International Journal for Housing Science and Its Applications
Publish September 16, 2025. Volume 47, Issue 1 Pages 1055-1067

https://doi.org/10.70517/ijhsa47191

Machine Learning in Teaching Piano Touch Technique and Tonal Expression

Da Li[1,*] and Qiyan Tan[1]
[1] School of Art, Nanchang University, Nanchang, Jiangxi, 330027, China
Corresponding authors: (e-mail: 18970945811@163.com).

Abstract The application of machine learning in the field of music teaching is emerging, especially in piano teaching shows obvious potential. In this paper, we propose a key touch action recognition model based on ADAG-SVM. The Gaussian kernel function is chosen to solve the problem of high-dimensional vectors in the input space and feature space of action representation. And the optimal penalty parameters as well as the kernel radius are obtained to further improve the performance of touch-key action recognition. In addition, the piano timbre feature matrix is extracted, and based on the discrete Fourier transform, a tone synthesis model with editable timbre is established to generate expressive demonstration clips to deepen the students' understanding of tonal expressiveness. The keystroke action recognition model in this paper can predict the change of the angle of the keystroke action of the student's fingers, and provide scientific action guidance and correction for the learners. The LSD and MSD values of piano tones generated by this paper's algorithm are 1.52 and 492.68 respectively, which are lower than those of the comparison algorithm. At the end of teaching, the scores of the C2 class under the intervention of this paper's machine learning method in the piano level evaluation dimension improved by 0.62-1.97 points compared with the traditional teaching C1 class. Students' satisfaction with the piano teaching class under this paper's method ranged from 4.35 to 4.81 points, and the overall satisfaction reached 4.59.The piano teaching method combined with machine learning significantly improved students' keystroke skills and tonal expression.

Index Terms Machine learning, ADAG-SVM, Fourier transform, Keystroke skill, Piano teaching

I. Introduction

With the continuous improvement of the living standard, people begin to pursue some quality life mode, and the emergence of music, to a certain extent, meets the public's needs in the spiritual aspect [1], [2]. Piano playing is a very important part of the music field, and in the process of playing, there are extremely high requirements for the players' key touching skills and tonal expression [3], [4].

In piano performance, keystroke is a very important part and the foundation of performance, which can well interpret the overall musical style [5], [6]. In order to play a complete piece of music and better explain their feelings to the audience, piano players not only need to deeply understand the style of the piece itself and the mood they want to express, but also should analyze the angle and speed of the key touch [7]-[10]. It is because each player has different keystrokes, which can produce different tone effects, and because the piano itself is a very special instrument, the players can give full play to their own advantages [11]-[13]. Therefore, they can combine their own actual situation, practice all kinds of different key touching techniques, according to the characteristics of the piece, take the strengths and make up for the weaknesses, so as to achieve a most idealized performance effect [14]-[16]. The cultivation of piano tonal expression is an organic part of piano teaching except for piano keystroke skills, and for a long time, many teachers have placed the focus on the training of skills and neglected the cultivation of students' tonal expression [17]-[20]. Under the deepening of the education reform, this drawback has gradually come to the fore. Therefore, it is necessary to explore the learning law of students and the teaching law of the piano, to combine piano skills, basic exercises and the cultivation of tonal expressiveness organically, and to promote the overall improvement of students' musical quality [21]-[24].

The article develops a piano key touch action recognition model, which takes the support vector machine algorithm in machine learning as the core, and introduces an adaptive directed acyclic graph to reduce node dependency and hierarchical depth, and improve data classification. The model chooses Gaussian function, which transforms the high-dimensional mapping problem from input space to feature space into the problem of finding the inner product of low-dimensional vectors. Through the piano timbre feature extraction, the piano timbre, pitch and other tonal expressive playing effects are accurately simulated. Based on this, a tone synthesis model with editable timbre is designed by using Fourier transform and combining the piano timbre features, and the editable timbre is realized by modifying the audio parameters. The proposed model is applied to a music academy to analyze the teaching effect of this paper's method on piano key touching skills and tonal expressiveness through teaching experiments.

II. Research on the application of machine learning interventions in piano teaching

II. A.Machine learning algorithms based on support vector machines

Support vector machine [25] is one of the most commonly used algorithms in machine learning, and its basic idea is to find an optimal classification hyperplane to separate the two types of samples as correctly as possible, and to maximize the classification interval while maintaining high classification accuracy, so as to obtain a better generalization ability. In this paper, we take the support vector machine algorithm as the core to construct a teaching method applicable to piano key touching skills and tonal expressiveness.

II. A. 1) Nonlinear support vector machines

For the nonlinear classification problem, the training samples are mapped to a high-dimensional feature space by introducing a nonlinear function $\varphi : \mathrm{R}^d \to \mathrm{H}$, and the optimal classification hyperplane is constructed in this high-dimensional space. Therefore, in the nonlinear case, the classification hyperplane is:

$$w \cdot \varphi(x) + b = 0 \tag{1}$$

The problem of solving the optimal categorical hyperplane can be described as:

$$\begin{aligned} &\min_{w,b,\xi_i} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i \\ &s.t.\ y_i\left[w \cdot \varphi(x_i) + b\right] \geq 1 - \xi_i \\ &\xi_i \geq 0, i = 1,2,\cdots,l \end{aligned} \tag{2}$$

Similar to linear support vector machines, the optimization problem (2)'s can be transformed into solving a bit of a dyadic problem using the Lagrange multiplier method:

$$\begin{aligned} &\max -\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j \varphi(x_i) \cdot \varphi(x_j) + \sum_{i=1}^{l}\alpha_i \\ &s.t.\ \sum_{i=1}^{l} y_i \alpha_i = 0 \\ &0 \leq \alpha_i \leq C, i = 1,2,\cdots,l \end{aligned} \tag{3}$$

Constructing an optimal hyperplane in a high-dimensional feature space requires only knowing the inner product operation $\varphi(x_i) \cdot \varphi(x_j)$ in this space. If a function $K(\cdot,\cdot)$ can be found such that $K(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$, linear classification can be achieved in the feature space.

