain relief, meditation, cognitive flexibility, and sleep quality. Listening to it, on the other hand, might have side effects, such as increasing feelings of sadness, anxiety, and anger in certain people.

As a result, much research had been done on the psychological and physical impacts of music's brainwave entrainment beats on people using the electroencephalogram (EEG) signals. But these research face the challenge of having each musical track and conducting EEG experiments on humans to identify it, as well as the EEG recordings (*Repovš, 2010*), can be effected by different types of noise such as the electrical interference, electromyographic (EMG), and electrooculogram (EOG). Thus, it is important to detect the existence of the brainwave entrainment beats within music files and to categorize music files according to how they affect brain waves using the music signal features themselves rather than the EEG.

The bidirectional deep long short term memory (BDLSTM) model (*Sadek, Khalifa & Elfattah, 2021*) is a modified version of the long short term memory (LSTM) model. This model was used for its ability to learn, process, and categorize sequential data; it can also learn long-term relationships between time steps of data. It is successful to detect the presence of entrainment beats in music files with an accuracy of 90.02%, which serves as an example of why deep learning techniques are the most effective models for this problem. The BDLSTM model was hampered by the tiny amount of data that it had to work with, but this problem can be solved by applying the transfer learning methodology (*Bali & Tyagi, 2020; Gemmeke et al., 2017*).

As a result, our contributions to this article are:

- The detection of the existence of brainwave entrainment beats within music files with higher accuracy than the BDLSTM model using the VGGish and YAMNET models; binary classification
- The categorization of music beats files according to how they affect brain waves based on five different brain waves using the BW-VGGish and BW-YAMNET models; Multi-classification
- A Brainwave Musical Manipulation (BWMM) dataset was provided.
- Carry out a real testing verification process.

This work's remaining sections are organized as follows: The Literature Review section contains some information on earlier research on this topic. The structure of the dataset is discussed in the Dataset section. The architecture of the classification models is discussed in the Proposed Models section. The experiment and its results are discussed in the

Experimental Results section. The model's verification is discussed in the Proposed Models Verification section, which is followed by sections on conclusion and references.

## LITERATURE REVIEW

Based on our knowledge of all the research that was previously conducted on the effect of entrainment beats on brain waves, it has been carried out in two directions. In the first direction, the effect of these beats and the possibility of using them from a medical point of view were studied, while in the second direction, it was studied how to analyze and classify the effect of these beats on brain waves using electroencephalogram (EEG) waves and artificial intelligence techniques.

**Medical view:** *Becher et al. (2014)* presented a study that examined the modification of the electroencephalogram (EEG) power and phase synchronization by stimulation through sound with monaural and binaural beats in epilepsy patients before surgery. The results showed that phase synchronization can be modulated using audio stimulation, and it offered a non-surgical model for the modulation of intracranial EEG syncing that could have medicinal applications.

*Beauchene et al. (2016)* presented a study that determined the response accuracy and cortical network topology, which were measured by EEG recordings using different acoustic stimulation. The results showed that listening to 15 Hz binaural beats increased the response accuracy and modified the strengths of the cortical networks.

*Shumov et al. (2017)* presented a study that provided a comparative analysis of the time to fall asleep using three types of files: binaural beats (BB) with pink noise, a sound with monaural beats (MB), and a sound without any beat (imitations, IM). The results showed that listening to the soundtrack with BBs has an average time to fall asleep less than the time needed when listening to the soundtrack with MBs or IM.

*Engelbregt et al. (2019)* presented a study that examined the effects on attention and working memory in high and low emotional participants when listening to the binaural beats compared with the monaural beats using the stimulation process. A flanker task to measure attention and a Klingberg task to measure working memory were performed, and the results showed that for the attention, the performance speed was higher in the MB and BB conditions relative to white noise (WN), while in the Klingberg task, no significant differences were found under the WN, MB, and BB conditions.

*Ross & Lopez (2020)* presented a paper that shows the possibility of improving the training outcome in an attentional blink (AB) task using binaural beat stimulation. Magnetoencephalography (MEG) recordings show that there is a high improvement of gamma waves during 40 Hz binaural beats stimulation and a smaller improvement with 16 Hz binaural beats.

*Yusim & Grigaitis (2020)* presented a study that had the goal of finding the effects of binaural beats on brain activity. They compared the effects of the binaural beats with those of monaural beats, the results showed that binaural beats are weakly entrained in the cortex in comparison with monaural beat stimuli.

*Engelbregt et al. (2021)* presented a study that showed the 40 Hz monaural beats (MB) and binaural beats (BB) effects on attention and electroencephalogram (EEG). The

result showed that the binaural beats improved attention on the Flanker task, but this improvement did not appear in the electroencephalography.

**Technical view:** Many studies have been conducted on EEG-based recognition, where [Tai & Lin \(2018\)](#) presented a study that compares between the EEG signals of the brainwaves influenced by classical music and heavy metal music. In this study, nine people wear headsets and listen to two music files, which are Sonata for Two Pianos in D major, K.448” and “Napalm Death Procrastination On The Empty Vessel” for 10 min then their brainwaves were collected. Based on the collected brainwaves, the results confirmed that Listening to the classical music led to a drastic improvement in mental performance.

[Johnson & Durrant \(2021\)](#) investigated the neural mechanisms of brain-wave music on sleep quality. In this study, 33 participants listened to brainwave music for 20 min before bedtime for 6 days, and their EEGs were collected. According to the results, slow wave sleep brain-wave music has a positive effect on sleep. Furthermore, an increase in the functional connectivity (FC) between the left frontal and parietal lobes and a drop in the delta band’s power spectral density may be responsible for enhanced sleep quality.

According to our understanding and former literature assessment, no other experimenter worked on how to detect the presence or absence of brainwave entrainment beats inside audio recordings in real time. All the research that was conducted was to prove the effect of entrainment beats on different brain waves using EEGs.

