Music Signal Processing and Optimization Algorithm Based on Wavelet Neural Network Model

Jie Tian[1] and Bing Tian[2]

[1]Faculty of Media and Music, Hainan Vocational University of Science and Technology, Hainan 571126, China, musictjie@163.com
[2]School of Music and Dance, Zhengzhou University of Science and Technology, Zhengzhou 450000, China, music.tian_bj@zit.edu.cn

Corresponding author: Bing Tian, music.tian_bj@zit.edu.cn

Abstract. In this article, the application of CAD (Computer Aided Design) technology and WNN (Wavelet Neural Network) model in music signal processing is studied, and the principle, method, and optimization ability to improve processing efficiency are explored. A music signal processing and optimization algorithm based on CAD and WNN model is constructed, and simulation experiments verify its effectiveness. Through comprehensive analysis, it is found that this model has good performance in music classification tasks, and its accuracy and processing efficiency are obviously higher than those of SVM (Support Vector Machine) and RNN (Recurrent Neural Network) algorithms. This is mainly due to the powerful feature detection and classification capabilities of CAD and WNN models, which can better capture the complex patterns and features in music signals. In contrast, SVM and RNN algorithms have some limitations when dealing with music classification tasks and can't make full use of the information in audio signals. Through research, this article provides a new idea and method for researchers in the field of music signal processing, which promotes the growth of this field.

Keywords: Computer-Aided Design; Wavelet Neural Network; Music Signal Processing; Optimization Algorithm
DOI: https://doi.org/10.14733/cadaps.2024.S18.222-238

1 INTRODUCTION

In recent years, with the rapid growth of digital music technology, music signal processing has become a hot research field. The application of CAD technology and the WNN model in music signal processing has gradually attracted the attention of researchers. Wavelet neural network is a neural network model based on wavelet transform, which can be used to process various nonlinear and non-stationary signals. In music signal processing, wavelet neural networks can be used for feature extraction, classification, and grading of audio data. By using wavelet transform to transform and decompose audio data, feature information is extracted, and then neural network technology is used to classify and grade the extracted features, achieving automatic sorting and grading of audio data. Bishop et al. [1] proposed an interactive music signal processing method based on CAD and wavelet neural network models for identifying and analyzing body movement patterns in music duo performances. This method combines the advantages of CAD models and wavelet neural network technology to achieve interpretable and efficient deep learning of body movement patterns in music duo performances. By using wavelet neural network technology to process and analyze the extracted features further, more abstract and meaningful feature information can be extracted. Using neural network technology to classify and segment the extracted features, achieving automatic recognition and processing of body movement patterns in music duo performances. WNN is a model based on wavelet transform and NN (Neural Network) theory, which has strong learning and optimization ability. The theory of music rhythm is an important tool for studying music rhythm and rhythm. It reveals the internal structure and laws of music by analyzing elements such as rhythm, beat, and time sign. However, traditional rhythm theory analysis methods are often abstract, and it is difficult to visually present the rhythm and rhythm of music. Therefore, Calilhanna [2] proposed a visualization and sound analysis method based on ski hills and loop diagrams to understand and present the theory of rhythm in Ogene Bunch music more intuitively. This is a method of presenting music rhythm and rhythm by drawing a ski hill diagram. In the ski hill plot, each small peak represents a beat or beat, and the height of the peak represents the intensity or sense of rhythm of that beat. By observing the ski hill plot, we can intuitively understand the rhythm and rhythmic structure of music. During the music creation process, Ogene Bunch can use visualization and sound analysis methods based on ski hill and loop diagrams to guide her own creations. He can intuitively understand the rhythm and rhythmic structure of music by drawing ski hill diagrams, thereby adjusting his creative ideas. At the same time, he can also use loop diagrams to analyze and adjust the loop structure and changing patterns of music in order to achieve more diverse musical expression. CAD technology can effectively improve the efficiency of music signal processing, while the WNN model has strong learning and optimization ability, which can deeply analyze and optimize music signals. Collective music listening is an important cultural activity that involves the process of multiple people appreciating and understanding music together. In collective music listening, people interact with music through physical movements, emotional expression, and other means, forming a unique energy field. Traditional collective music listening analysis methods often struggle to capture and enhance this movement energy accurately. Therefore, how to use deep learning techniques to analyze and enhance the movement energy in collective music listening has become an important research topic. CAD model is a computer-aided design model that can be used to process various complex data structures. In music signal processing, CAD models can be used for visualization, editing, and analysis of audio data. Through CAD models, we can convert audio data into visual graphics, making it convenient for users to edit and analyze. In addition, CAD models can also be used for audio data preprocessing and feature extraction, providing strong support for subsequent classification and grading. Dotov et al. [3] designed an interpretable and efficient deep-learning model to enhance motion energy in collective music listening. This model achieves accurate recognition and enhancement of movement energy in collective music listening by analyzing grooves and visual social cues in music signals. The experimental results indicate that the model has achieved good performance in collective music listening tasks.

CAD technology is a computer-aided design technique widely used in various fields. In music signal processing, CAD technology can be used for visualization, editing, and analysis of audio data. Through CAD technology, we can convert audio data into visual graphics, making it convenient for users to edit and analyze. In addition, CAD technology can also be used for audio data preprocessing and feature extraction, providing strong support for subsequent classification and grading. Farhadi et al. [4] introduced an intelligent music signal processing technology integration system based on CAD and wavelet neural network models, which can achieve automatic sorting and grading of music signals. By combining CAD technology and wavelet neural network models, the system can accurately extract features from music signals and perform intelligent classification and grading. The experimental results indicate that the integrated system has broad application prospects in the field of music signal processing. Wavelet neural network model is a machine learning method based on wavelet transform and neural network technology, which can be used to process various nonlinear and non-stationary signals. In music signal processing, wavelet neural network models can be used for feature extraction, classification, and grading of audio data. Through wavelet transform, we can transform audio data into wavelet coefficients at different scales, thereby extracting local features from audio signals. Then, neural network technology is used to classify and grade the extracted features, achieving automatic sorting and grading of audio data. However, at present, there is relatively little research on music signal processing and optimization based on CAD and WNN models, so it is necessary to conduct in-depth research. With the popularization of digital music, people's demands for music quality and user experience are becoming increasingly high. In order to meet these needs, music signal processing technology has been widely applied. Music signal processing involves the collection, encoding, decoding, analysis, and optimization of audio data. Traditional music signal processing methods often have some limitations, such as poor processing performance, high computational complexity, and lack of personalized recommendations. Therefore, how to optimize music signal processing using information technology has become an important research topic. Fu et al. [5] explored the revolution of music signal processing based on information technology and introduced how to optimize music signal processing using the Analytic Hierarchy Process (AHP) and Objective Programming Method (TOPSIS). With the continuous development of information technology, music signal processing has become an important component of the music industry. This article will introduce the application of AHP and TOPSIS in music signal processing and analyze their impact on music quality, user experience, and music recommendation systems. By setting multiple optimization objectives, such as audio quality, compression rate, decoding speed, etc., and considering various constraints, such as storage space, computing resources, etc., TOPSIS can solve for the optimal encoding and decoding scheme. These solutions can be used to improve audio quality and user experience while reducing computational complexity and resource consumption. The purpose of this study is to deeply explore the application of CAD and WNN models in music signal processing and optimization and to improve the efficiency and accuracy of music signal processing by constructing an algorithm based on CAD and WNN models.