With basically no increase in computational complexity, the support vector machine successfully uses the idea of the optimal classification surface in the linear classification problem for the nonlinear problem by introducing the kernel function, and realizes the more complex nonlinear classification. The quadratic optimization problem (3) at this point becomes:

$$\begin{aligned} &\max -\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i y_j \alpha_i \alpha_j K(x_i, x_j) + \sum_{i=1}^{l}\alpha_i \\ &s.t.\ \sum_{i=1}^{l} y_i \alpha_i = 0 \\ &0 \leq \alpha_i \leq C, i = 1,2,\cdots,l \end{aligned} \tag{4}$$

And the corresponding decision function and parameter b become respectively:

$$f(x) = sign\left[\sum_{i=1}^{l} y_i \alpha_i^* K(x_i, x) + b^*\right] \tag{5}$$

$$b^* = y_i - \sum_{j=1}^{l} y_j \alpha_j^* K(x_i, x_j) \tag{6}$$
$$(\text{where } x_i \text{ is the standard support vector})$$

II. A. 2) Kernel function

The introduction of kernel functions allows support vector machines to successfully solve nonlinear problems by transforming them into linear classification problems in the feature space, and at the same time can better overcome the dimensionality catastrophe problem. An inner product function $K(x_i, x_j)$ satisfying the following Mercer's condition [26] is called a kernel function.

Mercer's theorem: it is a sufficient condition for any given symmetric function $K(u,v)$ to be an inner product operation in some higher dimensional space that for any function $g(x)$ and $\int g^2(x)dx < \infty$ that is not constantly zero, there are:

$$\iint K(u,v)g(u)g(v)dudv > 0 \tag{7}$$

Establishment.

The choice of different kernel functions produces different support vector machines, and there are three main commonly used kernel functions as follows:

(1) Polynomial kernel function:

$$K(u,v) = (u \cdot v + 1)^d \tag{8}$$

where parameter d is the order of the polynomial.

(2) Gaussian radial basis kernel function (RBF):

$$K(u,v) = \exp\left(-\|u - v\|^2 / 2\sigma^2\right) \tag{9}$$

Where the kernel parameter σ is the width of the Gaussian function. The RBF kernel has excellent generalization properties and is therefore one of the most widely used kernel functions.

(3) Sigmoid kernel function:

$$K(u,v) = \tanh[a(u \cdot v) + b] \tag{10}$$

In addition, there are Gaussian kernel functions, Fourier kernel functions, and so on.

II. A. 3) Shortcomings of Support Vector Machines

As an emerging technology still under development, support vector machine itself has many shortcomings, and many theoretical studies, algorithm implementation and engineering applications of support vector machine still have a lot of problems to be further developed and improved. For example, how to design fast and effective algorithms that can meet the real-time requirements and adapt to a wider range of algorithms. Multi-valued problem classification algorithms based on support vector machines, the optimization problem of selecting kernel functions and kernel parameters, and the incremental learning problem of support vector machines. The existence of these problems has greatly limited the application of support vector machines in many fields.

II. B.ADAG-SVM based piano key touch action recognition method

When using multi-class support vector machines for simple action recognition of human hands, the larger number of action classes and rich feature sets, as well as the shortcomings existing in the support vector machines themselves, make it inevitable that there will be non-classification, misclassification and error accumulation of the sample classes. Aiming at this problem, this paper proposes a piano key touch action recognition method based on Adaptive Directed acyclic Graph with Support Vector Machine Model (ADAG-SVM) to realize the true classification of sample classes and improve the recognition rate. The framework of ADAG-SVM-based human hand simple action recognition is shown in Fig. 1.

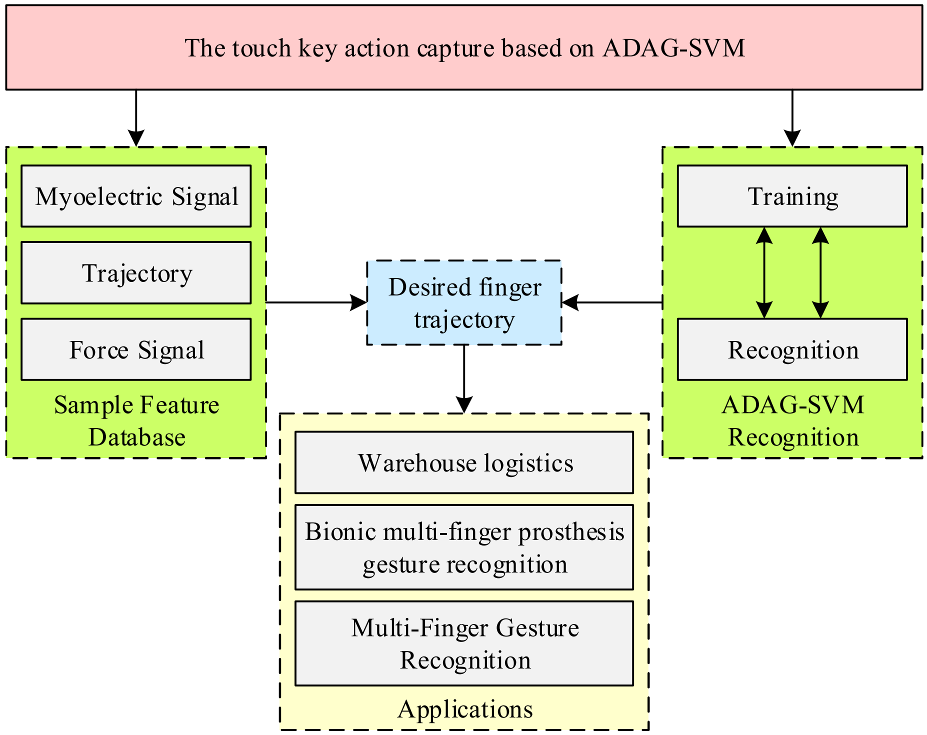

Figure 1: Based on ADAG-SVM's touch key action recognition framework

II. B. 1) ADAG-SVM based classification principle

In order to solve the problems of SVM, this paper proposes an adaptive directed acyclic graph (ADAG) [27] to reduce the ependence on the order of nodes and reduce the hierarchical depth of the directed acyclic graph, and to realize the correct classification of sample data. Adaptive directed acyclic graph uses a hierarchical decision tree structure with an inverted triangle structure and is trained in the same way as decision directed acyclic graph. For a k classification problem, the method includes $k(k-1)/2$ binary classifiers and $k-1$ internal nodes.