First, the bidirectional deep long short-term memory (BDLSTM) model ([Sadek, Khalifa & Elfattah, 2021](#)) was used to detect the presence of brainwave entrainment rhythms in music files using music features. This model was based on the LSTM model with gaining average accuracy 90.2%. However, the problem with the dataset’s small size can be solved by using the transfer learning method, according to a study by [Bali & Tyagi \(2020\)](#), which showed that VGG16 transfer learning is a powerful technique where the knowledge gained from the larger dataset is transferred to the new dataset. It also showed that it is preferable to use transfer learning techniques rather than CNNs as it’s hard for the CNNs to be trained using small datasets. Also, [Barman et al. \(2019\)](#) performed a comparison between transfer learning ResNet50 and normal learning, and it was explained that transfer learning requires less data as compared to a model built from scratch, requires less computation power, and does not require a long time, whereas the pre-trained model does most of the heavy work, like recognizing major patterns, and does not require heavy GPU’s. Even working with huge data can be done on a decent CPU without a GPU.

When looking at previous studies on transfer learning techniques, some studies have been found, including ([Haq, 2021](#)) offers research that uses deep learning algorithms to classify music genres using the AudioSet dataset. This study compares the Vgg 19 transfer learning model, which had a 75 percent accuracy rate, with the Vgg 19 fine-tuning, which had a 78 percent accuracy rate, and the Vgg feed-forward baseline, which had an 81% accuracy rate, while [Shi et al. \(2019\)](#) present a study that uses the VGGish-BiGRU model, which is a VGGish network paired with a bidirectional gated recurrent unit neural network, and the dataset of 1,152 samples to distinguish lung sounds. This model was chosen because the information contained in lung sounds is too complicated to be displayed using standard characteristics, and the temporal features of lung sounds cannot be recovered using a



convolutional neural network. The settings of the VGGish layers are fixed, while the BiGRU network's parameters are fine-tuned. In the recognition of asthma accuracy, this model performs better than state-of-the-art algorithms like CNN.

In order to get around the dataset's small size, transfer learning was utilized. We tested a number of models before settling on the best one for identifying entrainment beats in music files. The most accurate models were then adjusted for use in the process of categorizing music files based on their impact on brain waves.

## DATASETS

Due to the lack of previous research that studied the entrainment beats using the music features themselves, there is a lack of datasets that include music files with brainwave entrainment beats. As a result, the brainwave entrainment beats (BWEB) dataset and the Brainwave Musical Manipulation (BWMM) dataset were created to be used in the experiments that we conducted during our study.

The music files for the BWEB and BWMM datasets were taken from SoundCloud and YouTube audio files. Audience comments and number of views guided the selection of audio files for download (*Gemmeke et al., 2017*). Each track lasts 30 s and is sampled at a rate of 44,100 samples per second in 32 bits.

### Brainwave entrainment beats (BWEB) dataset

The BWEB dataset (*Khalifa, Sadek & Elfattah, 2022*), is a 300-track self-gathered collection organized into two classes: brainwave entrainment music and non-entrainment music. Each class has an aggregate of 150 tracks, bearing in mind that each class contains all the different sorts that make up these classes. The brainwave entrainment class contains music files with three different entrainment beats, which are monaural beats, binaural beats, and isochronic beats, and the non-entertainment class contains different music files that have no entrainment beats. This set of files was put together to diversify the dataset and contain all the available possibilities for each class. Also, by adding an equal number of files to each kind, a balance has been maintained between the various classes.

### Brainwave musical manipulation (BWMM) dataset

The BWMM dataset (*Sadek, Khalifa & Elfattah, 2021*), is a 240-track self-collected collection organized into six classes which are, alpha, beta, delta, theta, gamma, and no entrainment. There are 40 tracks overall in each category. There are different types of files in each category. The tracks in the No Entrainment category were put up using audio files from numerous music genres as well as ambient sounds, whereas the three types of brainwave entrainment monaural, binaural, and isochronic are included in the other five categories. This variation of files was created to make the dataset more diversified and to include all the available options for each category. Also, by adding an approximately equal number of files to each kind, a balance has been maintained between the various categories. For example, the alpha category contains 40 files from different three categories which are 13 files for monaural, 14 for binaural, and 13 for isochronic.

Samples for the dataset were taken from audio files on the servers of SoundCloud and YouTube. Based on audience feedback, the audio samples were picked for download. Each

track lasts for 30 s and is sampled at a rate of 44,100 samples per second at a bit depth of 32. Samples were taken from each audio file in a systematic method, covering the beginning, middle, and end of each audio file separately.

## PROPOSED MODELS

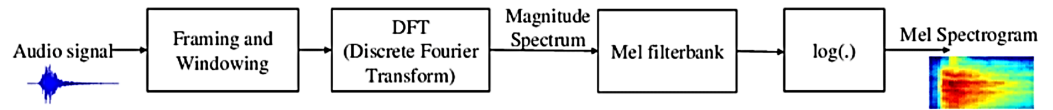
The proposed models that we reached have the ability to classify music files based on the presence of brainwave entrainment beats inside them, in addition to the ability to classify music files based on their effects on brain waves (five different brain waves). To get to these models, we have gone through two phases. In the first phase, we have gone through many models that had drawbacks that led us to use another model to overcome these drawbacks until we reached our proposed models. All these models have been trained and tested with the Brainwave Entrainment Beats (BWEB) dataset in the process of detecting the existence of entrainment beats inside music files. Then we went to the second phase, in which the models that achieved the highest accuracy were developed in the first phase for use in the process of classifying music files based on their effect on brain waves.

The models that we presented in the first phase are: AS-BDLSTM, VGGish, and YAMNET. All these models were trained and tested using the Brainwave Entrainment Beats (BWEB) dataset (*Khalifa, Sadek & Elfattah, 2022*). Firstly, we developed the AS-BDLSTM model using the transfer learning methodology by training it first on the AudioSet dataset and then using it with the target dataset, which improved the results but required a long training time and a lot of memory. This is what prompted us to consider employing the VGGish and YAMNET models, which were developed by Google.

