A Piano Single Tone Recognition and Classification Method Based on CNN Model

Miaoping Geng, Ruidi He*, Ziyihe Zhou
The Conservatory of Music, Hebei Institute of Communications, Shijiazhuang, Hebei 050010, China

***Abstract*—In order to improve the recognition and classification effect of piano single tone, this paper combines the CNN (Convolutional Neural Networks) model to construct the piano single tone recognition and classification model, and equalizes the uniformly irradiated parabolic tone transmission hardware. In this paper, the analytic method is used to calculate the direction diagram of the tone transmission hardware, and the analytical expression for calculating the gain of the tone transmission hardware is obtained. Moreover, this paper gives the calculation and analytical expression of the hardware gain of the tone transmission in the main lobe, and obtains the calculation method of the relative position of the two tone transmission hardware by using the conversion relationship between the global coordinate and the local coordinate. Finally, the variation law of the received power with the azimuth/elevation angle of the receiving tone transmission hardware and the incident high-power microwave frequency is given. The experimental study shows that the piano single tone recognition and classification method based on CNN model proposed in this paper can play an important role in piano single tone recognition. This article improves the note recognition algorithm for piano music by combining note features with frequency spectrum to obtain note spectrum, which improves the accuracy of audio classification recognition.**

Keywords—CNN model; piano; single tone recognition; classification

I. Introduction

For the interpretation of a beautiful piano piece, if the performer is the driver of the soul of the piece, then a piano with a beautiful tone is the carrier of the soul. Therefore, it is particularly important to understand and study the quality of the piano. The quality of a piano can be judged from six aspects: tone, touch, tuning stability, durability, appearance and tension. However, for players, tone and touch are undoubtedly the most important. It is not an easy task to make a correct judgment on the sound quality of a piano. First of all, it needs to ensure the piano intonation. In the case of intonation, the pronunciation of the bass region should be strong and powerful, the sound should be extended enough, and the pronunciation should not be short or weak. The mid-range requires that the extension of the sound should be as long as possible, the timbre should be beautiful and soft, and the pronunciation should not be dull or blunt. The pronunciation of the high-pitched area should be bright and clear, not too gorgeous or impure, and no reverberation is required. When the timbres of the three sound zones are satisfactory, it should be noted that the transition of the three sound zones should be natural, the timbre should be unified, and the volume ratio should be coordinated. The so-called tactile sensation refers to the responsive ability of the keyboard and the action mechanism to transmit the player's playing force to the strings. The touch of the piano should be comfortable for beginners and comfortable for accomplished players.

Music information signals themselves belong to fuzzy signals and require strict mathematical models to describe them. Traditional information processing methods are difficult to solve the fuzzy situation of music signals. Therefore, some studies use intelligent information analysis and processing methods such as fuzzy systems, neural networks, expert systems, and genetic algorithms to process music information [1].

Fuzzy system has the advantages of not needing precise mathematical model, easy to use human experience knowledge, nonlinear, robust and so on. Since there are many ambiguities in music information, language and thinking, the use of fuzzy sets to describe the characteristics of music, and the use of fuzzy logic and fuzzy reasoning for feature analysis and identification, should be said to be the closest to the cognitive process of people's music. At present, many studies have proved that the fuzzy system is an effective method for the study of music information. In the automatic chord analysis system of literature [2], the process of listening, feeling and understanding music all use the membership function of the fuzzy set, and the basic membership function The function is designed according to music theory and can be modified by a few simple parameters, which facilitates the study of music information from both music theory and human perception.

In order to improve the effect of piano single tone recognition and classification, this paper combines the CNN model to build a piano single tone recognition and classification model to improve the efficiency of intelligent processing of piano tones.

The contribution of this article is as follows: Section I is the introduction. Related work is given in Section II. Calculation and Analysis of front door of coupling quantity of piano tone transmission is given in Section III. Section IV delves in to the empirical model, results, analysis and discussion. Finally, Section V concludes the paper. This article improves the note recognition algorithm for piano music by combining note features with frequency spectrum to obtain note spectrum, which improves the accuracy of audio classification recognition.

*Corresponding Author.

II. Related Work

Neural network has the advantages of distributed information storage, parallel processing, self-organization, and self-learning, the existing research results use two types of neural networks [3]: (1) Multilayer perceptron (BP network): Using a layered neural network can make the system not only start from the notes, but also from the overall structure. Highly comprehensive processing of music information. The hierarchical recurrent neural network of [4], the first layer is for each measure, and the second layer is for each note. For a rhythm-regular piece like a waltz, both melody and rhythm characteristics can be taken into account. In addition, the use of Makov chains can also make the system predictive. (2) Hopfield network (feedback network): Research in [5] uses the associative memory function of the Hopfield network to recombine the existing melody data to achieve the purpose of composition. The expert system is suitable for embedding a large amount of music knowledge into the computer system, and at the same time enables the system to flexibly use this knowledge to make judgments, and has the ability to acquire and increase knowledge, and is best at simulating human experts to solve complex problems. Study [6] used 26 rules to complete the automatic playing expert system according to the performance characteristics of performers. Study [7] realized an expert system that automatically discriminates Bach's musical style. In the processing of musical information, intelligent methods play an important role. The choice of these methods should mainly be based on specific needs. For example, the analysis of the characteristics of music harmony, timbre, etc. is often inseparable from a fuzzy system that has a strong ability to describe music information. Automatic playing systems usually choose expert systems that are easy to acquire knowledge. Some auxiliary composing systems use expert systems, while others use expert systems. The neural network is used to simulate human image thinking, so as to compose music according to certain requirements [8].