The ADAG support vector machine is used to classify 8 sample classes, and the classification principle is shown in Fig. 2. In the figure, the first layer contains 4 nodes and evaluates the bipartite function of each node from this layer, and identifies the sample attributed class in each node by the output value of the bipartite function. After each round, the sample class will be reduced by half and the bi-parametric function of the next level node is selected based on the preferred class of the parent node. The classification process continues until the lowest level node.

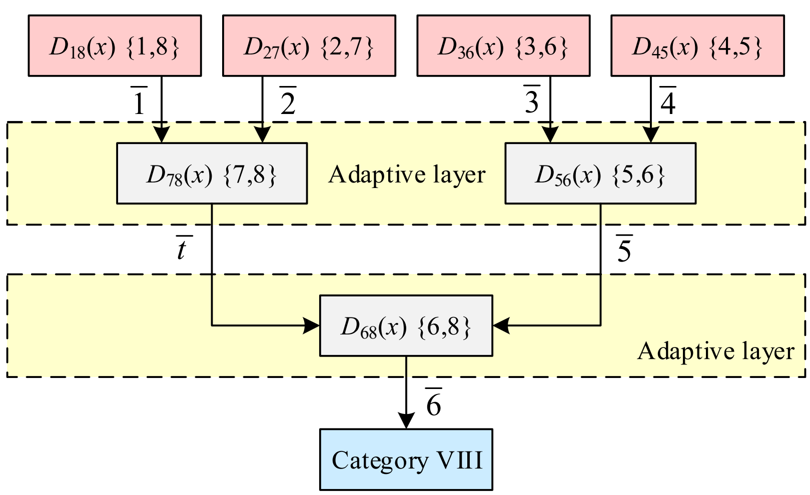

Figure 2: The classification principle of ADAG-SVM

II. B. 2) Kernel function of ADAG-SVM

ADAG-SVM classifies piano key touch action data samples with a recognition rate determined by the type of kernel function used. The main parameters of the kernel function include the penalty parameter and the kernel radius.

(1) Determination of kernel function

When solving linearly indivisible problems, a mapping function is needed to map the sample values to higher dimensional spaces or infinite dimensions. In the feature space, the sample inner product needs to be computed when using the previously mentioned method for solving classification problems in the linearly differentiable case for new samples that are linearly differentiable.

In this paper, the Gaussian kernel function [28] (radial basis kernel function) is used as the kernel function for this experiment as follows:

$$Z(x,y)=\exp\left(-\gamma\|x-y\|^2\right)\ \gamma>0 \tag{11}$$

For a given sample, the performance of adaptive directed acyclic graph is closely related to the type of kernel function selected. In this paper, the Gaussian kernel function (radial basis kernel function) is used as the kernel function for this experiment. After determining the kernel function, the next step will be how to select the best penalty parameter C and kernel radius γ.

Adopting the kernel function not only does not increase the computational complexity, but also characterizes the mapping function as a logical mapping characterizing the mapping relationship from the input space to the feature space. Therefore the choice of kernel function is important for constructing a support vector machine with good performance, if the kernel function is not chosen properly, it is mapped to an inappropriate feature space, which is likely to lead to poor performance in the end.

(2) Selection of Penalty Parameters

Penalty parameter C is a parameter that provides a compromise between obtaining wide boundary spacing and a small number of edge singularities for the constrained Lagrange multipliers. It has been proved by a large number of researchers that the way to find the optimal penalty parameter C is still to try to validate it by growing it exponentially in a certain region, such as $C=2^{-5},2^{-4},\cdots,2^{5}$.

As the value of C increases, the best classification error rate and the best number of support vectors can be obtained when large enough. Continuing to increase the value of C will only increase the training time with no additional improvement in the performance of the support vector machine and will result in an over-learning state affecting the accuracy of the final test set. Therefore, when selecting the optimal penalty parameter, it is necessary to combine past experience with multiple repetitions of trials in a certain region to select the parameter value with better results. Based on this, it is necessary to use a certain selection law together with the kernel radius to obtain the global optimal parameter values.

II. B. 3) Selection of nuclear radii

To obtain the desired kernel radius γ, it is first necessary to develop a reasonable criterion for model evaluation. Traditional support vector machines utilize the interval between two classes of samples as an indicator of generalization performance. However, different kernel functions as well as different kernel parameters lead to different feature spaces, and thus the intervals between them are not comparable. The cross-validation method is one of the classical methods for calculating the generalization error, and its basic idea is to group the sample data, part of which is used as the training set and the other part as the validation set, and then train the classifier through the training set, and then use the validation set to test the model obtained from the training, thus evaluating the performance metrics of the classifier.

In this paper, we use an improved grid search method to obtain the optimal (C,γ). The idea of the improved grid search method is that since there will be different (C,γ) all corresponding to the highest accuracy, the set of (C,γ) with the smallest penalty parameter is considered as the best choice, and then within the neighborhood of that set of parameters, a stepping method is used to find a more optimal (C,γ).

II. C.Models of pitch synthesis with editable timbres

II. C. 1) Extraction of the tone feature matrix

It is found that some specific function waveforms are similar to piano waveforms, and the specific functions that can be selected are sinc function, Gaussian function, and Cauchy function, which can be used to approximate the frequency-domain waveforms of piano music. Let $Y(j\omega)$ be the discrete Fourier transform of the discrete-time signal $y(n)$. Then $Y(j\omega)$ can be expressed by the following equation:

$$Y_i(j\omega) = S_i(j\omega)\cdot F_i(j\omega) \tag{12}$$

$$S_i(j\omega) = \begin{cases} A_i \dfrac{\alpha_i}{\alpha_i^2 + (\omega-\omega_i)^2}, \omega_i - \alpha_i \le \omega \le \omega_i + \alpha_i \\ A_i \dfrac{2}{|\omega-\omega_i|}, \text{Other situations} \end{cases} \tag{13}$$

In the above equation, ω_i is the fundamental or octave frequency of the specified tone, A_i is the amplitude of the fundamental or octave point, α_i is used to adjust the width of the waveform near the fundamental or octave frequency, and F_i is a function of sine and cosine.

When using the Cauchy function weighted by sine and cosine to reconstruct the piano timbre, selecting the appropriate parameters, a waveform can be constructed in the frequency and time domains that is quite similar to the waveform captured from the actual performance. In the frequency domain, the waveforms in the region from the fundamental frequency to the fifth octave and its vicinity well depict the four basic elements of music, namely, piano pitch, timing, intensity and its timbre. And the frequency-domain waveform construction based on the weighted Cauchy function can accurately simulate the timbre, pitch, and intensity of the piano tone by the inverse FFT transform.