Automatic music classification is an important component of the music industry, which can help musicians, artists, and consumers more conveniently manage and appreciate music. However, traditional music classification methods often have some limitations, such as poor processing performance, high computational complexity, and lack of personalized recommendations. Therefore, how to optimize the automatic music classification system using computer technology has become an important research topic. Ge et al. [6] explored the optimization method of a computer-aided design system for music automatic classification based on feature analysis. By deeply analyzing audio features and utilizing advanced computer technology, we can design a more efficient, accurate, and personalized music classification system. Integrate different audio features, such as pitch, timbre, rhythm, etc., to obtain a more comprehensive description of music features. By using machine learning algorithms to select extracted features, redundant and irrelevant features are removed, improving the quality and representativeness of the features. Based on the dynamic changes of audio data, dynamically adjust the method and parameters of feature extraction to improve adaptability and robustness. The use of deep learning techniques for feature extraction and classification of audio data can handle complex nonlinear problems and improve classification accuracy and efficiency. The significance of this investigation can be predominantly observed through three primary dimensions, as detailed below: (1) Theoretical significance: This study introduces CAD technology and the WNN model into music signal processing and optimization, which is helpful to promote the theoretical development in this field. (2) Practical significance: The algorithm constructed in this study can be applied to the practical application of music signal processing, which improves the processing efficiency and accuracy and has practical application value. (3) Academic significance: This study provides a new idea and method for researchers in the field of music signal processing, which is helpful to promote research and development in this field.

The specific innovations are as follows: (1) In this article, WNN is used for feature detection and representation learning of music signals, and the time-frequency features and emotional features of music signals are extracted, and CAD technology is used for visual display and analysis. Compared with the traditional machine learning algorithm, this model has stronger feature detection and classification ability and can better capture the complex patterns and features in music signals. (2) In this article, the music data sets of GTZAN and Million Song Dataset are selected for experimental verification. This data set contains various types and styles of music fragments, which is suitable for the research of music classification tasks. The generalization ability of the model can be better illustrated by using open data sets for experimental verification.

This article consists of six sections. The first section introduces the research background, purpose, significance, problems and assumptions, as well as research methods and technical routes. The second section combs and evaluates the application status of CAD technology and the WNN model in music signal processing and optimization. The third section expounds on the basic principles and technologies of CAD technology and the WNN model and the feasibility of their application in music signal processing and optimization. The fourth section describes in detail the construction of the music signal processing and optimization algorithm and the realization process of parameter setting based on CAD and WNN models. The fifth section introduces the environment, data set, design, implementation scheme, results and analysis of the simulation experiment, as well as discussion and explanation. The sixth section summarizes the research conclusions, discusses the limitations and shortcomings of the research, and puts forward future research prospects and suggestions.

2 RELATED WORK

Multidimensional scaling is a method used to describe high-dimensional data structures, which can be used to process multiple features in music works. In music information visualization, we can integrate multiple features into a two-dimensional plane through multidimensional scales, thereby more intuitively displaying the full picture of music works. For example, we can use features such as pitch, rhythm, and intensity as multidimensional dimensions to map musical works onto a two-dimensional plane. This visualization approach can help us better understand and analyze the style and emotions of music works. In music information visualization, Georges and Seckin et al. [7] present the structure and style of music works by constructing network graphs. For example, we can use melody lines as nodes and the relationships between notes as edges to construct a network diagram of musical works. This visualization approach can help us better understand and analyze the structure and style of music works. To verify the effectiveness of the above method, we conducted a large number of experiments. The experimental results show that the music information visualization method combining network graphs, multidimensional scales, and support vector machines has achieved good results in the discovery of classical composers. Through visualization technology, we can more intuitively understand and analyze the structure and style of music works. Through support vector machine classification and recognition techniques, we can more accurately discover and explore the potential value and influence of classical composers. With the continuous advancement of technology, multimodal education has become an important feature of modern education. Music, as an art form with profound emotions and cultural connotations, has an educational value that cannot be ignored. However, traditional music education methods often only focus on the auditory modality, neglecting the role of other sensory modalities, such as vision and touch. Therefore, Gorbunova and Plotnikov [8] proposed a music computer technology based on CAD and wavelet neural network models to achieve multimodal music perception, thereby improving students' comprehensive understanding and perception of music. CAD not only has extensive applications in the field of design but also has unique advantages in music education. By using CAD technology, we can transform abstract information, such as music symbols and scores, into intuitive visual images, helping students better understand the structure and elements of music. In addition, CAD can also be used to create interactive music learning environments, allowing students to perceive the rhythm and melody of music through visual and tactile interaction. Wavelet neural network is a new type of neural network model with excellent time-frequency characteristics and nonlinear approximation ability. In music perception, wavelet neural networks can effectively extract and process feature information of music, such as rhythm, melody, timbre, etc. By training wavelet neural networks, we can achieve automatic classification, recognition, and generation of music, thereby providing rich teaching resources and tools for music education. With the popularization of the Internet and technological progress, streaming media services have become one of the main forms of digital music consumption. Streaming services provide users with a convenient and fast music playback experience by transmitting audio data in real time. At the same time, streaming media services also have flexible pricing strategies and personalized recommendation functions, making digital music consumption more concentrated. However, the competition for streaming services is becoming increasingly fierce, and how to improve user experience and reduce costs is an important issue faced by streaming service providers. Im et al. [9] explored the impact of CAD and wavelet neural network models on the concentration of digital music consumption by studying their applications in streaming services. The article first introduces the basic concepts and principles of CAD and wavelet neural network models, then analyzes the advantages and challenges of streaming services in digital music consumption, and finally explores how CAD and wavelet neural network models affect the concentration of digital music consumption. In streaming services, CAD can be used to optimize the encoding and decoding process of audio data, improving audio quality. Wavelet neural network is a neural network model based on wavelet transform, which has multi-scale analysis ability and good time-frequency characteristics. In streaming media services, wavelet neural networks can be used for the classification, recognition, and synthesis of audio signals, improving the accuracy and efficiency of audio processing.