While in the second phase, we developed the BW-VGGish and BW-YAMNET models from the VGGish and YAMNET models by training them on the Brainwave Entrainment Beats (BWEB) dataset in addition to their pretraining on the AudioSet dataset, then using them with the target dataset, which is BrainWave Musical Manipulation (BWMM). Because the BWEB dataset contains data that is similar to the target dataset, BW-VGGish and BW-YAMNET gained more experience when training on it. This experience makes BW-VGGish and BW-YAMNET more accurate in the categorization of music files based on their effects on brain waves.

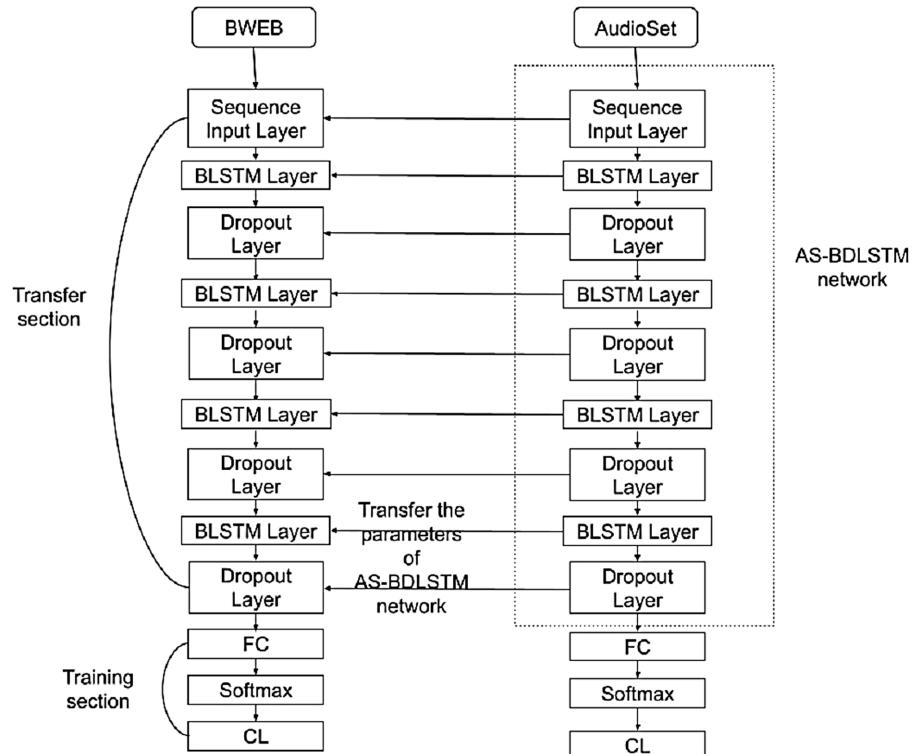
There are six different effects that brain waves can have.

- **Delta** (People in this state enter a state of deep sleep)  
Beats range: 0.5–4 Hz
- **Theta** (People in this state enter a first state of sleep and feel drowsiness)  
Beats range: 4–8 Hz
- **Alpha** (People in this state should be relaxed but alert)  
Beats range: 8–13 Hz
- **Beta** (People in this state should be highly alert and focus)  
Beats range: 13–30 Hz
- **Gamma** (People in this state should be able to recall memories)  
Beats range: >30 Hz
- **No effect** (No entrainment beats)



**Figure 1** Block diagram of Mel spectrogram of an audio signal.

Full-size DOI: 10.7717/peerjcs.1642/fig-1



**Figure 2** AS-BDLSTM model architecture.

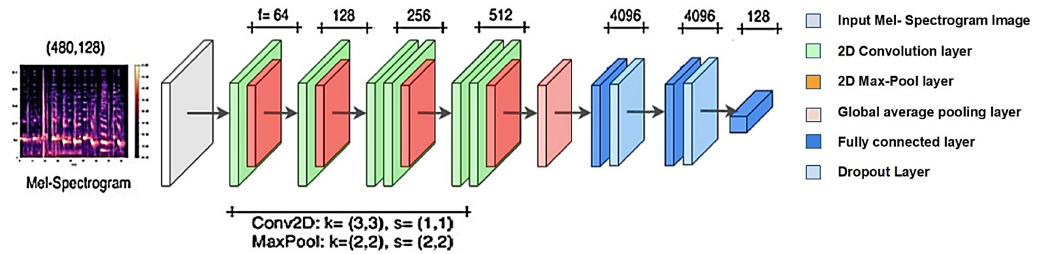
Full-size DOI: 10.7717/peerjcs.1642/fig-2

All the proposed models (VGGish, YAMNET, BW-VGGish, BW-YAMNET) were trained on the log mel spectrogram of the data, which made it necessary to pre-process both the BWEB and BWMM datasets with the following steps:

- Convert audio signals to the model's learned sampling rate.
- Obtain the log mel spectrogram as shown in Fig. 1, which is divided into 96 frames with 48-frame overlap.

### AS-BDLSTM

As shown in Fig. 2, the AS-BDLSTM approach was constructed by first training the BDLSTM approach on the AudioSet dataset, then retraining the output approach on the target dataset, which was a BWEB dataset. The retraining process is called the transfer learning process, which transfers the first training parameters to the target network. AudioSet and BWEB datasets were pre-processed, and the Mel-FrequencyCepstral Coefficients (MFCC) and Gammatone Cepstral Coefficients (GTCC) features were



**Figure 3** VGGish general model scheme.

Full-size DOI: [10.7717/peerjcs.1642/fig-3](https://doi.org/10.7717/peerjcs.1642/fig-3)

extracted as feature matrix in the form of sequence data, which was sent into the network as a sequence input through the sequence input layer.

### VGGish model

The VGGish model is a modification of the VGG model (*Hershey et al., 2017*) and it is appropriate to be used with tiny datasets (*Bali & Tyagi, 2020*; *Gemmeke et al., 2017*). It is pre-trained on the AudioSet dataset, which is a large-scale labeled audio dataset that Google opened in 2017. It consists of 24 layers. Six convolutional layers and three fully connected layers are among the nine layers with learnable weights, as shown in *Fig. 3*.