II. C. 2) Modeling

In the analysis of timbre composition, all the frequency components in the frequency domain work together to form a full and rich piano timbre, the synthesis of piano timbre can not only use the frequency components of overtones. The next step is to explore the different roles of different frequency components in the formation of piano timbre. Machine learning methods emphasize the role of the fundamental and overtone series in the sound. Based on the Fourier principle [29], different frequency signals are superimposed to form a new sound.

Depending on the degree of concordance of the intervals, the frequency amplitude of the different intervals can be adjusted to obtain a more concordant or discordant musical sound. In this paper, piano tone synthesis is based on the principle of machine learning synthesis, and the Fourier transform algorithm is used to calculate the frequency, amplitude, and phase of different sinusoidal signals in the signal in a cumulative manner using the original signal measured directly.

Based on the discrete Fourier transform and the multiplicative harmonic model of the monophonic signal, combined with the variability of the timbre characteristics, a model of pitch synthesis with editable timbre is established. The model is based on the frequency domain characteristics of the original audio to realize the timbre reconstruction. The timbre parameters in the model are constructed, and the editable timbre is realized by modifying the parameters. After a lot of analysis and experiments, the immutable features in the model are determined to ensure that the main timbre characteristics of the original audio remain unchanged, reflecting the main timbral characteristics of the instrument piano, and meeting the high-fidelity requirements of timbre synthesis. The specific model is shown as follows:

$$x_{f_1}(t) = \sum_{k=1}^{K} \sum_{f\in D_k} s_f v_k A_f \sin\left(2\pi ft + \varphi_f\right) + \sum_{f\in D_k} A_f \sin\left(2\pi ft + \varphi_f\right) \tag{14}$$

In the equation, $x_{f_1}(t)$ denotes the piano tone when the fundamental frequency is f_1, the first term on the right side of the equation denotes the variable feature, and the second term denotes the invariable feature. f denotes the frequency, φ_f denotes the phase angle of the corresponding frequency f in the frequency domain, A_f denotes the amplitude of the corresponding frequency f in the frequency domain, D_k denotes the frequency range of the variable feature in segment k, v_k denotes the amplitude modification coefficient of segment k, and s_f is the weighting coefficient of the corresponding frequency f at the frequency $f \in D_k$, which is determined by the modification line pattern s. The product $(s_f * v_k)$ indicates the actual modification coefficient for the variable frequency of segment k, the product $\left(s_f * v_k * A_f\right)$ indicates the amplitude of the corresponding frequency f after modification, and the constant K indicates that the maximum octave frequency selected is K octave.

II. C. 3) Model analysis

The model divides the timbre feature information of the audio into two parts: the variable feature and the immutable feature. The variable feature enhances or attenuates the effect on the variable frequency (in the range of D), and when the coefficients v_k in the variable feature are all constant 1, the original piano audio timbre remains unchanged, i.e., the reconstruction of the original piano audio is realized.

The amplitude modification coefficient v_k is selected according to the interval concordance relationship. When the octave frequency that is not in harmony with the fundamental is removed (i.e., Order $v_k = 0$), the octave frequency that is not in full harmony with the fundamental is reduced (i.e., Order $0 < v_k < 1$), and the octave frequency that is in full harmony with the fundamental is increased or unchanged (i.e., Order $v_k \geq 1$), then a more harmonic tone is obtained. Conversely, increasing the dissonant octave frequency and decreasing the consonant octave component results in a less consonant tone.

Variable frequency range D is satisfied:

$$D = \bigcup_{k=1}^{K} D_k = D_1 \cup D_2 \cup \ldots \cup D_k \cup \ldots \cup D_K, k = 1, 2, \ldots, K \tag{15}$$

where k is the k nd octave and K is the selected maximum octave of K.

The maximum octave ordinal K satisfies the constraint that K is an integer and $K \leq \frac{f_s}{2f_1}$, where f_s is the sampling frequency. When performing tone synthesis, it is necessary to combine the constraints on the value of K and select according to the bass, midrange, and treble regions.

The k th variable frequency range D_k is determined by the synthesis parameter ∂, expressed as follows:

$$f_k - \partial f_1 \leq D_k < f_k + \partial f_1 \tag{16}$$

where f_1 represents the fundamental frequency, f_k represents the k-octave frequency, and the synthesis parameter ∂ has a value in the range $\partial \in [0, 0.5]$.

The larger the value of the synthesis parameter ∂, the larger the variable frequency range D_k of the k nd segment. When $\partial = 0$, $D_k = f_k$, then only a few isolated points in the entire frequency range $f_k\left(k = 1, 2, 3, \ldots, K\right)$ are modified tonally. When $\partial = 0.5$, $D_k = \left[f_k - 0.5f_1, f_k + 0.5f_1\right)$, $D_{k+1} = \left[f_{k+1} - 0.5f_1, f_{k+1} + 0.5f_1\right)$, because of the theory $f_k = kf_1$, $f_{k+1} = f_k + f_1$, then the variable frequency range D is continuous over the entire frequency range and can be expressed as follows:

$$D = D_1 \cup D_2 \cup D_3 \cup \cdots \cup D_K = \left[0.5f_1, \min\left(f_s, \left(f_K + 0.5f_1\right)\right)\right) \tag{17}$$

When the amplitude modification coefficient v_k of paragraph k is set large or small, it may lead to the following situation: the modification coefficient of paragraph k is very large, and in the frequency domain, the amplitude of $f \in D_k$ is too high, and the amplitude of the non-variable frequency in the vicinity of $f \notin D_k$ remains unchanged, resulting in the amplitude of the frequency in the entire frequency domain range having a sudden and precipitous increase or decrease, and the amplitude of the frequency in the frequency domain changes in a steep and discontinuous manner, resulting in the formation of a murmur and a loss of tone quality.

II. D.Piano Music Pitch Expression Dataset

II. D. 1) Data acquisition

Three performers were recruited to record music data for this dataset. The first performer, earned a doctorate in piano performance in 2022. The second performer, a graduate piano student at a conservatory. The third performer, a graduate engineering student at a university, has been studying piano for 16 years. In order to cover a diverse range of piano tonal expressiveness, the dataset contains Chinese and foreign classical compositions, folk songs, and popular compositions. Recordings of each performer were collected over a period of 30 to 70 days.