Study [9] argues: "The typical sound quality of an instrument should be attributed to the relative intensities of all harmonics in a definite or relatively definite region in the musical scale." Although this view is not entirely correct, it points out that the sound quality of an instrument is the importance of harmonic amplitudes to the sound quality of musical instruments. In terms of using computer to synthesize piano sound and improving the sound quality of piano sound through computer processing simulation, literature in [10] pointed out that harmonics are an important factor of sound quality, but this paper discusses how to improve the sound quality from the perspective of harmonic amplitude and phase changing with time, piano sound. Study in [11] believes that the harmonic amplitude is an important factor that constitutes the sound quality of the piano, and then uses the method of simulating multiple strings to reasonably adjust the frequency spectrum of the piano to study how to improve the sound of the piano.

When the hammers hit the strings, the piano sound reaches a peak of vibration soon after a brief onset. From a musical point of view, it is better to have a shorter time in this stage. If the time is too long, the sound will appear soft and lack the feeling of rigidity; but it should not be too short, otherwise the sound will have a stiff feeling. When the peak of the sound amplitude is reached, there will be a rapid decay. This period can be divided into two stages: early decay 2 and late decay 3. Of course, the decay speed of this stage will vary greatly with different keys and different percussion strengths [12]. An excellent pianist can well control the percussion intensity and time during the performance, so that the piano can make a full sound and get a good timbre effect. The strings are restrained by the sound felt and the vibration is rapidly attenuated. For the treble keys of a piano, these processes are less complicated, and the time domain graph of the treble keys looks like a straight line with only a sloping downward trend. The time domain characteristics of all these piano sounds are closely related to the piano hardware itself, such as felt, hammer shape, soundboard, etc. That is to say, the hardware of the piano itself is a very important part of the time domain characteristics of the piano sound. determinants [13].

Regarding the influence of the harmonic amplitude of the piano on the sound quality, the literature [14] first recorded the piano, and then took one of the signals, and performed simple processing on the harmonic amplitude, for example, the harmonic amplitude was formed proportionally, or the harmonic amplitude was decreased according to the law form a set of synth sounds, etc. Then compare these synthetic sounds with the original recording of the piano, and finally find that the sound quality after such simple processing is not as good as the original sound, at most only close to the original recording. , processing the harmonic amplitude of the piano is not a simple process, and it also shows that the composition of the harmonic amplitude does affect the sound quality of the piano sound to a great extent. In addition to studying the influence of harmonic amplitude on sound quality, study in [15] found the law of piano harmonics changing with time, that is, the law that different harmonics have different time delays. For example, the first harmonic and the second harmonic appear at different times, and there is a small delay, which means that the second harmonic will come later than the first harmonic. And as the harmonic order increases, their delay time increases.

Using a computer to analyze piano music signals, the processed music information must first be digitally processed, that is, convert the analog music signals collected by the recording equipment into digital music signals. In this process, two aspects are mainly considered: conversion accuracy and operation efficiency. The A/D conversion accuracy is realized by the number of bits of the A/D conversion device; the operation efficiency is related to the digitization accuracy and sampling rate of the music. However, if the data accuracy and sampling rate are too low, it will cause relatively large waveform distortion in the digitization process of the signal [16].

The collected piano music will be studied from the perspectives of physical acoustics, rhythm, fast Fourier transform, wavelet analysis, etc., to study the specific extraction methods and analysis methods of different musical features, and complete the design of pitch, duration, intensity and other feature extraction methods. and its computer implementation. The extraction of musical features provides a practical basis for the design of subsequent piano evaluation systems [17].

The keystroke sensitivity of the piano keys, that is, the response frequency of the keyboard, is one of the main parameters affecting the touch feeling of the piano keys, and it is also a problem that piano manufacturers are concerned about at present. By analyzing whether the response times of the piano keys can reach the standard, the quality of the percussion performance of the piano keys can be judged. To test this parameter, we must first ensure that the strength of hitting the piano keyboard is the same, the interval is even, and the frequency is adjustable, so that accurate test results can be obtained. In terms of testing algorithms, the keystroke sensitivity of piano keys can be effectively detected through waveform normalization, endpoint detection, and single-note separation [18].

III. Calculation and Analysis of Front Door Coupling Quantity of Piano Tone Transmission Hardware Port

On the basis of studying the characteristics of high-power microwaves, in order to quantitatively obtain the coupling number of high-power microwaves through the front door to the port of the piano tone transmission hardware, it is necessary to study the calculation method of the coupling amount of the front door.

A. Description of Calculation Method of Front Door Coupling

The HPM (High Performance Computing) generated in the high-power microwave transmitter transmits the hardware radiation through piano tone. As shown in Fig. 1, piano tone transmission hardware 1 is used as receiving, and piano tone transmission hardware 2 is used as piano tone transmission hardware for transmitting HPM.

When the distance between the received piano tone transmission hardware and the HPMpiano tone transmission hardware is R, the power density S from the HPM to the receiving piano tone transmission hardware is:

$$S = \frac{P_t G_t\left(\theta_2, \varphi_2\right)}{4\pi R^2}\left[\mathrm{W/m^2}\right] \tag{1}$$

The content calculated by Formula (1) is the power density of the far-field field, and the condition for satisfying the far-field field is:

$$R \geq 2d^2/\lambda \tag{2}$$

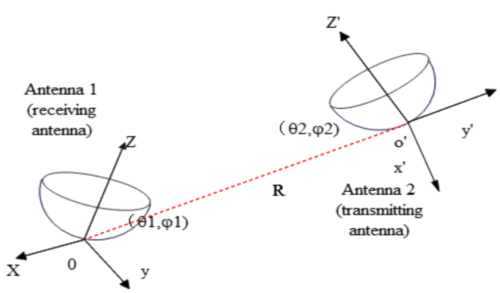


Fig. 1. The relative position of the two piano tone transmission hardware.