Music is an important feature in-car entertainment systems, but users often need to manually select music genres or search for keywords to find their favourite music when enjoying music. This not only increases the operational burden on users but also limits personalized music recommendations. Therefore, how to achieve automatic music genre recognition in-car entertainment systems has become an important research topic. In car infotainment systems, audio data is the main source of music signals. In order to improve the accuracy of music recognition, it is necessary to preprocess audio data. CAD technology can extract useful features from audio signals by compressing, denoising, and enhancing audio data, providing a better data foundation for subsequent music genre recognition. To meet this requirement, Jakubec and Chmulik [10] proposed an automatic music genre recognition method for in-car entertainment systems based on CAD and wavelet neural network models. This method preprocesses audio data through CAD technology and then uses wavelet neural network models to learn and classify audio features, thereby achieving automatic genre recognition of music. Using wavelet neural network models to learn and classify audio features, thereby achieving automatic genre recognition of music. This method can provide personalized music recommendation services for in-car entertainment systems, improving user satisfaction with the music experience. Music audio is often compressed during transmission and storage to reduce file size and improve transmission efficiency. However, this compression can lead to loss of music audio quality, such as sound distortion and increased noise. To address this issue, we propose a method for randomly recovering recompressed music audio using generative adversarial networks. Lattner and Nistal [11] proposed a method for randomly recovering recompressed music audio using Generative Adversarial Networks (GANs). This method trains a generator network and a discriminator network, where the generator network is used to generate uncompressed music audio and the discriminator network is used to distinguish the generated music audio. The experimental results show that this method can effectively restore the quality of recompressed music audio and improve its listening experience and subjective rating. Generative Adversarial Network (GAN) is a deep learning model consisting of a generator network and a discriminator network. The generator network is used to generate new data samples, and the discriminator network is used to distinguish between real data and generated data. During the training process, the generator and discriminator constantly engage in confrontation until reaching a balanced state. At this point, the samples generated by the generator can be distorted, while the discriminator cannot distinguish between real data and generated data.

Music signal processing is an important component of the music industry, which involves the analysis, processing, and annotation of audio data. Traditional music signal processing methods often have some limitations, such as poor processing performance, high computational complexity, and lack of personalized recommendations. Liu et al. explored the application of model network optimization in music signal processing and machine learning. By delving into the characteristics of music signals and combining advanced machine learning techniques, we can design more efficient, accurate, and stable music signal processing models. Liu et al. [12] will introduce the application of model network optimization in music signal processing, including feature extraction, model training, and optimization, and analyze its impact on the performance of music signal processing. By optimizing the structure and parameters of neural network models, the accuracy and efficiency of feature extraction can be improved. For example, convolutional neural networks (CNNs) can be used to extract local features of audio signals, or recurrent neural networks (RNNs) can be used to extract temporal features of audio signals. With the continuous development of digital music technology, music signal processing has become an important field of research. Music signal processing involves the collection, encoding, decoding, analysis, and optimization of audio data. In order to improve audio quality and user experience, it is necessary to use effective algorithms to process and optimize music signals. CAD and wavelet neural network models are two widely used techniques in music signal processing in recent years, each with its own advantages and characteristics. Maba [13] introduced music signal processing and optimization algorithms based on CAD and wavelet neural network models. By using CAD technology to preprocess audio data, noise interference can be reduced, and audio quality can be improved. Meanwhile, CAD can also be used for encoding and decoding audio data, enabling the transmission and playback of audio data between different devices and platforms. In music signal processing, wavelet neural networks can be used to extract time-frequency features of audio signals and, through training and optimizing models, achieve classification and recognition of audio data. By learning and comparing different genres of music, wavelet neural networks can automatically identify the genre types of music and provide personalized recommendation services for music recommendation systems. Meanwhile, wavelet neural networks can also be used for the synthesis and generation of audio signals, providing users with more music choices.

Deep learning is a powerful machine learning technique that can be used to process various nonlinear and non-stationary signals. In music signal processing, deep learning can be used for feature extraction, classification, and grading of audio data. Automatic sorting and grading of audio data can be achieved by using neural network technology to classify and grade the extracted features. In addition, deep learning can also be used for classification, segmentation, and recognition in image processing. Using techniques such as convolutional neural networks (CNN) to extract and classify features from image data, achieving automatic recognition and processing of images. Monga et al. [14] explored the application of music signal processing based on CAD network models in signal and image processing. By combining deep learning and CAD network models, we have designed an interpretable and efficient deep learning model for processing music signals and image data. The experimental results show that the model has achieved good performance in both music signal processing and image processing tasks. By using deep learning techniques to process and analyze the extracted features further, more abstract and meaningful feature information can be extracted. Using neural network technology to classify and segment the extracted features, achieving automatic recognition and processing of audio and image data. Visualize the results of classification and segmentation to facilitate user viewing and analysis. Music classification and labelling are important components of the music industry, involving the analysis, processing, and annotation of audio data. Traditional music classification and labelling methods often have some limitations, such as poor processing efficiency, high computational complexity, and lack of personalized recommendations. Therefore, how to use deep learning techniques to optimize music classification and labelling has become an important research topic. Deep learning is a neural network-based machine learning method that can be used to solve various complex nonlinear problems. In music classification and labelling, deep learning can be used to extract audio features, establish classification models, and achieve automated labelling. By training deep learning models, they can be equipped with the ability to distinguish different types and styles of music, thereby achieving automatic classification and labelling of music. Nam et al. [15] explored deep learning techniques for audio-based music classification and labelling and demonstrated through examples how to teach computers to distinguish between rock and Bach. With the continuous development of deep learning technology, music classification and labelling have become important research areas. This article will introduce the application of deep learning in music classification and labelling and analyze its impact on the music industry.