II. D. 2) Data pre-processing

In order to unify the experimental conditions, the audio files in the dataset were first downsampled and converted to ensure that all the audio was monophonic, and the sampling rate was set to 44.1 kHz. Meanwhile, considering that as a public dataset, the data should be in a format that occupies less storage space, the audio format was set to MP3 format.

Subsequently, audio files that are too short in duration, contain too many muted clips or have too much ambient noise are excluded through filtering. At the same time, according to the sheet music score, each song was sliced by sentence, and the total number of music fragments after slicing was 1496. Mel spectrograms were extracted for 1496 music clips. The image was also enhanced, including adjusting the contrast, brightness and sharpness of the image. The Gamma correction values for contrast and brightness were 1.8 and 1.6 respectively.Gamma value is a nonlinear method of adjusting the brightness of an image and is commonly used in image processing and display. In digital images, the brightness value of a pixel is usually encoded in a linear manner, i.e., there is a direct linear relationship between the brightness value and the actual brightness. Adjust the sharpness by setting radius=1.5 and AMOUNT=1.5. Combined Enhancement may have a somewhat moderate localized effect, making certain details of the image more prominent or sharper, without causing overly significant changes to the overall image.

III. Analysis of the results of model simulation and pedagogical applications

III. A. Prediction of average inclination of piano touch action

In order to assess students' playing movements in real time while performing piano keystroke technique learning, this paper utilizes the finger movement recognition model described above in conjunction with the design of a piano teaching robot. The thumb exoskeleton was utilized to realize the teaching of thumb piano touching movement, and the standard change curve of angle θ was developed at the same time. Through the analysis of the sample data, it is found that the change of angle θ is more flexible, so after eliminating the singular value caused by the deformation of the hand movement, the single cycle of the thumb piano touching action is divided into four phases: hovering, dropping the finger, touching the key and lifting the finger, and take the average value of the moments of hovering and touching the key θ to formulate the standard curve of the change of the angle θ, which is shown in Fig. 3. The average inclination angle of finger θ= -1.31° during a single playing key touch cycle. The expression for θ is given below:

$$\begin{cases} \theta = -4.2 & 0 \le t \le 0.1167 \\ \theta = 29.5t - 7.64 & 0.1167 < t \le 0.3167 \\ \theta = 1.7 & 0.3167 < t \le 0.533 \\ \theta = -29.5t + 17.43 & 0.5333 < t \le 0.7333 \\ \theta = -4.2 & 0.7333 < t \le 0.85 \end{cases} \tag{18}$$

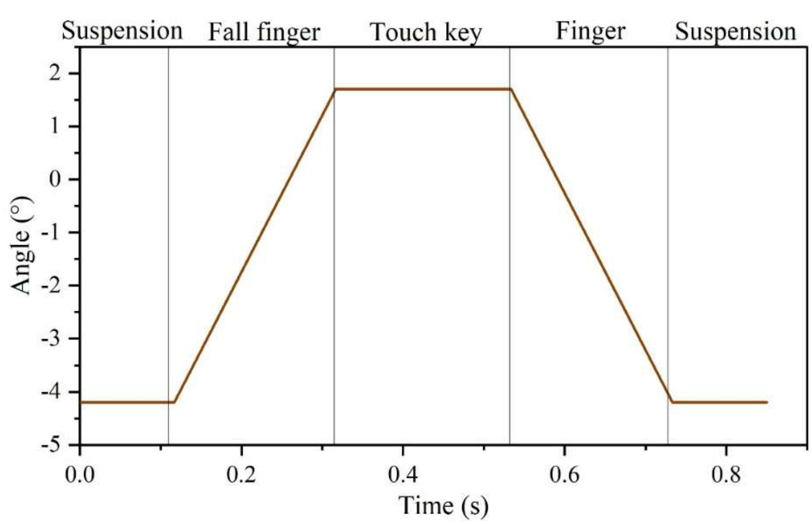

Figure 3: Standard curve

Since the high lifting finger needs to make a uniform motion during the lifting and lowering process, the motion setting of the lifting and lowering phase can be performed by predicting the angle values of the thumb at the moment of hovering and key touching, and then following the uniform motion. For the two extreme thumb lengths, H=90mm and H=100mm, the rest of the angles were predicted for the hovering and keystroke stages, and the angle prediction results are shown in Table 1. The deviation of each angle under the two extreme thumb lengths was not more than 0.5°, so the average value was taken as the prediction angle of 90~100mm thumb length for keystroke, and the prediction curves of the angle changes were plotted as shown in Fig. 4.

In the evaluation of the teaching effect of piano teaching, by measuring the change curve of the tilt angle θ of the middle finger of the learner, the change curve of the rest of the angle can be predicted, and by comparing the predicted curve with the actual angle curve of the learner, scientific and intuitive movement guidance and error correction can be provided for the learner to improve the learning efficiency. In the use of finger exoskeleton for piano teaching, the standard motion curve for the rest of the thumb angle of 90~100mm length can be formulated according to the standard motion curve of θ, which can provide the ideal curve of joint motion for the design of the structure and control system of the piano teaching manipulator in the future.

Table 1: The result of angle prediction

Angle	H=90mm Angle (Suspension to Touch key)(°)	H=100mm Angle (Suspension to Touch key)(°)	Average angle (Suspension to Touch key)(°)
1	48.1~35.2	47.9~34.8	48~35
2	28~17.3	28~16.7	28~17
3	149.8~157.1	150.2~156.9	150~157
4	136.1~154.2	135.9~153.8	136~154