The VGGish model is a ready-made model; when using it, we use the same layers of the model with the removal of the last layer to match the size of the number of the out classes and adding a group of fully connected, softmax and classification layers. Then we train it on the target data set with an input size of  $96 \times 64$ , as shown in *Fig. 4*. During the training process, the training parameters are transformed from the VGGish network to the target network, which enhances the performance result.

### YAMNET model

YAMNET (Yet Another Mobile Network) is a pre-trained neural network that uses the MobileNetV1 depthwise-separable convolution architecture. MobileNet is a mobile and embedded visual convolutional neural network. They are based on a condensed architecture that builds thin deep neural networks using depthwise separable convolutions (*Harjoseputro, Yuda & Danukusumo, 2020*). It is characterized by low latency and low power consumption. The scheme of a generic YAMNET deep neural network for classification tasks is represented in *Fig. 5*, as the input features go through the input layer and then the input features are gone through multiple convolutions. Finally, the output of the convolution layers go through the fully connected layers, which send the outs to the output layer that is out the results.

YAMNET is made up of 86 layers. There are 28 layers with weights that may be learned: There are 27 convolutional layers and one fully connected layer. It is a pre-built model trained on the AudioSet dataset from Google. With the exception of deleting the last layer to match the number of out classes and adding a set of entirely connected, softmax, and classification layers, we use the same model layers as previously. Then, with a  $96 \times 64$

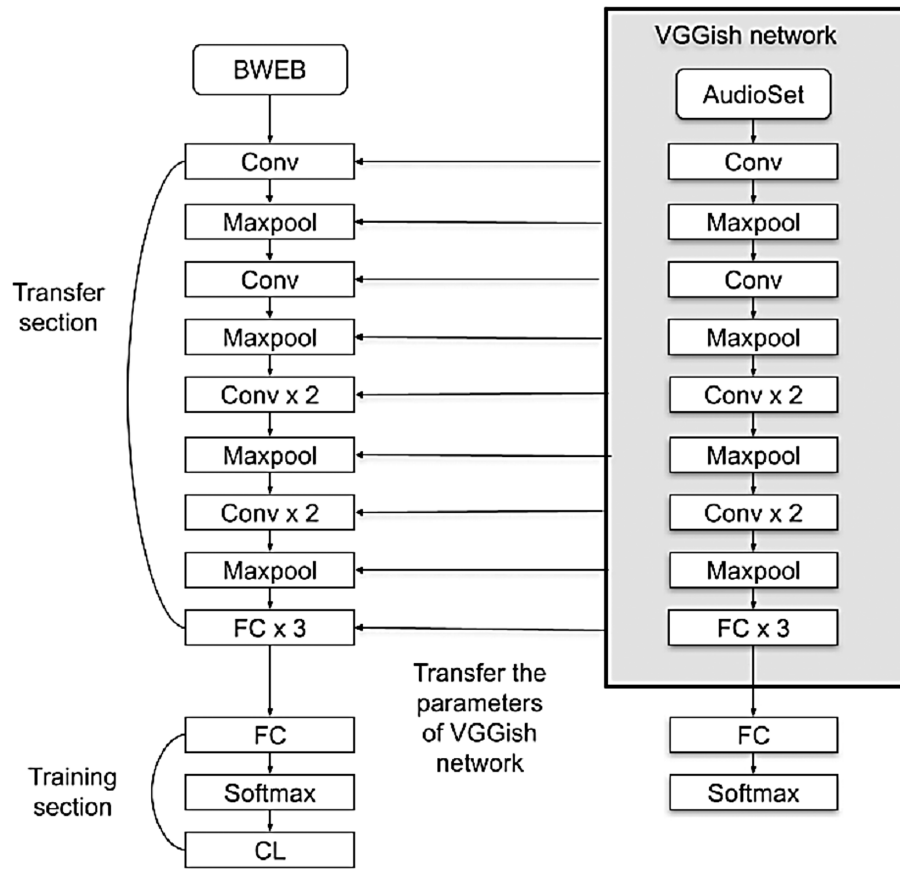


Figure 4 VGGish model architecture.

Full-size DOI: 10.7717/peerjcs.1642/fig-4

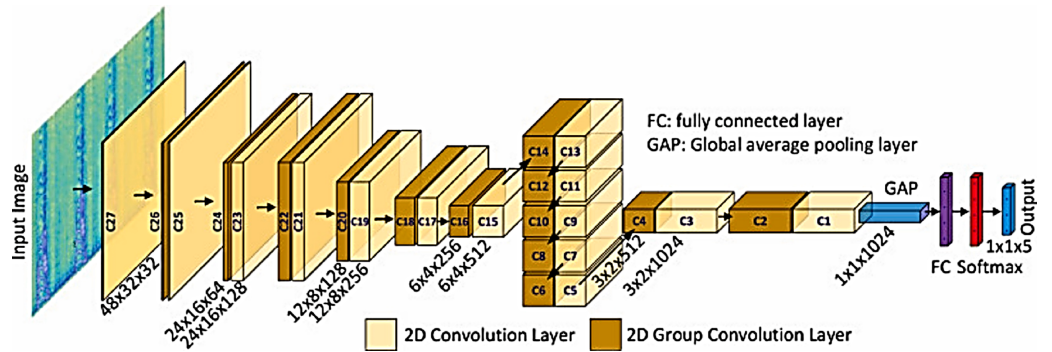


Figure 5 YAMNET model architecture.

Full-size DOI: 10.7717/peerjcs.1642/fig-5

input size, we train it on the target data set. Throughout the training process, the training parameters are switched from the YAMNET network to the target network, which enhances the performance outcome.

### **BW-VGGish model**

BW-VGGish model is a developed version of the VGGish model. It is pre-trained on the BWEB dataset in addition to its pre-training on the AudioSet dataset. Unlike the VGGish model which pretrained on only the AudioSet dataset. The training on the BWEB dataset has been done on to make the model gain more experience in classifying this type of music that contains brainwave entrainment beats inside it.