Instrument recognition is an important application direction in music signal processing, aimed at accurately identifying the types of instruments contained in audio signals. With the rise of deep learning, instrument recognition methods based on deep convolutional neural networks (DCNN) have received widespread attention. However, traditional DCNN models still face some challenges when dealing with complex audio signals, such as noise interference and uneven audio quality. Therefore, how to optimize the DCNN model to improve the accuracy and efficiency of instrument recognition has become a focus of research. Solanki and Pandey [16] studied a signal processing method for instrument recognition based on deep convolutional neural networks (DCNN) and optimized its performance. By introducing advanced signal processing techniques and optimization algorithms, we have improved the accuracy and efficiency of DCNN in instrument recognition tasks. The experimental results show that the optimized DCNN model has achieved significant improvement in instrument recognition. DCNN is a powerful deep-learning model that can automatically extract features from audio signals and perform classification. In instrument recognition, DCNN can extract feature information at different levels by performing layer-by-layer convolution and pooling operations on audio signals. This feature information can include pitch, timbre, rhythm, etc., which helps to accurately identify the type of instrument in the audio signal. ECG music signal is a biological signal with rich information content, widely used in medical diagnosis, health monitoring and other fields. However, traditional ECG music signal processing methods often have some limitations, such as poor processing performance, high computational complexity, and lack of personalized recommendations. Therefore, how to use deep learning techniques to classify and analyze ECG music signals has become an important research topic. Wasimuddin et al. [17] proposed a two-dimensional convolutional neural network (2D-CNN) modelling method based on a global average for multi-class ECG music signal analysis. By using global average pooling technology, the number of model parameters is effectively reduced, and the generalization ability of the model is improved. The experimental results show that this method has achieved good performance in multi-class ECG music signal classification tasks. This method uses global average pooling technology to average the global information of each feature map, thereby obtaining a global feature vector. This global feature vector can effectively capture the common features of ECG music signals of different categories, improving the model's generalization ability. By using global average pooling technology, the number of model parameters is effectively reduced, and the generalization ability of the model is improved. The experimental results show that this method has achieved good performance in multi-class ECG music signal classification tasks.

3 THEORETICAL BASIS

3.1 Basic Principles and Techniques Of CAD and WNN Models

CAD is a technology that uses computer technology to assist in design work. Its basic principle is to input the geometric shape, size, material, and other information of the design object into the computer and design, modify, optimize, and output it through specific CAD software. CAD technology mainly includes the following aspects: graphic processing technology, data management technology, calculation and analysis technology, and interactive technology.

WNN is a model based on wavelet transform and NN theory. In WNN, the input signal is transformed by wavelet to obtain wavelet coefficients at different scales, and then these wavelet coefficients are used as the input of NN for learning and optimization. WNN model mainly includes the following technologies:

(1) Wavelet transform technology: The wavelet transform represents a multi-scale analytical technique capable of decomposing signals into wavelet coefficients across various scales, enabling the extraction of time-frequency characteristics inherent in the signal.

(2) NN technology: NN is a computational model to simulate the connection of human brain neurons, which can automatically adjust the connection weight through learning and realize the functions of classification, recognition and prediction of input signals.

(3) Learning algorithm technology: WNN needs to adopt appropriate learning algorithms for training and optimization, such as the backpropagation algorithm, genetic algorithm and particle swarm optimization.

3.2 Basic Principle and Technology of Music Signal Processing

Music signal processing is a method based on digital signal processing technology to analyze, process, and create music signals deeply. This technology first digitizes the music signal and then carries out diversified processing and analysis in the computer environment. The basic process diagram of music acquisition and expression is as follows.

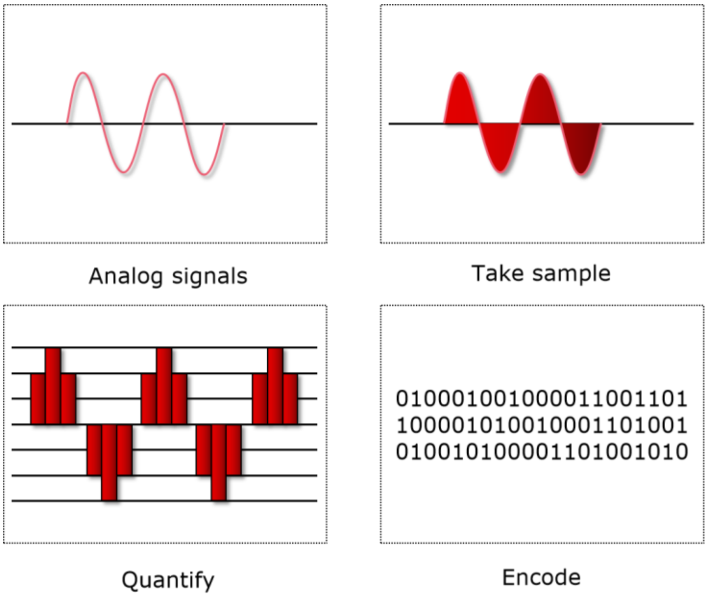

Figure 1: Basic process diagram of music acquisition and representation.

Figure 1 shows the basic flowchart of music acquisition and representation. There are three key technologies involved in music signal processing: (1) Music signal representation and analysis technology: music signals are represented and analyzed by different methods such as time domain, frequency domain and cepstrum, and advanced signal processing technologies such as Fourier transform, short-time Fourier transform and wavelet transform are also needed to extract and analyze the core features of music signals. (2) Audio content editing and synthesis technology: In the process of music signal processing, the audio needs to be carefully edited and synthesized, such as audio clip cutting, connection, speed adjustment, tone change and other operations, and advanced technologies such as audio synthesis and mixing should also be used. (3) Interpretation and transmission technology of music emotion: Music signal processing should not only analyze the emotional elements in music but also effectively convey them, including key technologies such as classification, identification and visualization of music emotion.

3.3 Feasibility Analysis of Music Signal Processing and Optimization Based on CAD and WNN Model

The optimization algorithm is a calculation method to find the optimal solution, which is widely used in various fields. Within the realm of music signal processing and optimization, optimization algorithms play a pivotal role in identifying optimal parameter combinations and processing tactics. Some widely used optimization algorithms encompass genetic algorithms, particle swarm optimization, and ant colony algorithms. These methodologies operate on the fundamental principle of mimicking optimization processes observed in natural or societal phenomena to pinpoint optimal solutions, exhibiting robust global search capabilities.