According to the Frith transmission formula, the power P_r received as the received piano tone transmission hardware is expressed as:

$$P_r = A_e\left(\theta_1, \varphi_1\right) \cdot S \tag{3}$$

The power received by the piano tone transmission hardware can be obtained from Formulas (1) and (3). In Formulas (1) and (3), the specific gain and effective receiving area of the two piano tone transmissions hardware need to be obtained.

B. Achieve Method of the Gain of the Piano Tone Transmission Hardware

For any form of piano tone transmission hardware, some commercial electromagnetic simulation software (HFSS (High Frequency Structure Simulator), FEKO etc.) can be used to establish the model of piano tone transmission hardware. The direction diagram and gain of piano tone transmission hardware are obtained through simulation.

The parabolic piano tone transmission hardware belongs to a class of piano tone transmission hardware with symmetrical structure, and its direction map and gain can be obtained by analytical method.

As shown in Fig. 2, for the larger rotating paraboloid piano tone transmission hardware with uniform illumination aperture shown in Fig. 2(a), the radiated electromagnetic waves. In essence, it can be equivalent to the electromagnetic wave radiated by the circular aperture (same diameter as the paraboloid) on a metal plate with an infinite size irradiated by a uniform plane wave as shown in Fig. 2(b).

For the equivalent model, the method for obtaining the far-field pattern can be based on the Huygens principle. According to this method, the normalized field strength pattern $E(\theta)$ obtained is shown in Formula (4).

$$E(\theta) = \frac{J_1\left[\pi d_\lambda \sin\theta\right]}{\sin\theta} \cdot \frac{2}{\pi d_\lambda} \tag{4}$$

In the formula, J_1 represents the first-order Bessel function, and θ refers to the angle relative to the focal axis.

d_λ is the wavelength number of the diameter d of the circular mouth, which can be expressed as

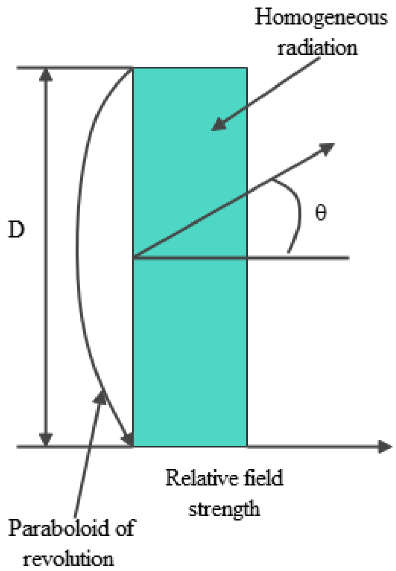
$$d_\lambda = d/\lambda \tag{5}$$

The n-order Bessel function can be expressed as:

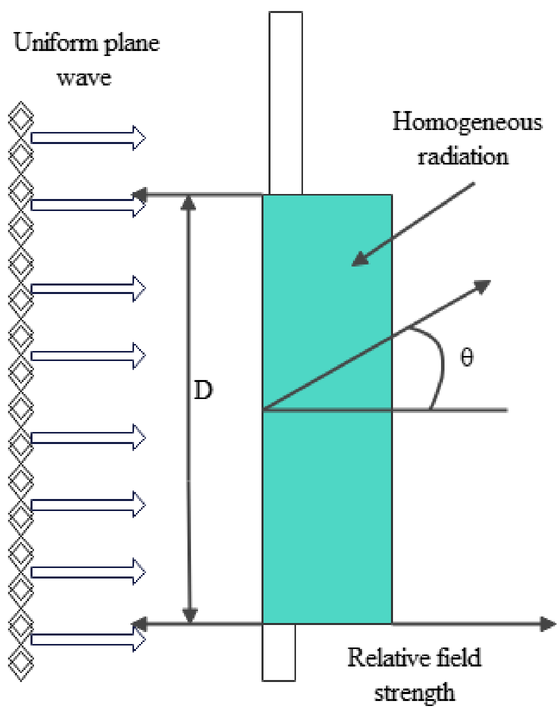
$$J_n(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(n\theta - z\sin\theta)d\theta \tag{6}$$

The expression of the first-order Bessel function is shown in Formula (7).

$$J_1(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\cos(\theta - z\sin\theta)d\theta \tag{7}$$



(a) Dish antenna.



(b) Equivalent model.

Fig. 2. The paraboloid model of revolution and its equivalent model.

Fig. 3 is the normalized field strength pattern when the center frequency of the parabolic piano tone transmission hardware is $f = 10\text{GHz}$ and the aperture is $d = 1\text{m}$.

Formula (8) represents the formula for obtaining the directivity coefficient of the parabolic piano tone transmission hardware:

$$D(\theta,\varphi) = (\tilde{E}(\theta,\varphi))^2 \cdot \frac{4\pi}{\theta_{HP1}\theta_{HP2}} \quad (8)$$

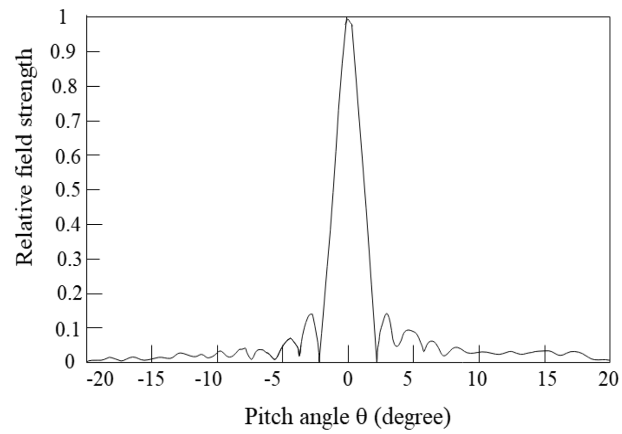


Fig. 3. Normalized direction diagram.