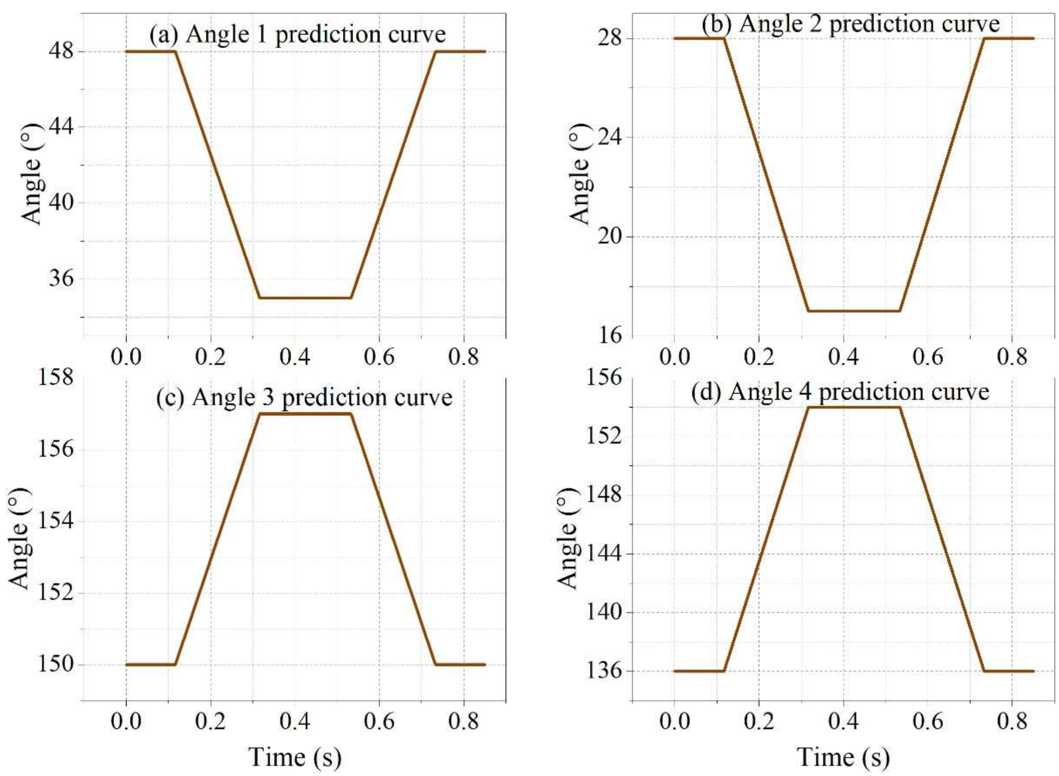

Figure 4: Angular change prediction curve

III. B. Analysis of objective and subjective evaluation of tone generation

The objective evaluation of tonal expressiveness is usually done by choosing some specific parameters to represent the degree of distortion of the timbral signal after processing. In order to intuitively distinguish the differences from the spectra, this paper adopts Log Spectral Distance (LSD) and Mel Spectral Distortion (MSD) based on the auditory model to analyze the experimental results with objective indexes. The music speech signal was divided into frames according to the frame length of 30ms, and the LSD and MSD eigenvalues were calculated for each frame, and then divided by the average of the total number of frames as the distortion level. The formula for calculating LSD and MSD for each frame is as follows:

$$LSD = \sqrt{\frac{\sum_{k=0}^{N-1}\left[10\log_{10}\left(\frac{|X(k)|^2}{|\hat{X}(k)|^2}\right)\right]^2}{N}} \tag{19}$$

where N is the frequency point, $|X(k)|^2$ and $|\hat{X}(k)|^2$ are the power spectra of the original and estimated signals, respectively.

$$MSD = \sqrt{\frac{\sum_{m=1}^{M}\log E_m - \log \hat{E}_m}{M}} \tag{20}$$

where M is the number of Mel filter banks, and E_m and $\hat{E}_m$ are the energy of the original and estimated signals passing through the mth each Mel filter, respectively.

In the objective evaluation experiments, 10 sound sources as well as 4 sets of comparison models are selected from the above dataset and the average LSD and MSD values are measured separately. The LSD and MSD results of different algorithms are shown in Table 2. From the data in the table, it can be seen that the LSD and MSD of the method used in this paper are the smallest, which are 1.52 and 492.68, respectively, indicating that the algorithm in this paper generates tones in a better quality and degree of distortion than the other four algorithms.The MSD of LPC is smaller than that of PSOLA, indicating that the LPC method ensures that the spectral envelope is not distorted, and the LSD is the largest, which indicates that the quality of the tones is distorted to a greater extent. Variable pitch processing is a distorted signal than the original source, so Mel spectral distortion are relatively large.

Table 2: LSD and MSD results for different algorithms

Algorithm	LSD	MSD
TDRS	2.17	937.26
PSOLA	1.86	856.13
LPC	2.44	706.26
ILPPS	1.71	579.94
Ours	1.52	492.68

Naturalness and clarity are difficult to derive from the mapping and indicators, this paper adopts the mean opinion score (MOS) method to listen to the processed speech master observations and give a comparative evaluation of the resultant scores. The MOS value is divided into 1-5 points, with a total of five levels, and the mean opinion score criteria and definitions are shown in Table 3.

Table 3: Average opinion indexing and definition

MOS	Grade	Definition of quality assessment
5	Superior	Unperceived distortion
4	Good	I just noticed the distortion, but didn't hate it
3	Medium	Detect distortion and hate it a little
2	Worse	Hate, but it's not offensive
1	Very bad	It's disgusting, it's disgusting

In order to ensure that the experimental results are comparable, the test tones used in the subjective evaluation experiment are consistent with those in the objective evaluation experiment. In this paper, 30 people were invited to evaluate the pitch-altered speech using the average opinion score method, and the results of the MOS scores of the generated tones are shown in Figure 5. As can be seen from the comparison in the figure, all the methods used in this paper outperform the other compared methods. For example, compared with the ILPPS algorithm, this paper's algorithm improves significantly. The average opinion score of piano tones generated by this paper's method is between 3 and 5, with an average score of 4.1. In contrast, the score of ILPPS is distributed between 2 and 4, with an average score of only 3.3. To summarize, the generation algorithm proposed in this paper has better naturalness and clarity in the expressiveness of piano tones.