The BW-VGGish model architecture is shown in Fig. 6, where the VGGish model parameters transferred to the BW-VGGish network while training of the VGGish model on BWEB dataset, then the model parameters transferred to the target network while the training of the BW-VGGish model on a BWMM dataset.

### **BW-YAMNET model**

BW-YAMNET model is a developed version of the YAMNET model. It is pre-trained on the BWEB dataset in addition to its pre-training on the audioset dataset. Unlike the YAMNET model, which is trained on only the AudioSet dataset, the training on a BWEB dataset has been done to give the model more experience in classifying this type of music that contains brainwave entertainment beats inside it.

In the BW-YAMNET the YAMNET model parameters are transferred to the BW-YAMNET network while training the YAMNET model on BWEB dataset, then the model parameters are transferred to the target network while training the BW-YAMNET model on a BWMM dataset.

## **EXPERIMENTAL RESULTS**

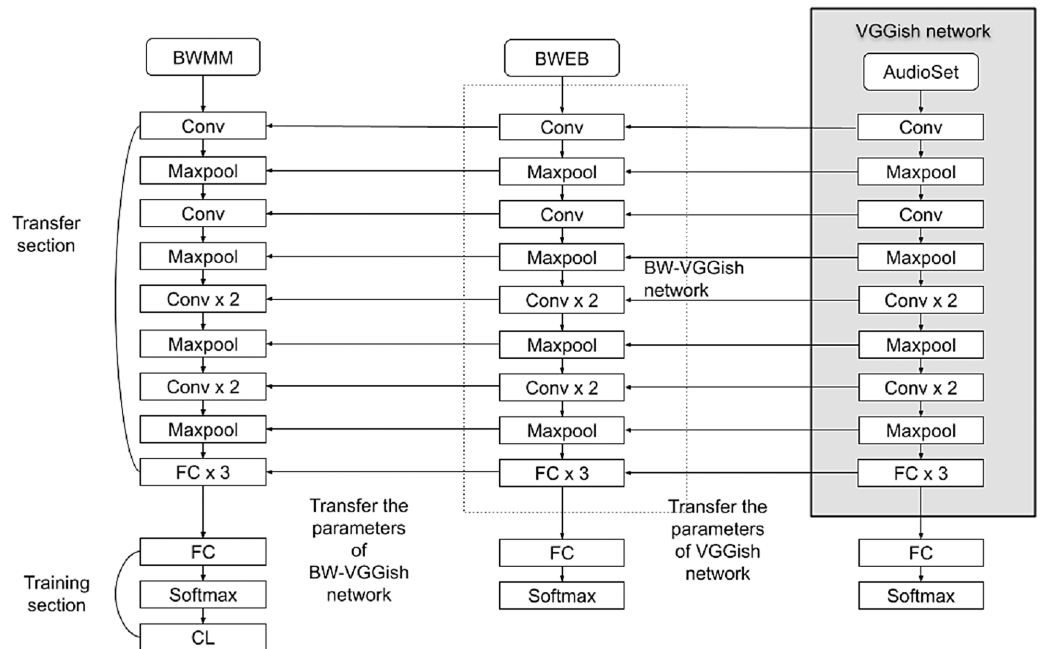
To test the generalization ability of the model, the experiments were run three times, and the mean of the results was calculated as the average precision. In each experiment, the k-fold cross-validation (*Haq, 2021; Shi et al., 2019*) with  $k = 5$  was used, where the data was separated into five subgroups at random and the ratio of samples in each subset was the same for each category. Four subsets were taken as the training data in turn, and one subset as the testing data. From each training subset, 20% of the data was taken for validation. All the experiments in this article have been done on windows 10.1 with a hardware core of I7, 16 GB of RAM using MATLAB R2020b (*The MathWorks, Inc., 2020*).

The hyper-parameters that were utilized to train the models are:

- The learning rate equals 0.0001
- Mini-batch size equals 32
- Optimizer equals adam
- Training using one GPU

After that, we verified the effectiveness of the proposed models using data that was outside the data set used in the training process. We tested music files that were used and





**Figure 6** BW-VGGish model architecture.

Full-size DOI: [10.7717/peerjcs.1642/fig-6](https://doi.org/10.7717/peerjcs.1642/fig-6)

classified in previous research using EEG, then compared the results of their classification with our results, as will be discussed in the Proposed Models Verification section.

### Performance evaluation metrics of model

Accuracy, recall, precision, F-score, and the confusion matrix are all used to evaluate the model's performance.

$$Accuracy = \frac{(TN + TP)}{(TP + TN + FP + FN)} \quad (1)$$

$$Precision = \frac{TP}{(TP + FP)} \quad (2)$$

$$Sensitivity/Recall = \frac{TP}{(TP + FN)} \quad (3)$$

$$F - score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (4)$$

where TP indicates how many times the model properly predicts the positive category, TN indicates how many times the model has accurately predicted the negative category, FP indicates the number of times the model has predicted the positive category incorrectly, and FN indicates how many times the model has predicted the negative category erroneously.

**Table 1** Model comparison between VGGish, YAMNET, and BDLSTM.

Model	Precision (%)	Recall (%)	F-score (%)	Accuracy (%)
VGGish	99.12	99.33	99.22	98.5
YAMNET	100	98%	98.98	98.4
AS-BDLSTM	92	90.67	91.33	91.1
BDLSTM	89	92	90.47	90.2
Unidirectional - LSTM	86	90.6	88.24	88.4

**Table 2** Model comparison between VGGish and BW-VGGish with the BWMM dataset.

Model	Precision (%)	Recall (%)	F-score (%)	Accuracy (%)
VGGish	93.9	93.75	93.82	93.3
BW-VGGish	95	95	95	94.6

**Table 3** Model comparison between YAMNET and BW-YAMNET with the BWMM dataset.

Model	Precision (%)	Recall (%)	Accuracy (%)
YAMNET	93.66	93.83	93.3
BW-YAMNET	94.66	94.66	94.6

## RESULTS DISCUSSION

Table 1 shows the performance metrics for the unidirectional LSTM, BDLSTM, AS-BDLSTM, VGGish, and YAMNET models (Sadek, Khalifa & Elfattah, 2021). The AS-BDLSTM model achieved higher percentages for precision, and accuracy with 92%, and 91.1%, respectively when compared with the BDLSTM model. On the other hand, it has a lower recall than the BDLSTM model, but this can be bypassed as we care more about the precision value. However, our greatest concern is the correct detection of the entrainment beats inside the music files and we can disregard mistakenly labeling a music file as containing entrainment beats. Although the AS-BDLSTM model improved efficiency, it required a long time for training and a large memory space, which made us move on to thinking about using the VGGish and YAMNET models.