In the formula, $\tilde{E}(\theta,\varphi)$ is the normalized electric field intensity in the direction away from the piano tone transmission hardware, and θ_{HP1} and θ_{HP2} are the half-power beam widths of the main lobes of the piano tone transmission hardware in the two main planes.

The gain of the piano tone transmission hardware in a certain direction θ and φ is defined as:

$$G(\theta,\varphi) = \eta D(\theta,\varphi) \quad (9)$$

In the formula, η is the efficiency factor of piano tone transmission hardware.

When one piano tone transmission hardware (transmit/receive) is located in the main lobe of another piano tone transmission hardware (receive/transmit), the gain of piano tone transmission hardware can be calculated by the following method.

When the receiving piano tone transmission hardware is in the main lobe of the HPMpiano tone transmission hardware, the gain of the transmitting piano tone transmission hardware can be expressed as:

$$G_t = \varepsilon_a \frac{4\pi}{\lambda^2} A_{pt} \quad (10)$$

In the formula, ε_a is the aperture efficiency of the piano tone transmission hardware, and A_{pt} is the actual aperture area of the piano tone transmission hardware.

Similarly, when the HPM transmitting piano tone transmission hardware is located in the main lobe of the receiving piano tone transmission hardware, the gain of the receiving piano tone transmission hardware can be expressed as:

$$G_r = \varepsilon_a \frac{4\pi}{\lambda^2} A_{pr} \quad (11)$$

In the formula, A_{pr} is the actual aperture area of the receiving piano tone transmission hardware.

C. Calculation Method of Effective Receiving Area of Piano Tone Transmission Hardware

The effective receiving area $A_e(\theta,\varphi)$ of the piano tone transmission hardware is related to the gain of the receiving piano tone transmission hardware and the polarization mismatch coefficient, which can be expressed as:

$$A_e(\theta,\varphi) = \rho \cdot G_r(\theta,\varphi) \cdot A_d \tag{12}$$

In the formula, ρ is the polarization mismatch coefficient, the value is between $0 \sim 1$, and A_d is the integral effective area of piano tone transmission hardware. When calculating the integral effective area of the piano tone transmission hardware, it is necessary to consider the relationship between the frequency f_t of the HPM transmitting piano tone transmission hardware and the center frequency f_r of the receiving piano tone transmission hardware. When $f_t < f_r$, it is called down-band coupling. When $f_t \approx f_r$, it is called in-band coupling. When $f_t > f_r$, it is called on-band coupling. The up-band coupling and the down-band coupling are collectively referred to as out-of-band coupling.

When calculating the in-band coupling, λ is not only the working wavelength, but also the wavelength corresponding to the center frequency of the receiving piano tone transmission hardware. When discussing the case of out-of-band coupling, the wavelength λ_r corresponding to the center frequency of the receiving piano tone transmission hardware should be used.

In the calculation model of the front door coupling amount, it is necessary to calculate the gain $G_t(\theta_2,\varphi_2)$、$G_r(\theta_1,\varphi_1)$ of the transmitting piano tone transmission hardware and the receiving piano tone transmission hardware. Because when calculating the direction map and gain of the piano tone transmission hardware, the angle is the relative coordinate system of the piano tone transmission hardware itself. Therefore, when calculating (θ_1,φ_1), the coordinates of the transmitting piano tone transmission hardware should be converted into the relative coordinate system of the receiving piano tone transmission hardware. When calculating (θ_2,φ_2), the coordinates of the receiving piano tone transmission hardware should be converted into the relative coordinate system of the transmitting piano tone transmission hardware.

For any point Q on the piano tone transmission hardware, its coordinate in the local coordinate system of the piano tone transmission hardware is (x',y',z'), and its coordinate in the overall coordinate system is (x, y, z). Then, the overall coordinates (x, y, z) of the point Q can be regarded as obtained by the local coordinate (x',y',z') through two rotation transformations and one translation transformation.

1) When the x angle is rotated counterclockwise around the θ -axis, the corresponding coordinate transformation matrix is:

$$T_1 = \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{13}$$

2) When the angle φ is rotated counterclockwise around the z-axis, the corresponding coordinate transformation matrix is:

$$T_2 = \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{14}$$

3) When (x_0,y_0,z_0) is translated along the x-axis, y-axis, and z-axis, the corresponding coordinate transformation matrix is:

$$T_3 = \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{15}$$

4) In the global coordinates, the corresponding relationship between the point Q (x, y, z) and the point in the local coordinates $Q'(x',y',z')$ is:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = T \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \tag{16}$$

The transformation matrix T is:

$$\begin{aligned} T &= T_3 \cdot T_2 \cdot T_1 \\ &= \begin{bmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & y_0 \\ 0 & 0 & 1 & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi & -\sin\varphi & 0 & 0 \\ \sin\varphi & \cos\varphi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & 0 & \sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\theta & 0 & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} \cos\theta\cos\varphi & -\sin\varphi & \sin\theta\cos\varphi & x_0 \\ \cos\theta\sin\varphi & \cos\varphi & \sin\theta\sin\varphi & y_0 \\ -\sin\theta & 0 & \cos\theta & z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{aligned} \tag{17}$$

In the actual simulation calculation, the overall coordinates of a certain point on the piano tone transmission hardware are usually known. Therefore, it is necessary to convert the point coordinates in the overall coordinates into the local coordinate system, that is,