Music signal processing and optimization based on CAD and WNN models has the following feasibility:

CAD technology can improve the efficiency and quality of music signal processing and reduce design time and cost through automatic design processes and optimization algorithms. At the same time, CAD technology can also provide a visual design and editing environment so that designers can operate and modify more intuitively.

WNN model has strong learning and optimization ability and can deeply analyze and optimize music signals. By training and optimizing the WNN model, we can extract the time-frequency characteristics and emotional characteristics of music signals so as to realize the functions of music classification, recognition, and prediction. At the same time, the WNN model can be combined with other optimization algorithms to improve the efficiency and accuracy of music signal processing further. Therefore, the music signal processing and optimization based on CAD and WNN models have broad application prospects and research value.

4 MODEL CONSTRUCTION

4.1 Music Signal Processing and Optimization Model Construction Based on CAD and WNN Model

The construction of an optimization algorithm is a systematic and orderly process. Firstly, the problem is defined. That is, the optimization goal is clearly set. The optimization objectives of this article include improving the classification accuracy and reducing the loss function value. Then, data preparation is carried out, including collecting and arranging data sets for training and testing the model. These data sets should contain input data and corresponding labels or expected outputs. Then comes the network design, which involves setting the structure and parameters of WNN (network depth, number of nodes, selection of activation function, etc.) and selecting appropriate wavelet basis functions and scales to perform wavelet transform. Continuous wavelet transform is a wavelet transform in which the expansion and contraction of the scale parameter a and the translation of positioning parameters b take continuous values. In practical application, a they b are discretized, so all discrete wavelet bases form a function group, which can be used as basis functions. The wavelet basis function used here is:

$$h_{a,b}(x) = h\left[\frac{x-b}{a}\right] \tag{1}$$

$$h(x) = \cos(1.75x)^{*}\exp(-x^2/2) \tag{2}$$

In this article, the backpropagation algorithm is selected to train and optimize the network, and the appropriate learning rate and iteration times are set according to the actual needs. For the short-term average energy $E(i)$ of the music signal in the i frame, the formula is as follows:

$$E(i) = \sum_{n=1}^{N} X_i^2(n) \tag{3}$$

Where N is the frame length and $X_i\left(n\right)$ is an amplitude energy of the music signal at the n point. The relationship between Mel frequency and music signal frequency f is as follows:

$$Mel\left(f\right)=1125\ln\left(1+f/700\right) \tag{4}$$

By extracting mel-frequency cepstral coefficients, all features in music can be represented as a part of a vector, and each frame of music can be regarded as an independent vector. The spectrogram already contains the spectrum information of the whole piece of music, which is a dynamic spectrum display method. Through the fast Fourier transform, we can get the following spectrum performance:

$$X\left(n,k\right)=\sum_{M=0}^{N-1}X_n\left(m\right)e^{-j\frac{2\pi km}{N}} \tag{5}$$

$$0\leq k\leq N-1 \tag{6}$$

Among them, $X_n\left(m\right)$ is the n frame signal of framing music. $\left|X\left(n,k\right)\right|$ is the short-time amplitude spectrum estimation of $X\left(n\right)$, and the spectrum energy density function $p\left(n,k\right)$ at m is:

$$p\left(n,k\right)=\left|X\left(n,k\right)\right|^2 \tag{7}$$

Finally, the performance evaluation is carried out; that is, the trained network model is evaluated using the test data set. This will be explained in detail in subsequent chapters.

When building a music signal processing model based on CAD, it is necessary to determine the processing objectives and tasks first and then select appropriate CAD software and tools to design and build the model according to the requirements of the processing tasks. Specifically, this article first carries out music signal input: the music signal to be processed is input into the CAD system, which can be audio files, MIDI files and other formats. Then, preprocessing the music signal: preprocessing the music signal, such as noise reduction, equalization, filtering and other operations, in order to improve the quality and processability of the signal. The CAD design: using CAD software to design the music signal processing model, including signal processing flow, algorithm selection, parameter setting and so on. CAD system provides abundant design tools and resources, which can be used for model design and modification conveniently. Finally, the music signal is processed: according to the designed model, the music signal is processed and analyzed, such as audio editing, sound effect processing, feature detection and so on. The CAD system can display the processing results and effects in real-time, which is convenient for users to adjust and optimize. The spectrum of musical signal idioms is shown in Figure 2.

In order to build a music signal processing and optimization model based on CAD and WNN, this section combines CAD technology with the WNN model to form a complete processing and optimization process. Specifically, the construction process of the model includes the following steps:

(1) Music signal input and preprocessing: input the music signal to be processed into the system and perform preprocessing operations such as noise reduction and equalization. You can use CAD technology to edit and synthesize audio.

(2) Feature detection and representation: WNN is used for feature detection, and representation learning of the preprocessed music signal to extract the time-frequency features and emotional features of the music signal, and CAD technology is used for visual display and analysis. In the music signal processing and optimization based on CAD and WNN, its transfer function $H\left(z\right)$ can be expressed as:

$$H\left(z\right)=U\left(z\right)V\left(z\right)R\left(z\right) \tag{8}$$

Where $U\left(z\right)$ represents an excitation signal $V\left(z\right)$ is the transfer function of the sound channel and $R\left(z\right)$ is a first-order high-pass filter.

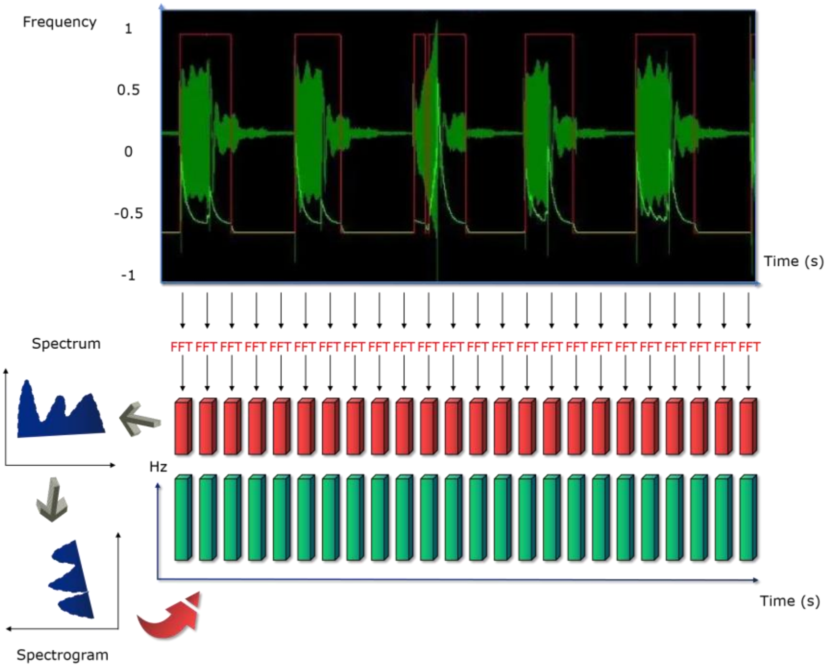

Figure 2: Spectrum of idioms generated by music signals.