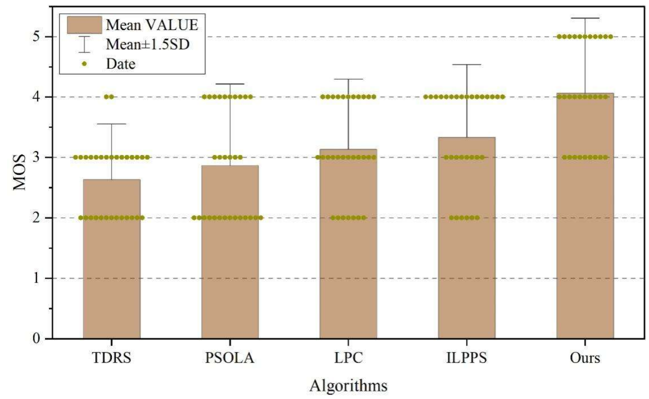

Figure 5: MOS score generated in the generator

III. C. Effectiveness of piano teaching based on machine learning

This section relies on the machine learning algorithm proposed above and applies it to the teaching of keystroke skills as well as tonal expressiveness in the piano program of a conservatory. Two parallel classes with no significant difference in piano ability were selected for control. The experimental class (C2) was taught keystroke technique and tonal expressiveness under the intervention of this paper's algorithm, while the control class (C1) was taught in a traditional way, and the piano instructor was the same in both classes.

At the end of the teaching action, a performance test was given to the students based on the teaching content. In the same piano room, students in both classes used the same piano pieces and rated each dimension from 1-10 according to the teacher in terms of the seven dimensions of playing, intonation, fingering, rhythm, phrasing, spectral notation, and completeness. In order to understand the learning outcomes of the students and the level of skill acquisition under both teaching styles.

Figure 6 demonstrates the results of the learning outcome score statistics for the two classes of students at the end of the instructional action. From the figure, it can be seen that there are large differences between the two classes in playing, intonation, fingering, rhythm, phrasing, notation, and completeness. Specifically analyzed, the average scores of the students in class C2 in each dimension increased by 0.62 to 1.97 points compared to class C1. The largest gap between the two classes was in the dimensions of piano intonation and fingering. Classroom observations revealed that this paper's algorithm for teaching piano keystroke technique and tonal expressiveness provides feedback and captures students' poor keystroke habits and optimizes tonal control.C2 classes can be taught in a richer and more lively way so that students can better understand and grasp the stylistic moods of the piece, which in turn provides a musical treatment of the piece as a whole in terms of their own understanding.

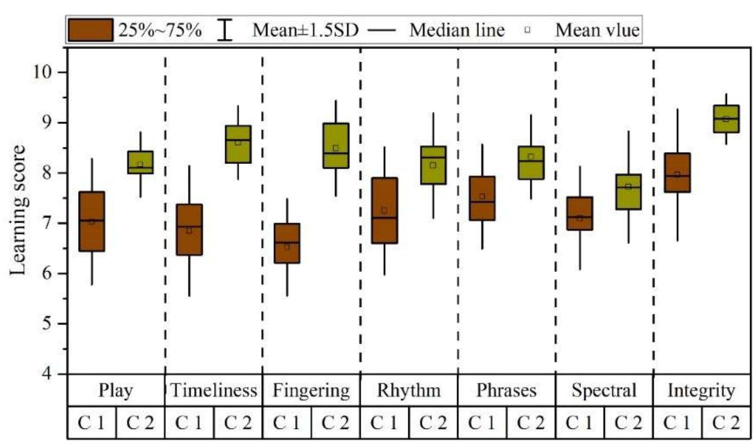

Figure 6: The results of the results of the results of the students' learning scores

III. D. Feedback from Student Satisfaction Survey on Teaching Quality

This study used the Student Teaching Quality Evaluation Scale, which investigated the teaching satisfaction of 30 experimental classes conducting this paper's algorithms for piano key touching skills and tonal expressiveness through two phases of data collection and modeling tests, and achieved good indices in content validity, internal consistency reliability, convergent validity, discriminant validity and predictive validity. The questionnaire survey as well as descriptive statistical analysis were used to quantify the satisfaction for each dimension. The survey dimensions, including teaching content, teaching attitude, teaching effect, and teaching method, satisfaction was scored according to a five-level scale, with higher scores representing higher satisfaction.

The results of the average rating of students' satisfaction with each dimension in the statistical experimental class are shown in Figure 7. From the figure, it can be clearly concluded that students have the highest satisfaction with the teaching effect in the piano classroom with the intervention of the algorithm in this paper, with an average score of 4.81. It shows that in the piano classroom assisted by the machine learning model, students can master the piano key touching skills well and apply them in other repertoire independently, and the students' knowledge can be expanded more, and they get more opportunities to think independently. Students' satisfaction with the teaching content, teaching attitude and teaching method were 4.66, 4.35 and 4.53 respectively, which were slightly lower than the teaching effect, but also showed higher satisfaction. In terms of teaching content, the methodological interventions in this paper resulted in higher scores as teachers would design more interesting and student-friendly content. They also create a relatively relaxed classroom atmosphere based on students' personality characteristics to mobilize students' motivation and participation. In terms of teaching methods, students are given a positive learning experience through diverse symbolic interactions to strengthen their intrinsic motivation.

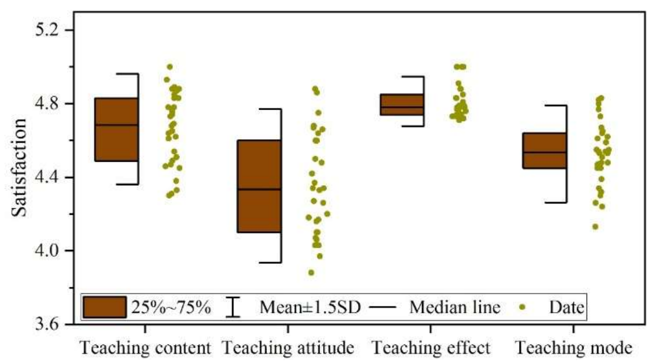

Figure 7: The average score of satisfaction of each dimension

IV. Conclusion

The research aims to utilize machine learning methods to achieve the teaching of piano keystroke skills and tonal expressiveness, for which an adaptive directed acyclic support vector machine-based model is constructed for the recognition and instruction of students' keystroke movements. The model chooses Gaussian function to map the sample values to a higher dimensional space to solve the linear indivisibility problem. The grid search method is used to determine the minimum penalty parameter as the range of kernel radius to optimize the model performance. In addition, based on the Fourier principle, different frequency signals of piano tones are superimposed to construct a tone synthesis model with editable timbre. The feasibility of this paper's method in piano teaching is evaluated through simulation experiments and practical applications.