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = T^{-1} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{18}$$

Among them, the transformation coordinate T^{-1} is:

$$T^{-1}=\left(T_3T_2T_1\right)^{-1}=T_1^{-1}T_2^{-1}T_3^{-1}=\begin{bmatrix}\cos\theta & 0 & -\sin\theta & 0\\ 0 & 1 & 0 & 0\\ \sin\theta & 0 & \cos\theta & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}\cos\varphi & \sin\varphi & 0 & 0\\ -\sin\varphi & \cos\varphi & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & -x_0\\ 0 & 1 & 0 & -y_0\\ 0 & 0 & 1 & -z_0\\ 0 & 0 & 0 & 1\end{bmatrix}$$

$$=\begin{bmatrix}\cos\theta\cos\varphi & \cos\theta\sin\varphi & -\sin\theta & -x_0\cos\theta\cos\varphi-y_0\cos\theta\sin\varphi+z_0\sin\theta\\ -\sin\varphi & \cos\varphi & 0 & x_0\sin\varphi-y_0\cos\varphi\\ \sin\theta\cos\varphi & \sin\theta\sin\varphi & \cos\theta & -x_0\sin\theta\cos\varphi-y_0\sin\theta\sin\varphi-z_0\cos\theta\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{19}$$

D. Conversion Relationship between Received Power and Field Strength

Since the calculation model of the coupling quantity of the front door obtains the received power, the required parameter is the field strength value when designing the waveguide plasma limiter. Therefore, it is necessary to obtain the conversion relationship between the power and the field strength.

The conversion relationship between field strength and power is shown in Formula (20).

$$E_{[\text{dB}(\mu V/m)]}=P_{[\text{dBm}]}+AF+107 \tag{20}$$

In the formula, AF is the piano tone transmission hardware factor.

In Formula (20), the received power can be obtained from the calculation model of the front door coupling, and the unknown parameter is the piano tone transmission hardware factor.

The piano tone transmission hardware factor is defined as the ratio between the electric field E and the piano tone transmission hardware terminal voltage V_L, namely:

$$AF=\frac{E}{V_L} \tag{21}$$

It is converted to decibels and expressed as,

$$E_{[\text{dB}(\mu V/m)]}=V_{[\text{dB}(\mu V)]}+AF_{\left[\text{dB}\left(m^{-1}\right)\right]} \tag{22}$$

The effective receiving area A_e of the piano tone transmission hardware can be described as the ratio of the output power P_{out} of the piano tone transmission hardware to the incident power density P_d of the electromagnetic wave, namely:

$$A_e=\frac{P_{\text{out}}}{P_d} \tag{23}$$

The effective area of the piano tone transmission hardware can also be calculated according to Formula (24).

$$A_e=\frac{G_r\lambda^2}{4\pi} \tag{24}$$

The output power of the piano tone transmission hardware can be expressed as:

$$P_{out}=\frac{V_L^2}{Z} \tag{25}$$

The power density of the incident electromagnetic wave is:

$$P_d=\frac{E^2}{120\pi} \tag{26}$$

From Formula (23) (26), we can get:

$$\frac{V_L^2}{Z}=\frac{E^2}{120\pi}\times\frac{G_r\lambda^2}{4\pi} \tag{27}$$

Therefore, the piano tone transmission hardware factor can be expressed as:

$$AF=\frac{E}{V_L}=\sqrt{\frac{480\pi^2}{Z\lambda^2G_r}} \tag{28}$$

For the 50 ohm piano tone transmission hardware system, the piano tone transmission hardware factor can be expressed as:

$$AF=\frac{9.37}{\lambda\sqrt{G_r}} \tag{29}$$

It is expressed in decibels as:

$$AF = 19.8 - 20\times\log(\lambda) - 10\times\log\left(G_r\right) \quad (30)$$

E. Scope of Application of the Calculation Model of Front Door Coupling

The distance from the piano tone transmission hardware is different, so that the field around the piano tone transmission hardware can be divided into three areas: the induction field area, the radiation near field area (Fresnel area) and the radiation far field area (see Fig. 4).

In the induction field area, piano tone transmission hardware does not radiate power. In the radiation near-field area, where it is close to the piano tone transmission hardware, it is difficult to distinguish the main lobe and side lobe of the pattern. As the distance becomes farther, the envelope of the pattern becomes clearer.

The direction map of the piano tone transmission hardware in the far field almost no longer changes with the distance. The measurement results at infinity from the piano tone transmission hardware are compared with the measurement results at a certain distance from the piano tone transmission hardware. If the difference between the two results is acceptable in engineering, the distance is called the Rayleigh distance, and here it is called the inner boundary of the far field.

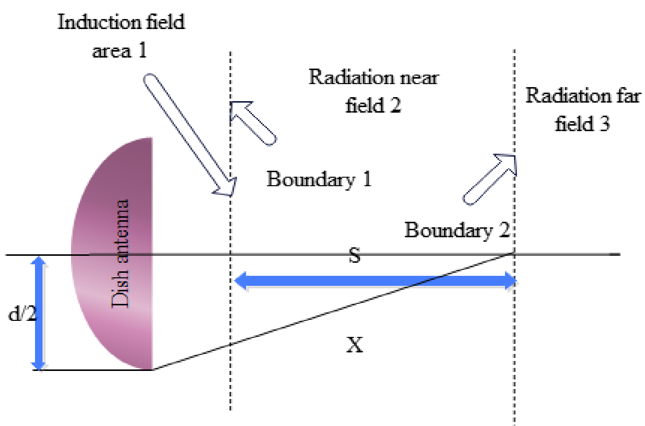


Fig. 4. Division of the piano tone transmission hardware field.