The transfer function $V(z)$ of the sound channel is an all-pole model;

$$V(z)=\frac{1}{\sum_{i=0}^{p} a_i z^{-i}} \tag{9}$$

In the formula p is the order, with a value in the range of 8~12, and each pair of poles corresponds to a formant. a_i is the parameter of sound channel moulding.

(3) Optimization algorithm selection and training: Select the appropriate WNN structure and learning algorithm for network training and optimization so as to find the optimal parameter combination and processing strategy. Set appropriate learning rate and iteration times for training and use CAD technology to visualize and adjust the model.

(4) Signal processing and optimized output: according to the trained WNN model and optimization algorithm, the music signal is processed and optimized to obtain the processed music signal or optimized result. You can use CAD technology to output and save audio.

4.2 Model Parameter Setting and Implementation Process

In the process of model construction, it is necessary to set appropriate parameters and implementation processes to ensure the performance and efficiency of the model. Firstly, this article determines the type and format of music signal to be processed, such as audio file, MIDI file, etc. Then, according to the requirements of processing tasks, select appropriate CAD software and tools for model design and construction. Then, determine the structure and parameters of WNN, including the number of network layers, nodes, activation function, etc., for network training and optimization. Finally, select the appropriate performance evaluation index to evaluate and adjust the trained model. Table 1 shows the performance evaluation results under different WNN structures.

Hidden layer number	Number of hidden layer nodes	Accuracy rate	Recall rate	F1 value
1	16	0.85	0.80	0.82
1	32	0.88	0.85	0.86
1	64	0.89	0.86	0.87
2	16-16	0.90	0.87	0.88
2	32-32	0.91	0.94	0.92
2	64-64	0.92	0.90	0.91

Table 1: Performance evaluation results under different WNN structures.

In the above experiments, the structure with two hidden layers and 32-32 nodes in each layer performs well in the performance index. Therefore, this article adopts an NN structure with two hidden layers, and the number of nodes in each layer is 32-32.

Table 2 shows the performance evaluation results of WNN under different parameter settings.

Learning rate	Iterations	Batch size	Accuracy rate	Recall rate	F1 value
0.01	50	32	0.85	0.80	0.82
0.01	100	32	0.93	0.89	0.91
0.01	200	32	0.90	0.87	0.88
0.01	100	64	0.89	0.86	0.87
0.01	100	128	0.87	0.83	0.85
0.001	100	32	0.84	0.80	0.82
0.0001	100	32	0.79	0.75	0.77

Table 2: Performance evaluation results of WNN under different parameter settings.

In the experiment, it can be found that when the learning rate is 0.01, the number of iterations is 100, and the batch size is 32, the performance of the model is better. Therefore, this article adopts this data. Figure 3 shows the time domain waveform of the music signal collected by the model.

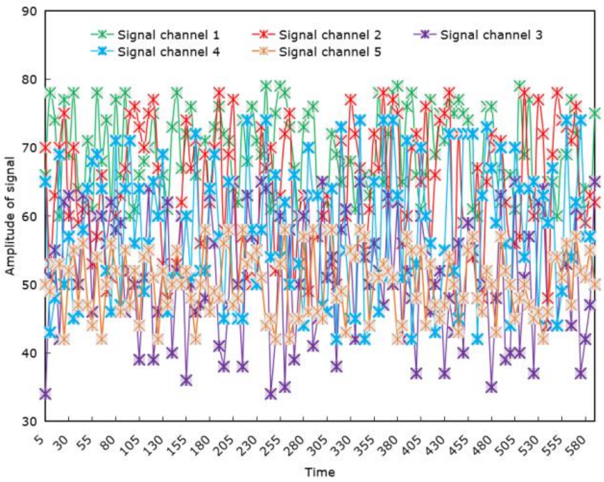

Figure 3: Time domain waveform diagram of music signal.

From the time domain waveform, we can see that the music signal has a certain periodicity and sense of rhythm, and the waveform presents different forms in different time periods, reflecting the rhythm and melody changes of music. This accords with the general characteristics of music signals.

5 SIMULATION EXPERIMENT

In order to verify the validity of the music signal processing and optimization model based on CAD and WNN, this section systematically carries out simulation experiments. The experiment uses Python as the main programming language and uses Keras deep learning framework to construct and train the WNN model. At the same time, CAD software such as AutoCAD is used to design and construct the music signal processing model. In order to train and test the model in this article, the GTZAN music data set and the Million Song Dataset are selected. The data set is shown in Table 3 below.

Dataset name	*Number of pieces of music*	*Music types and styles*	*The length of each piece of music*	*Sampling rate*
GTZAN music Dataset	10,000 paragraphs	Classical, pop, rock, jazz, electronic and many other types and styles.	30 seconds	44.1kHz
Million Song Dataset	-	A variety of styles and genres	50 seconds	-

Table 3: Performance evaluation results of WNN under different parameter settings.

Different control groups were set up, including only using CAD technology for music signal processing, only using the WNN model for optimization, and combining CAD and WNN models for music signal processing and optimization. On the GTZAN music data set, the classification accuracy of different methods is shown in Figure 4. On the Million Song Dataset, the classification accuracy of different methods is shown in Figure 5.

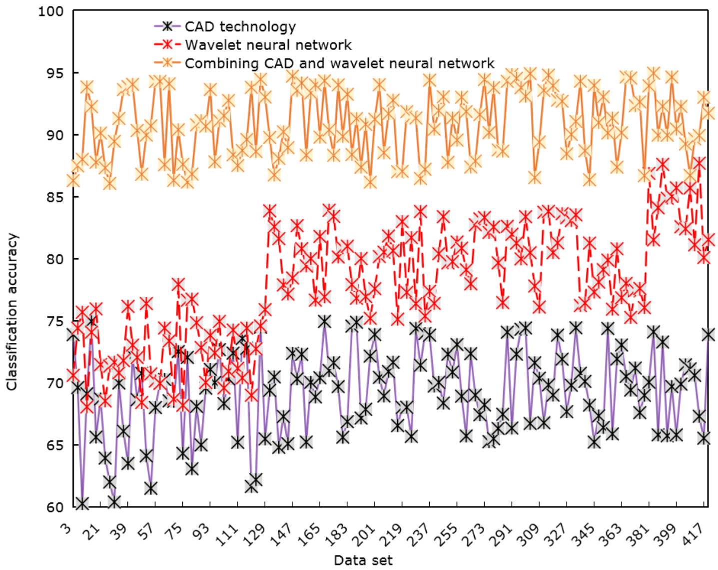

Figure 4: Classification accuracy -GTZAN dataset.