The VGGish and YAMNET models achieved higher percentages for precision, recall, and accuracy when compared with the BDLSTM model. VGGish and YAMNET models provide similar performance metrics, but the YAMNET model is recommended for usage with mobile apps because of the nature of its structure. It is a light - parameter model that needs less power, which is a requirement for mobile applications.

Tables 2 and 3 show the performance metrics for the VGGish, BW-VGGish, YAMNET, and BW-YAMNET models when used with the BWMM dataset in the classification of the music files depending on their effect on the brain waves.

The performance of the models can be visualized using the confusion matrix. Figures 7 and 8 represent the confusion matrix for the transfer learning of the VGGish and YAMNET

Output Class	brainwave entrainment	148 49.3%	1 0.3%	99.3% 0.7%
	no entrainment	2 0.7%	149 49.7%	98.7% 1.3%
		98.7% 1.3%	99.3% 0.7%	99.0% 1.0%
		brainwave entrainment	no entrainment	
		Target Class		

**Figure 7** Confusion matrix of the VGGish model with the BWEB dataset.

[Full-size](#)  DOI: 10.7717/peerjcs.1642/fig-7

Output Class	brainwave entrainment	147 49.0%	0 0.0%	100% 0.0%
	no entrainment	3 1.0%	150 50.0%	98.0% 2.0%
		98.0% 2.0%	100% 0.0%	99.0% 1.0%
		brainwave entrainment	no entrainment	
		Target Class		

**Figure 8** Confusion matrix of the YAMNET model with the BWEB dataset.

[Full-size](#)  DOI: 10.7717/peerjcs.1642/fig-8

alpha	37 15.4%	1 0.4%	2 0.8%	0 0.0%	0 0.0%	0 0.0%	92.5% 7.5%
beta	0 0.0%	38 15.8%	0 0.0%	1 0.4%	0 0.0%	2 0.8%	92.7% 7.3%
delta	2 0.8%	1 0.4%	38 15.8%	0 0.0%	1 0.4%	0 0.0%	90.5% 9.5%
gamma	1 0.4%	0 0.0%	0 0.0%	38 15.8%	1 0.4%	2 0.8%	90.5% 9.5%
no effect	0 0.0%	0 0.0%	0 0.0%	0 0.0%	38 15.8%	0 0.0%	100% 0.0%
theta	0 0.0%	0 0.0%	0 0.0%	1 0.4%	0 0.0%	36 15.0%	97.3% 2.7%
	92.5% 7.5%	95.0% 5.0%	95.0% 5.0%	95.0% 5.0%	95.0% 5.0%	90.0% 10.0%	93.8% 6.3%
	alpha	beta	delta	gamma	no effect	theta	
	Target Class						

**Figure 9** Confusion matrix of the VGGish model with the BWMM dataset.

Full-size  DOI: [10.7717/peerjcs.1642/fig-9](https://doi.org/10.7717/peerjcs.1642/fig-9)

models with the BWEB dataset, while Figs. 9–12 represent the VGGish, BW-VGGish, YAMNET, and BW-YAMNET models when used with the BWMM dataset.

## PROPOSED MODELS VERIFICATION

In this section, we will review the experiments and results that we carried out to verify the effectiveness of the proposed models using data that is outside the data set used in the training process. We tested music files that were used and classified in previous research, then compared the results of their classification with our results. In this experiment, we downloaded two music files, which are “Sonata for Two Pianos in D major, K.448” for classical music and “Napalm Death Procrastination on The Empty Vessel” for metal music. These music files were used by *Tai & Lin (2018)* to study the effect of classical and metal music on brain waves. In *Tai & Lin (2018)*, nine volunteers listened to the two music files for 10 min. Then their brain waves were collected for 8 min only by removing the data for the first and last minute. When entering these music files into our models, we found that “Sonata for Two Pianos in D major, K.448” affected the delta waves, while “Napalm Death

**Confusion Matrix:**

<b>Output Class</b>	alpha	36 15.0%	2 0.8%	1 0.4%	0 0.0%	0 0.0%	0 0.0%	92.3% 7.7%
	beta	1 0.4%	37 15.4%	0 0.0%	0 0.0%	0 0.0%	1 0.4%	94.9% 5.1%
	delta	1 0.4%	0 0.0%	39 16.3%	0 0.0%	0 0.0%	0 0.0%	97.5% 2.5%
	gamma	1 0.4%	0 0.0%	0 0.0%	39 16.3%	0 0.0%	2 0.8%	92.9% 7.1%
	no effect	0 0.0%	0 0.0%	0 0.0%	0 0.0%	40 16.7%	0 0.0%	100% 0.0%
	theta	1 0.4%	1 0.4%	0 0.0%	1 0.4%	0 0.0%	37 15.4%	92.5% 7.5%
			90.0% 10.0%	92.5% 7.5%	97.5% 2.5%	97.5% 2.5%	100% 0.0%	92.5% 7.5%
		<b>Target Class</b>						
		alpha	beta	delta	gamma	no effect	theta	

**Figure 10** Confusion Matrix of the BW-VGGish Model with the BWMM Dataset.

Full-size  DOI: [10.7717/peerjcs.1642/fig-10](https://doi.org/10.7717/peerjcs.1642/fig-10)

Procrastination on The Empty Vessel” affected the beta waves. And this result is the same that has been proven in *Tai & Lin (2018)* as shown in [Table 4](#).