(1) Model Simulation Experiment

The keystroke action model divides the keystroke action into four stages: hovering, finger drop, key touch, and finger lift, and the average values of the hovering and key touch moments can be calculated to formulate a scientific standard change curve of the keystroke angle. The average MOS value of the pitch generated by this method is 4.1 points, which is more than 24.24% higher than the comparison algorithm.

(2) Piano Teaching Example

The method of this paper can improve the students' piano intonation and fingering dimensions the most, the difference between the experimental class and the control class is 1.78 and 1.97 respectively, and the students' satisfaction with the piano teaching effect under the method of this paper is the highest 4.81 points, which indicates that the method of this paper can provide students with good teaching of keystroke skills and pitch expression.

References

[1] Ivanović, S. (2015). Influence of music on spiritual development of peoples. QUALITY OF LIFE (BANJA LUKA)-APEIRON, 12(3-4).

[2] Creech, A., Larouche, K., Generale, M., & Fortier, D. (2023). Creativity, music, and quality of later life: A systematic review. Psychology of Music, 51(4), 1080-1100.

[3] Springer, D. G., & Sorenson, R. A. (2024). Evaluations of solo piano performances: The role of performing with and without a musical score. Psychology of Music, 03057356241259119.

[4] Kim, S., Park, J. M., Rhyu, S., Nam, J., & Lee, K. (2021). Quantitative analysis of piano performance proficiency focusing on difference between hands. PloS one, 16(5), e0250299.

[5] Goebl, W. (2017). Movement and touch in piano performance. Handbook of human motion, 1-18.

[6] Xu, H., Luo, Y., Wang, S., Darrell, T., & Calandra, R. (2022, October). Towards learning to play piano with dexterous hands and touch. In 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (pp. 10410-10416). IEEE.

[7] Hu, Y. (2024). Research on Piano Performance and Emotional Expression. The Educational Review, USA, 8(9), 1151-1155.

[8] Morijiri, Y., & Welch, G. F. (2022). Decisions on the quality of piano performance: Evaluation of self and others. Frontiers in Psychology, 13, 954261.

[9] Cao, X. (2020). Negative emotions and psychological problems of piano learners and countermeasures. Revista Argentina de Clínica Psicológica, 29(1), 887.

[10] Gao, H. (2024). Revolutionary Virtual Reality Applications Empowering Machine Learning for Teaching Piano Touch Technique and Tonal Expression. Computer-Aided Design and Applications, 21(S17), 63-75.

[11] Flückiger, M., Grosshauser, T., & Tröster, G. (2018). Influence of piano key vibration level on players' perception and performance in piano playing. Applied Sciences, 8(12), 2697.

[12] Verdugo, F., Pelletier, J., Michaud, B., Traube, C., & Begon, M. (2020). Effects of trunk motion, touch, and articulation on upper-limb velocities and on joint contribution to endpoint velocities during the production of loud piano tones. Frontiers in Psychology, 11, 1159.

[13] Şuteu, C. L. (2021). The collaborative pianist–Skills and excellence. Bulletin of the Transilvania University of Braşov, Series VIII: Performing Arts, 14(Suppl), 177-182.

[14] Simones, L., Rodger, M., & Schroeder, F. (2017). Seeing how it sounds: observation, imitation, and improved learning in piano playing. Cognition and Instruction, 35(2), 125-140.

[15] Khachatryan, D. (2022). Some piano performance and technique difficulties. Sciences of Europe, (108), 3-6.

[16] Stana, A. (2023). Aspects of Technique, Mechanics, and Piano Performance in a Diachronic Perspective. Învăţământ, Cercetare, Creaţie, 9(1), 443-452.

[17] Thompson, M. B., & Thompson, M. B. (2018). Piano Tone and Technique. Fundamentals of Piano Pedagogy: Fuelling Authentic Student Musicians from the Beginning, 47-60.

[18] James, B. (2018). Pianism: Performance communication and the playing technique. Frontiers in psychology, 9, 2125.

[19] Meng, C. (2024). Expressive movements in piano performance: The inducing factors. Journal of Human Movement Science, 5(1), 74-79.

[20] Li, Y. (2022). Application of Composite Nanomaterials in the Teaching of Piano Touch Technology and Timbre Expressiveness. Advances in Materials Science and Engineering, 2022.

[21] Hatipova, I. (2021). Exercises for piano and their role in the pedagogical activity of Vitalii Sechkin. Review of Artistic Education, (21), 26-30.

[22] Dogantan-Dack, M. (2017). Expressive freedom in classical performance: insights from a pianist-researcher. Musicians in the Making: Pathways to Creative Performance, 1, 131.

[23] Smith, S. (2021). Expressive piano technic: basics of healthy movement, technic and interpretation. Rast Müzikoloji Dergisi, 9(2), 2757-2768.

[24] Yang, X. (2024). Teaching Design and Practical Effect Exploration of Incorporating Chinese Piano Works into Piano Teaching in Colleges and Universities. Revista Electronica De Leeme, (54).

[25] K. González León,R.J. Delgado Macuil,B. Vertti Cervantes,S. Muñoz Aguirre,J. Castillo Mixcóatl,M. García Juárez... & G. Beltrán Pérez. (2025). Application of support vector machine technique to optical fiber biosensors for neuroprotector (IL-10) detection in serum samples of murine model. Optics and Laser Technology,186,112629-112629.

[26] Kudaybergenov Karimbergen,Arziev Allabay,Orinbaev Parakhatdiin & Tangirbergen Aisulu. (2023). THE MERCER'S THEOREM FOR PARTIAL INTEGRAL OPERATORS. Journal of Mathematical Sciences,271(6),749-761.

[27] Xue Yaxu,Yu Yadong,Yin Kaiyang,Du Haojie,Li Pengfei,Dai Kejie & Ju Zhaojie. (2022). Using Adaptive Directed Acyclic Graph for Human In-Hand Motion Identification with Hybrid Surface Electromyography and Kinect. Symmetry,14(10),2093-2093.

[28] Caiping Jin,Jingxin Zhang & Yonglin Sun. (2025). Numerical simulation of the fluid and flexible rods interaction using a semi-resolved coupling model promoted by anisotropic Gaussian kernel function. Theoretical and Applied Mechanics Letters,15(1),100520-100520.

[29] Jesús Villa,Jorge Luis Flores,Guillermo Garcia Torales & Manuel Montes Flores. (2025). Fourier transform digital moiré interferometry. Measurement Science and Technology,36(3),035208-035208.