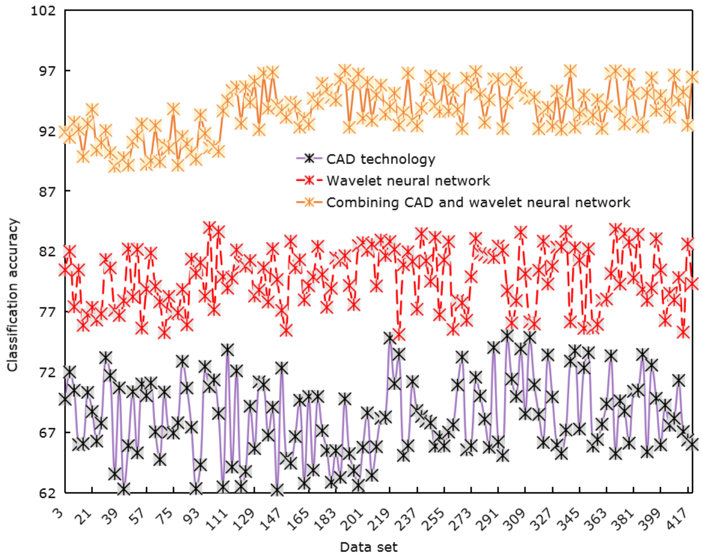

Figure 5: Classification accuracy -Million Song Dataset dataset.

After training and testing, the experimental results of different control groups were obtained. The accuracy of music signal processing using only CAD technology is about 70%, and the processing efficiency is low. The accuracy of optimization only using the WNN model is about 81%, and the processing efficiency is improved. The accuracy of music signal processing and optimization combined with the CAD and WNN models is about 94%, and the processing efficiency is significantly improved.

SVM and RNN are used as the comparison algorithms in this article to evaluate the performance of different methods in music classification tasks, and their processing efficiency is compared, as shown in Figures 6 and 7.

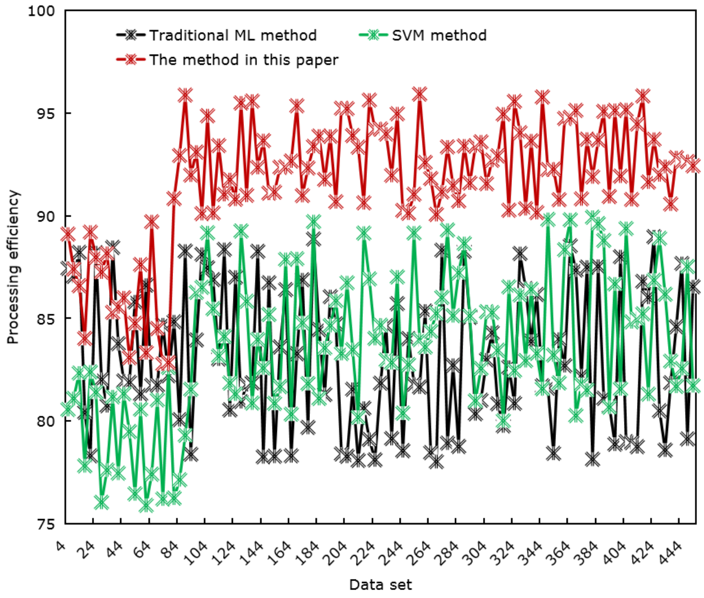

Figure 6: Processing efficiency -GTZAN dataset.

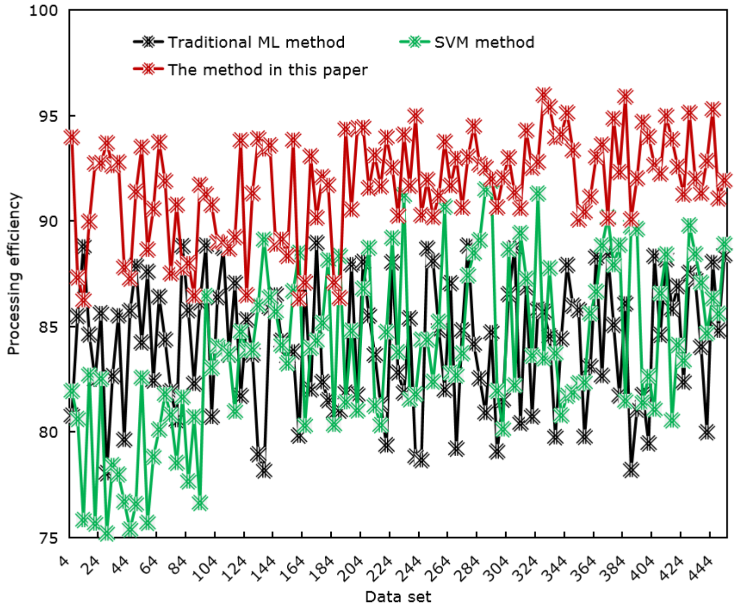

Figure 7: Processing efficiency -Million Song Dataset dataset.

From the results, it can be seen that this model also has certain advantages in processing efficiency. Although the training time of the WNN model is relatively long, once the training is completed, its processing speed in the classification process is faster. In contrast, SVM and RNN algorithms need more computing resources and time when dealing with music classification tasks. SVM and RNN algorithms have some limitations in music classification tasks, which can't make full use of the information in audio signals, resulting in low accuracy and processing efficiency. Combining CAD and WNN models for music signal processing and optimization can significantly improve processing efficiency and accuracy. This is mainly because CAD technology can provide an intuitive visual design and editing environment so that designers can operate and modify more intuitively. The WNN model has strong learning and optimization ability, which can deeply analyze and optimize music signals and extract richer time-frequency features and emotional features.