## CONCLUSIONS

Music has become a significant part of people’s lives as a result of the growing quantity of songs available and the widespread usage of the Internet by people of all ages. They listen to music at various times throughout the day and engage in diverse activities. A new kind of music has emerged in recent years that combines brainwave entrainment beats that can alter human moods and induce seizures in epileptic sufferers. This kind of music has grown rapidly, as evidenced by the fact that there are multiple sites that sell these audio files for a price, in addition to their free availability on YouTube and the SoundCloud server.

In this article, we present models to be employed in two key categorization processes. The VGGish and YAMNET models are proposed to be employed in the first classification process, which is the categorization of music files based on the presence of entrainment

<b>Output Class</b>	<b>alpha</b>	36 15.0%	2 0.8%	1 0.4%	1 0.4%	0 0.0%	0 0.0%	90.0% 10.0%
	<b>beta</b>	0 0.0%	37 15.4%	0 0.0%	1 0.4%	0 0.0%	0 0.0%	97.4% 2.6%
	<b>delta</b>	2 0.8%	1 0.4%	38 15.8%	0 0.0%	2 0.8%	0 0.0%	88.4% 11.6%
	<b>gamma</b>	1 0.4%	0 0.0%	1 0.4%	37 15.4%	0 0.0%	1 0.4%	92.5% 7.5%
	<b>no effect</b>	1 0.4%	0 0.0%	0 0.0%	0 0.0%	38 15.8%	0 0.0%	97.4% 2.6%
	<b>theta</b>	0 0.0%	0 0.0%	0 0.0%	1 0.4%	0 0.0%	39 16.3%	97.5% 2.5%
			90.0% 10.0%	92.5% 7.5%	95.0% 5.0%	92.5% 7.5%	95.0% 5.0%	97.5% 2.5%
		<b>alpha</b>	<b>beta</b>	<b>delta</b>	<b>gamma</b>	<b>no effect</b>	<b>theta</b>	
		<b>Target Class</b>						

**Figure 11** Confusion matrix of the YAMNET model with the BWMM dataset.

Full-size  DOI: [10.7717/peerjcs.1642/fig-11](https://doi.org/10.7717/peerjcs.1642/fig-11)

beats within them. We suggested the BW-VGGish and BW-YAMNET models for the second classification, which is the categorization of music files based on their impact on brain waves.

The VGGish and YAMNET models overcome the BDLSTM model with an accuracy of 98.5% and 98.4%, respectively. BW-VGGish and BW-YAMNET models are modifications of the VGGish and YAMNET models. This modification happened by pre-training them on the BWEB dataset in addition to the AudioSet dataset to give them additional experience with music files containing brainwave entrainment beats. This knowledge has given them a leg up on the competition when it comes to categorizing music files based on their effect on brain waves. Both the BW-VGGish and BW-YAMNET models achieved an accuracy of 94.6%.



<b>alpha</b>	<b>37</b> 15.4%	<b>2</b> 0.8%	<b>1</b> 0.4%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>0</b> 0.0%	92.5% 7.5%
<b>beta</b>	<b>1</b> 0.4%	<b>38</b> 15.8%	<b>1</b> 0.4%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>1</b> 0.4%	92.7% 7.3%
<b>delta</b>	<b>1</b> 0.4%	<b>0</b> 0.0%	<b>38</b> 15.8%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>0</b> 0.0%	97.4% 2.6%
<b>gamma</b>	<b>1</b> 0.4%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>38</b> 15.8%	<b>1</b> 0.4%	<b>2</b> 0.8%	90.5% 9.5%
<b>no effect</b>	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>39</b> 16.3%	<b>0</b> 0.0%	100% 0.0%
<b>theta</b>	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>0</b> 0.0%	<b>2</b> 0.8%	<b>0</b> 0.0%	<b>37</b> 15.4%	94.9% 5.1%
	92.5% 7.5%	95.0% 5.0%	95.0% 5.0%	95.0% 5.0%	97.5% 2.5%	92.5% 7.5%	<b>94.6%</b> 5.4%
	<b>alpha</b>	<b>beta</b>	<b>delta</b>	<b>gamma</b>	<b>no effect</b>	<b>theta</b>	
	<b>Target Class</b>						

**Figure 12** Confusion matrix of the BW-YAMNET model with the BWMM dataset.

Full-size  DOI: [10.7717/peerjcs.1642/fig-12](https://doi.org/10.7717/peerjcs.1642/fig-12)

**Table 4** A comparison between the results of our models and previous research.

Music file	BW-VGGish /BW-YAMNET	Previous research
Sonata for Two Pianos in D major, K.448	Delta waves	Delta waves
Napalm Death Procrastination On The Empty Vessel	Beta waves	Beta waves

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

The authors received no funding for this work.

### Competing Interests

Rowayda A. Sadek is an Academic Editor for PeerJ.

### Author Contributions

- Rowayda A. Sadek conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.

- Alaa A. Khalifa performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Marwa M.A. Elfattah analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The data is available at GitHub and Zenodo:

- <https://github.com/AlaaKhalif/BWEB.git>

- AlaaKhalif. (2023). AlaaKhalif/BWEB: Deep learning Binary /Multi classification for music's brainwave entrainment beats (BWMM). Zenodo. <https://doi.org/10.5281/zenodo.8341595>

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.1642#supplemental-information>.