It can be seen from the measurement results that at the position where the Rayleigh distance is located, the distance from the center of the aperture to the observation point is:

$$x - s = \frac{\lambda}{16} \quad (31)$$

From the Pythagorean Theorem, we can get:

$$\left(\frac{d}{2}\right)^2 + s^2 = x^2 \quad (32)$$

Combining Formulas (31) and (32), the expression for the Rayleigh distance S is obtained as:

$$s = \frac{2d^2}{\lambda} - \frac{\lambda^3}{2048} \approx \frac{2d^2}{\lambda} \quad (33)$$

Therefore, the analytical calculation method of the parabolic piano tone transmission hardware pattern obtained in this paper is suitable for the far field.

IV. The Piano Single Tone Recognition and Classification Method Based on CNN Model

A. Empirical Model

The system environment adopts Windows 10+Python 3.5, and the model training adopts a more concise and effective TFLearn model library. Due to limitations in conditions, the TensorFlow used in this article is the CPU (Central Processing Unit) version (the training effect may be better using the GPU (graphics processing unit) version), and the training status information comes from the auxiliary function output of TFLearn.

In order to realize the effective identification and classification of piano single tones, this paper combines the CNN algorithm to carry out research, and adopts the note-level method based on convolutional neural network (CNN) for multi-tone steel piano tone frequency recognition. The model flow is shown in Fig. 5.

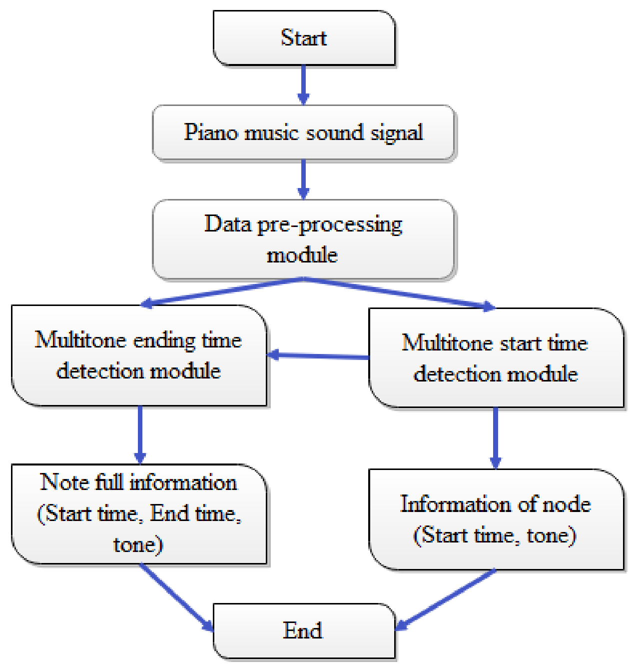


Fig. 5. Structure design of the algorithm system.

Compare the model proposed in this article with the method proposed in study [4], and explore their performance in piano single tone recognition and classification. Quantitative evaluation was conducted using expert evaluation methods to obtain the results shown in Table I.

B. Results

The multi-tone onset time model detects the onset time points of multiple notes, and the note onset time points can see obvious frequency variation characteristics on the spectrogram, as shown in Fig. 6. The CQT video feature map of the F#4 notes has obvious edge mutation, which is very suitable for CNN.

Using FEKO software, the parabolic piano tone transmission hardware is modeled. The piano tone transmission hardware size is: the aperture is 1m, and the center frequency is 5GHz. The normalized direction diagram of the piano tone transmission hardware is obtained as shown in Fig. 7.

The hardware gain analysis calculation method of piano tone transmission in the main lobe is compared with the results of FEKO software simulation. The FEKO software is applied to model the piano tone transmission hardware of the speaker, and calculate the stereo pattern and gain of the piano tone transmission hardware when the frequency is 10GHz, as shown in Fig. 8.

After the above model is constructed, the effect of the model is verified by the piano single tone recognition and classification method based on the CNN model, and the verification is carried out through multiple sets of piano single tone recognition experiments, and the classification results are counted, and the schematic results shown in Fig. 9 are obtained.

TABLE I. PIANO MONOTONE CLASSIFICATION RESULTS

	The method of this article	The method of reference [4]
1	79.51	77.86
2	88.75	72.33
3	80.99	73.95
4	83.72	74.99
5	84.06	71.84
6	87.76	71.30
7	82.92	70.25
8	86.26	68.16
9	78.73	68.42
10	80.32	73.31
11	88.74	73.88
12	83.59	76.15

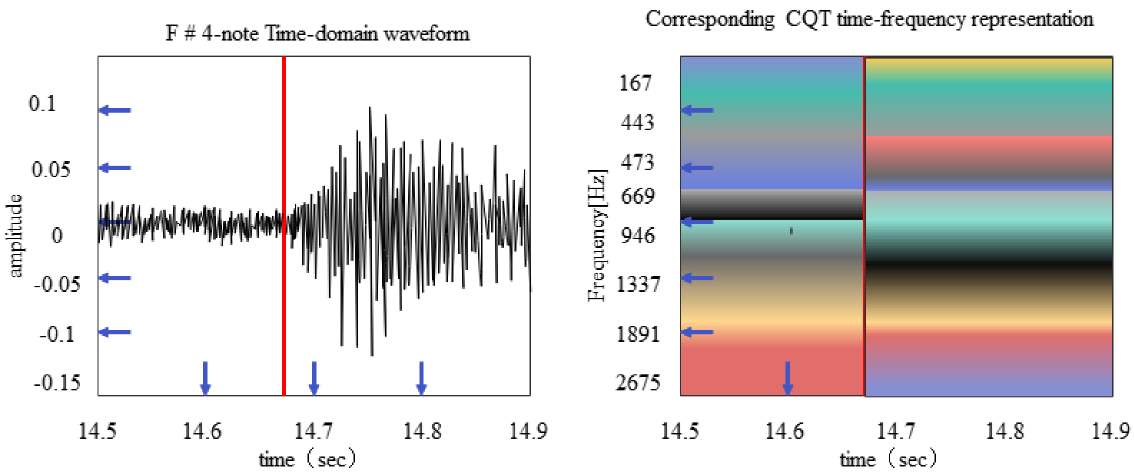


Fig. 6. F#4 note time domain diagram and corresponding CQT spectrum diagram.