The experimental results in this section show that the music signal processing and optimization model based on CAD and WNN has obvious advantages in music signal processing tasks. This model can combine the visualization and editing ability of CAD technology and the deep learning and optimization ability of the WNN model to realize more efficient and accurate music signal processing and optimization. At the same time, the model has good generalization performance and can adapt to different types and styles of music signal processing tasks. The reasons for the experimental results can be summarized as follows: firstly, CAD technology can provide an intuitive visual design and editing environment so that designers can operate and modify more intuitively; Secondly, WNN model has strong learning and optimization ability, which can deeply analyze and optimize music signals and extract more abundant time-frequency features and emotional features; Ultimately, the integration of these two elements allows for the maximization of their individual strengths, leading to more efficient and precise processing and optimization of music signals.

6 CONCLUSIONS

In this article, CAD and WNN models are deeply studied and successfully applied to music signal processing and optimization. By combining their advantages, an effective music signal processing and optimization model is constructed, and its feasibility and effectiveness are verified by simulation experiments. The results show that the model can achieve high-precision feature detection and classification in music signal processing, as well as remarkable results in music optimization.

The research contribution of this article is mainly reflected in the following aspects: CAD technology is introduced into the field of music signal processing, which provides a visual and structured processing method for complex music signals. Using the advantages of WNN in music signal processing, high-precision feature detection and classification of music signals are realized. A music signal processing and optimization model combining CAD and WNN is constructed, which provides a useful reference for research in related fields. The feasibility of the model is verified by simulation experiments, which provides strong support for practical application.

Although this study has achieved some beneficial results, there are also some limitations and shortcomings. In the future, we can further expand the application research of CAD and WNN models in other fields and explore their application potential in more scenarios. At the same time, we can strengthen cooperation with practical music projects and apply the model proposed in this study to practical projects for verification and evaluation so as to promote its popularization and application in practical applications.

Jie Tian, https://orcid.org/0009-0004-6375-0332
Bing Tian, https://orcid.org/0009-0009-3508-1772

REFERENCES

[1] Bishop, L.; Cancino, C.-C.; Goebl, W.: Moving to communicate, moving to interact: Patterns of body motion in musical duo performance, Music Perception: An Interdisciplinary Journal, 37(1), 2019, 1-25. https://doi.org/10.1525/mp.2019.37.1.1

[2] Calilhanna, A.: Ogene Bunch music analyzed through the visualization and sonification of beat-class theory with ski-hill and cyclic graphs, The Journal of the Acoustical Society of America, 148(4), 2020, 2697-2697. https://doi.org/10.1121/1.5147469

[3] Dotov, D.; Bosnyak, D.; Trainor, L.-J.: Collective music listening: movement energy is enhanced by groove and visual social cues, Quarterly Journal of Experimental Psychology, 74(6), 2021, 1037-1053. https://doi.org/10.1177/1747021821991793

[4] Farhadi, M.; Abbaspour, G.-Y.; Mahmoudi, A.; Mari, M.-J.: An integrated system of artificial intelligence and signal processing techniques for the sorting and grading of nuts, Applied Sciences, 10(9), 2020, 3315. https://doi.org/10.3390/app10093315

[5] Fu, Y.; Zhang, M.; Nawaz, M.; Ali, M.; Singh, A.: Information technology-based revolution in music education using AHP and TOPSIS, Soft Computing, 26(20), 2022, 10957-10970. https://doi.org/10.1007/s00500-022-07247-w

[6] Ge, M.; Tian, Y.; Ge, Y.: Optimization of the computer-aided design system for music automatic classification based on feature analysis, Computer-Aided Design and Applications, 19(S3), 2021, 153-163. https://doi.org/10.14733/cadaps.2022.S3.153-163

[7] Georges, P.; Seckin, A.: Music information visualization and classical composers discovery: an application of network graphs, multidimensional scaling, and support vector machines, Scientometrics, 127(5), 2022, 2277-2311. https://doi.org/10.1007/s11192-022-04331-8

[8] Gorbunova, I.-B.; Plotnikov, K.-Y.: Music computer technologies in education as a tool for implementing the polymodality of musical perception, Musical Art and Education, 8(1), 2020, 25-40. https://doi.org/10.31862/2309-1428-2020-8-1-25-40

[9] Im, H.; Song, H.; Jung, J.: The effect of streaming services on the concentration of digital music consumption, Information Technology & People, 33(1), 2020, 160-179. https://doi.org/10.1108/ITP-12-2017-0420

[10] Jakubec, M.; Chmulik, M.: Automatic music genre recognition for in-car infotainment, Transportation Research Procedia, 40(11), 2019, 1364-1371. https://doi.org/10.1016/j.trpro.2019.07.189

[11] Lattner, S.; Nistal, J.: Stochastic restoration of heavily compressed musical audio using generative adversarial networks, Electronics, 10(11), 2021, 1349. https://doi.org/10.3390/electronics10111349
[12] Liu, S.; Chen, P.-Y.; Kailkhura, B.; Zhang, G.; Hero, III.-A.-O.; Varshney, P.-K.: A primer on zeroth-order optimization in signal processing and machine learning: Principals, recent advances, and applications, IEEE Signal Processing Magazine, 37(5), 2020, 43-54. https://doi.org/10.1109/MSP.2020.3003837
[13] Maba, A.: Computer-aided music education and musical creativity, Journal of Human Sciences, 17(3), 2020, 822-830. https://doi.org/10.14687/jhs.v17i3.5908
[14] Monga, V.; Li, Y.; Eldar, Y.-C.: Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing, IEEE Signal Processing Magazine, 38(2), 2021, 18-44. https://doi.org/10.1109/MSP.2020.3016905
[15] Nam, J.; Choi, K.; Lee, J.; Chou, S.-Y.; Yang, Y.-H.: Deep learning for audio-based music classification and tagging: Teaching computers to distinguish rock from Bach, IEEE signal processing magazine, 36(1), 2018, 41-51. https://doi.org/10.1109/MSP.2018.2874383
[16] Solanki, A.; Pandey, S.: Music instrument recognition using deep convolutional neural networks, International Journal of Information Technology, 14(3), 2022, 1659-1668. https://doi.org/10.1007/s41870-019-00285-y
[17] Wasimuddin, M.; Elleithy, K.; Abuzneid, A.; Faezipour, M.; Abuzaghleh, O.: Multiclass ECG signal analysis using global average-based 2-D convolutional neural network modeling, Electronics, 10(2), 2021, 170. https://doi.org/10.3390/electronics10020170