## REFERENCES

- Bali S, Tyagi SS. 2020.** Evaluation of transfer learning techniques for classifying small surgical dataset. In: *2020 10th international conference on cloud computing, data science & engineering (confluence)* DOI [10.1109/confluence47617.2020](https://doi.org/10.1109/confluence47617.2020).
- Barman R, Deshpande S, Agarwal S, Inamdar U. 2019.** Transfer learning for small dataset. In: *National conference on machine learning, 26th 2019*. 132–137.
- Beauchene C, Abaid N, Moran R, Diana RA, Leonessa A. 2016.** The effect of binaural beats on visuospatial working memory and cortical connectivity. *PLOS ONE* **11(11)**:e0166630 DOI [10.1371/journal.pone.0166630](https://doi.org/10.1371/journal.pone.0166630).
- Becher AK, Höhne M, Axmacher N, Chaieb L, Elger CE, Fell J. 2014.** Intracranial electroencephalography power and phase synchronization changes during monaural and binaural beat stimulation. *European Journal of Neuroscience* **41(2)**:254–263 DOI [10.1111/ejn.12760](https://doi.org/10.1111/ejn.12760).
- Engelbregt H, Barmentlo M, Keeser D, Pogarell O, Deijen JB. 2021.** Effects of binaural and monaural beat stimulation on attention and EEG. *Experimental Brain Research* **239**:2781–2791 DOI [10.1007/s00221-021-06155-z](https://doi.org/10.1007/s00221-021-06155-z).
- Engelbregt H, Meijburg N, Schulten M, Pogarell O, Deijen JB. 2019b.** The effects of binaural and monaural beat stimulation on cognitive functioning in subjects with different levels of emotionality. *Advances in Cognitive Psychology* **15(3)**:199–207 DOI [10.5709/acp-0268-8](https://doi.org/10.5709/acp-0268-8).
- Gemmeke JF, Ellis DPW, Freedman D, Jansen A, Lawrence W, Moore RC, Plakal M, Ritter M. 2017.** Audio set: an ontology and human-labeled dataset for audio events. In: *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. Piscataway: IEEE DOI [10.1109/icassp.2017.7952261](https://doi.org/10.1109/icassp.2017.7952261).

- Haq M.** 2021. Music genre classification using deep learning techniques. *International Journal for Research in Applied Science and Engineering Technology* **9(VI)**:4157–4160 DOI [10.22214/ijraset.2021.35965](https://doi.org/10.22214/ijraset.2021.35965).
- Harjoseputro Y, Yuda IP, Danukusumo KP.** 2020. MobileNets: efficient convolutional neural network for identification of protected birds. *International Journal on Advanced Science, Engineering and Information Technology* **10(6)**:2290 DOI [10.18517/ijaseit.10.6.10948](https://doi.org/10.18517/ijaseit.10.6.10948).
- Hershey S, Chaudhuri S, Ellis DPW, Gemmeke JF, Jansen A, Moore RC, Plakal M, Platt D, Saurous RA, Seybold B, Slaney M, Weiss RJ, Wilson K.** 2017. CNN architectures for large-scale audio classification. In: *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)*. Piscataway: IEEE DOI [10.1109/icassp.2017.7952132](https://doi.org/10.1109/icassp.2017.7952132).
- Jirakittayakorn N, Wongsawat Y.** 2017. Brain responses to a 6-Hz binaural beat: effects on general theta rhythm and frontal midline theta activity. *Frontiers in Neuroscience* **11**:365 DOI [10.3389/fnins.2017.00365](https://doi.org/10.3389/fnins.2017.00365).
- Johnson JM, Durrant SJ.** 2021. Commentary: SWS brain-wave music may improve the quality of sleep: an EEG study. *Frontiers in Neuroscience* **15**:609169 DOI [10.3389/fnins.2021.609169](https://doi.org/10.3389/fnins.2021.609169).
- Khalifa AA, Sadek RA, Elfattah MMA.** 2022. Brainwave entrainment beats (BWEB). IEEE Dataport. DOI [10.21227/30m9-9f90](https://doi.org/10.21227/30m9-9f90).
- Repovš G.** 2010. Dealing with noise in EEG recording and data analysis. *Informatica Medica Slovenica* **15(1)**:18–25.
- Ross B, Lopez MD.** 2020. 40-Hz Binaural beats enhance training to mitigate the attentional blink. *Scientific Reports* **10**:7002 DOI [10.1038/s41598-020-63980-y](https://doi.org/10.1038/s41598-020-63980-y).
- Sadek RA, Khalifa AA, Elfattah MM.** 2021. Detecting digital stimulant music using bidirectional deep long short term memory. In: *2021 9th international Japan-Africa conference on electronics, communications, and computations (JAC-ECC)* DOI [10.1109/jac-ecc54461.2021.9691449](https://doi.org/10.1109/jac-ecc54461.2021.9691449).
- Sadek RA, Khalifa AA, Elfattah MMA.** 2022. BrainWave musical manipulation (BWMM). IEEE Dataport. DOI [10.21227/1qvb-zn13](https://doi.org/10.21227/1qvb-zn13).
- Shi L, Du K, Zhang C, Ma H, Yan W.** 2019. Lung sound recognition algorithm based on VGGish-BiGRU. *IEEE Access* **7**:139438–139449 DOI [10.1109/access.2019.2943492](https://doi.org/10.1109/access.2019.2943492).
- Shumov DE, Arsen'ev GN, Sveshnikov DS, Dorokhov VB.** 2017. Comparative analysis of the effect of stimulation with a binaural beat and similar kinds of sounds on the falling asleep process: a brief note. *Moscow University Biological Sciences Bulletin* **72(1)**:33–36 DOI [10.3103/s0096392517010047](https://doi.org/10.3103/s0096392517010047).
- Tai SK, Lin YK.** 2018. The influence of different kinds of music on brainwave signals. In: *2018 IEEE international conference on advanced manufacturing (ICAM)*. Piscataway: IEEE, DOI [10.1109/amcon.2018.8614971](https://doi.org/10.1109/amcon.2018.8614971).
- The MathWorks, Inc..** 2020. *MATLAB R2020b*. Natick: Mathworks, Inc. Available at <https://www.mathworks.com>.

- Wahbeh H, Calabrese C, Zwickey H. 2007.** Binaural beat technology in humans: a pilot study to assess psychologic and physiologic effects. *The Journal of Alternative and Complementary Medicine* **13**(1):25–32 DOI [10.1089/acm.2006.6196](https://doi.org/10.1089/acm.2006.6196).
- Yusim A, Grigaitis J. 2020.** Efficacy of binaural beat meditation technology for treating anxiety symptoms. *Journal of Nervous & Mental Disease* **208**(2):155–160 DOI [10.1097/nmd.0000000000001070](https://doi.org/10.1097/nmd.0000000000001070).