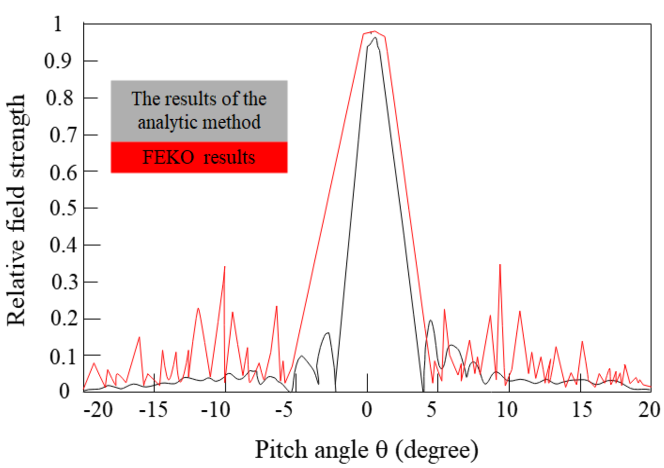


Fig. 7. The normalized direction diagram of the parabolic piano tone transmission hardware simulated in FEKO.

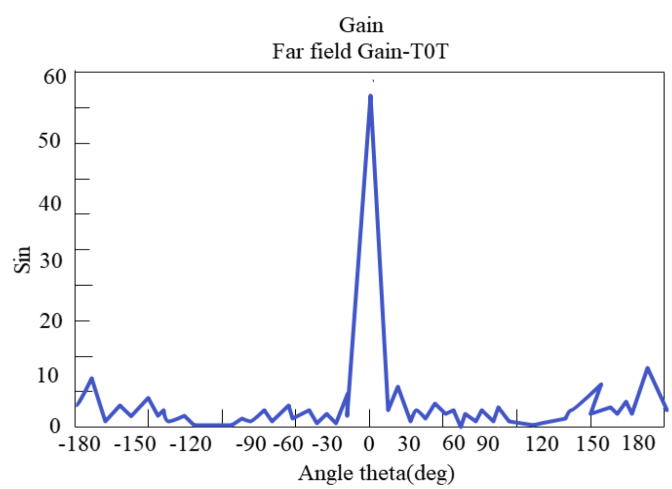


Fig. 8. Gain of piano tone transmission hardware.

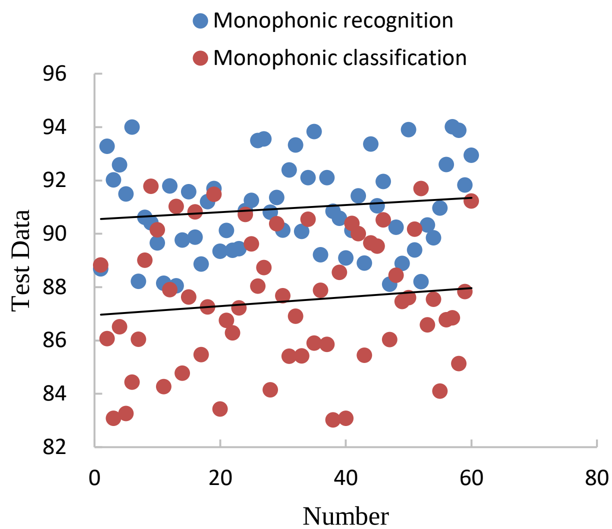


Fig. 9. Verification of the effect of the piano single tone recognition and classification method based on the CNN model.

C. Analysis and Discussion

Nowadays, piano music on the internet generally stores digital audio information, which is generated by sampling analog audio data. The higher the sampling frequency, the larger the amount of digital audio data, and the better the fidelity of digital audio. Digital audio, as a manifestation of acoustic signals, can also be analyzed in the frequency domain and decomposed into pitch components of different frequencies. The frequency domain is composed of many sine functions (or cosine functions, or a combination of both), which have different amplitudes and phases, representing the rich information of audio in the frequency domain angle. The digital audio information stored in piano music is directly unfolded as the energy amplitude trend that changes over time, which is called the time-domain characteristics of audio signals and is the most intuitive representation of the signal. Time domain analysis and frequency domain analysis consider signal characteristics from two perspectives. For digital audio, the amount of data in the time domain is very large. The higher the sampling frequency of audio, the larger the amount of data, which often leads to a large computational workload. Compared to time-domain analysis, frequency-domain analysis has a smaller amount of data and can better reflect some substantive features. Therefore, frequency domain analysis has gradually become the mainstream of signal analysis. Frequency domain features are the results of frequency domain analysis, which describe the basic characteristics of audio signals. Using frequency domain features to represent music is not only easy to implement, but also reduces data volume and facilitates data processing.

FEKO has an algorithm that combines the method of moments as well as the more classical high-frequency analysis methods (physical optics) and consistent diffraction theory. It is extremely suitable for application in analyzing the design layout of piano tone transmission hardware, radar cross section (RCS) and other electromagnetic field analysis problems.

For directional piano tone transmission hardware such as parabolic piano tone transmission hardware, the main consideration is the main lobe pattern and gain of the piano tone transmission hardware. Therefore, the calculation results in this paper can be applied to engineering calculations.

From the results of piano single tone classification, the quantitative evaluation results of the method proposed in this article are distributed between [78, 89], while the quantitative evaluation results of the method in study [4] are distributed between [68, 8]. Therefore, it can be seen that the method proposed in this article has certain advantages compared to traditional methods

From the above research, it can be seen that the piano single tone recognition and classification method based on CNN model proposed in this paper can play an important role in piano single tone recognition.

V. Conclusion

Music and speech are both sound signals, and the basic principles of their recognition are similar. That is to say, all of the sound signals are analyzed, and processing processes such as noise processing, feature analysis, and recognition must be applied. The single-tone signal of the piano, as a sound signal, follows the basic laws of acoustics. In the field of speech research, the current technology is relatively stable and mature, so this paper draws on and refers to the technology of speech recognition and applies it to single-speech recognition. At the same time, there are obvious differences between speech and music: individual differences in speech are large. Even if the same person says the same sentence twice, the sound is quite different and cannot be exactly the same. However, for the same musical instrument, such as the piano studied in this paper, the same key is pressed by anyone at any time, the sound difference is very small, and it has a high degree of acoustic similarity. This paper combines the CNN model to build a piano single tone recognition and classification model to improve the efficiency of piano tone intelligent processing. The research shows that the piano single tone recognition and classification method based on the CNN model proposed in this paper can play an important role in piano tone recognition. The follow-up work of this article is as follows: In terms of note feature extraction, it is necessary to further address the interference of harmonic waves caused by fast rhythms. In the design of convolutional neural networks, further in-depth research is needed on the impact of network structure and loss function design on classification performance.

References

[1] Wang, X. (2018). Research on the improved method of fundamental frequency extraction for music automatic recognition of piano music. Journal of Intelligent & Fuzzy Systems, 35(3), 2777-2783.

[2] Shuo, C., & Xiao, C. (2019). The construction of internet+ piano intelligent network teaching system model. Journal of Intelligent & Fuzzy Systems, 37(5), 5819-5827.

[3] Liu, M., & Huang, J. (2021). Piano playing teaching system based on artificial intelligence–design and research. Journal of Intelligent & Fuzzy Systems, 40(2), 3525-3533.

[4] Lampe, R., Turova, V., & Alves-Pinto, A. (2019). Piano jacket for perceiving and playing music for patients with cerebral palsy. Disability and Rehabilitation: Assistive Technology, 14(3), 221-225.

[5] Johnson, D., Damian, D., & Tzanetakis, G. (2020). Detecting hand posture in piano playing using depth data. Computer Music Journal, 43(1), 59-78.

[6] Gruhn, W., Ristm, R., Schneider, P., D'Souza, A., & Kiilu, K. (2018). How stable is pitch labeling accuracy in absolute pitch possessors? Empirical Musicology Review, 13(3-4), 110-123.

[7] Schutz, M., & Gillard, J. (2020). On the generalization of tones: A detailed exploration of non-speech auditory perception stimuli. Scientific reports, 10(1), 1-14.

[8] Nan, Y., Liu, L., Geiser, E., Shu, H., Gong, C. C., Dong, Q., ... & Desimone, R. (2018). Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children. Proceedings of the National Academy of Sciences, 115(28), E6630-E6639.

[9] Dean, R. T. (2022). The Multi-Tuned Piano: Keyboard Music without a Tuning System. Leonardo, 55(2), 166-169.

[10] Mahanta, S. K., Khilji, A. F. U. R., & Pakray, P. (2021). Deep Neural Network for Musical Instrument Recognition Using MFCCs. Computación y Sistemas, 25(2), 351-360.

[11] Hong, Y., Chau, C. J., & Horner, A. (2018). How often and why mode fails to predict mood in low-arousal classical piano music. Journal of New Music Research, 47(5), 462-475.

[12] Bando, Y., & Tanaka, M. (2022). A Chord Recognition Method of Guitar Sound Using Its Constituent Tone Information. IEEJ Transactions on Electrical and Electronic Engineering, 17(1), 103-109.

[13] Jiam, N. T., & Limb, C. (2020). Music perception and training for pediatric cochlear implant users. Expert Review of Medical Devices, 17(11), 1193-1206.

[14] Foley, L., & Schutz, M. (2021). High time for temporal variation: Improving sonic interaction with auditory interfaces. IEEE Instrumentation & Measurement Magazine, 24(7), 4-9.

[15] Chari, D. A., Barrett, K. C., Patel, A. D., Colgrove, T. R., Jiradejvong, P., Jacobs, L. Y., & Limb, C. J. (2020). Impact of auditory-motor musical training on melodic pattern recognition in cochlear implant users. Otology & Neurotology, 41(4), e422-e431.

[16] Dissegna, D., Sponza, M., Falleti, E., Fabris, C., Vit, A., Angeli, P., ... & Toniutto, P. (2019). Morbidity and mortality after transjugular intrahepatic portosystemic shunt placement in patients with cirrhosis. European Journal of Gastroenterology & Hepatology, 31(5), 626-632.

[17] Pamies, S. (2021). Deconstructing modal jazz piano techniques: The relation between Debussy's piano works and the innovations of post-bop pianists. Jazz Education in Research and Practice, 2(1), 76-105.

[18] Miyazaki, K. I., Rakowski, A., Makomaska, S., Jiang, C., Tsuzaki, M., Oxenham, A. J., ... & Lipscomb, S. D. (2018). Absolute pitch and relative pitch in music students in the East and the West: Implications for aural-skills education. Music Perception: An Interdisciplinary Journal, 36(2), 135-155.