

Research Article

Multimedia Vocal Performance Automation Evaluation Model Based on RBF Network

Yu Wang 

Shanghai Normal University, Shanghai, 200233, China

Correspondence should be addressed to Yu Wang; wangyu0619@shnu.edu.cn

Received 29 November 2021; Revised 13 December 2021; Accepted 10 January 2022; Published 28 January 2022

Academic Editor: Gengxin Sun

Copyright © 2022 Yu Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Aiming at the problems of the radial basis network model, this paper proposes a multimedia vocal singing automation evaluation network model, combined with the characteristics of multimedia modeling innovation design, and proposes a two-level comprehensive model. First of all, the theory and algorithm of analytic hierarchy process and radial basis function network are researched and analyzed, and the RBF is predicted for the mature area of multimedia development based on the three indicators of the total amount of classified vocals. The prediction scheme evaluation system is then used to fit the prediction data and influencing factors using the RBF network, and then the classified vocals are adjusted and synthesized hierarchically, and a multimedia vocal classification prediction model is established. Finally, this paper uses an example to verify the feasibility performance and prediction accuracy. Based on the above theory, the experiment uses VC++ 6.0 and Matlab 6.5 combined with database technology to initially realize the evaluation system and achieves a good evaluation effect. The simulation results show that three different algorithms are used to establish RBFO content prediction models. The correlation coefficient limit, root mean square error prediction, and relative analysis error (RED) reached 0.9937, 15.5095, and 8.216, respectively. At the same time, the evaluation results have high accuracy and credibility, which not only provide designers with ideas and improvement basis for innovative designs but also ensure design quality, improve design efficiency, and show that RBF networks have good generalization capabilities.

1. Introduction

Since the RBF network has strong self-organization, self-learning, and associative memory capabilities, it can obtain the weight of the importance of the evaluation target through the learning of existing samples, and it can weaken the human factor in the determination of the weight to ensure it. Therefore, the current radial basis network model is widely used to solve the problems of actual prediction and program evaluation [1]. For this type of problem, it is mainly through learning, memorizing, and associating existing actual data or evaluation results (called learning samples) and then achieving the purpose of prediction or evaluation. RBF network is a type of forward radial basis function network; it has powerful vector classification function and fast calculation ability [2–5]. Evaluation work is an important link in computer-aided innovation design. It helps to

open up new areas of computer-aided innovation design research. Its systematic research will lay the foundation for the development of computer-aided innovation design research based on multimedia vocal music to establish and improve the multimedia design theory and method system will help the scientific, standardized, and quantified multimedia design and evaluation.

RBF network is widely used because of its good approximation characteristics. It is applied to the fields of function approximation, communication adaptive equalization, and pattern recognition [6–9]. However, there are still many problems to be solved in RBF network. The most prominent one is the determination of the network topology; in particular the selection of the number of nodes in the hidden layer of the network and the performance-speed ratio requires further research. The determination of the center value of the layer node still needs the development of the

clustering algorithm. Many scholars have carried out improved research on this, but many algorithms are only suitable for the situation where the network structure is determined, and the network parameters cannot be optimized at the same time. The recursive algorithm is the premise that the RBF network can be applied online and in real time in the control field. At present, it is mainly used to combine the recursive K-mean algorithm and the recursive least squares method, which has achieved certain results, but more effective improvements are needed. The above issues will further promote the application of RBF networks in various fields [10].

This paper proposes a prediction method based on RBF network combined with fuzzy control. The model has fast training speed and good convergence and can greatly reduce the number of hidden layer neurons, and fuzzy control can further correct the prediction error, so this method effectively improves the prediction accuracy. If so, there will be similar functions, causing trouble in processing and time consumption in reasoning. In terms of optimizing the content of fuzzy control, the main consideration is whether the antecedent and subsequent parts of the control rules are in the most reasonable state and whether there are contradictions between different rules. Using the program compiled by the method proposed in this paper, a short-term vocal prediction test was carried out on the actual regional power grid. The results showed that the prediction accuracy was satisfactory.

2. Related Work

The evaluation of multimedia vocal music singing design is a multiobjective decision-making problem. To solve this problem, we must first analyze the characteristics, requirements, and various constraints of the design object and then comprehensively search for various design options to make decisions and select those that meet the design goals and requirements. Recently, many scholars have done in-depth research on the evaluation methods of multimedia vocal music, such as fuzzy comprehensive evaluation method, evaluation method based on gray theory, and artificial radial basis network evaluation method. Although these researches on multimedia design evaluation that serve the entire life cycle of multimedia have achieved certain results, they still have different degrees of limitations, which are generally manifested in the following aspects [11–13].

RBF network model prediction is a method to predict systems with uncertain factors. In forecasting, Pandeya [14] regards the observed data sequence as a random time-varying gray quantity or process, rather than a random process, and gradually whitens the gray quantity through accumulation and subtraction generation, thereby establishing a model relative to the solution of the differential equation, as well as making predictions. Hindawi [15] believes that it is suitable for analysis and prediction under poor information conditions. One of its advantages is as follows: it requires fewer vocal data, simple principles, no consideration of distribution rules, no consideration of change trends, convenient calculations, and

high short-term prediction accuracy. The exponential solution of the gray model differential equation is more suitable for vocal sequences with an exponential growth trend. When the degree of data dispersion is greater, the greater the gray level of the data, the lower the prediction accuracy. Lataifeh [16] found that the function definition of the regression model is similar to that of the classification model. The main difference is that the regression model uses continuous predicted values, while the classification model uses discrete predicted values (such as class labels). In this view, both regression and classification are prediction problems. However, the data mining industry generally believes that predicting continuous values (e.g., using regression methods) is prediction, and using prediction methods to predict class labels is classification. Many problems can be solved by linear regression, and many nonlinear problems can be solved by changing variables and transforming them into linear problems. Based on this qualitative and quantitative analysis, Priyadarshani [17] mathematically processed decision-making problem thinking. RBF uses a relative scale method to process various evaluation factors. It has the advantages of being simple, easy to understand, practical, flexible, and systematic. The field of decision-making problems has been widely used.

After establishing the hierarchical structure, the membership relationship between the upper and lower elements is determined. Rosner [18] believes that the next step is to determine the weights of the elements at each level. For most of the socioeconomic issues, especially the more complex issues, the weights of the elements are not easy to obtain directly. At this time, it is necessary to derive their weights through appropriate methods. Analytic hierarchy process uses the method that the decision-maker gives the judgment matrix to derive the weight. The trend method is based on historical vocal music data and establishes a model by analyzing the development trend of vocal music and predicts the total demand and geographical distribution of vocal music in the future. The prerequisite for using this method is to have enough small area vocal statistics [19–24]. The advantages of the trend method are as follows: the method is simple, only the historical data of the vocal music of the district is needed, and the required amount of data is small. At the same time, its disadvantages are as follows: the smoothness and continuity of the vocal growth curve are poor, and it is not suitable for the vacant district in the past. The land-use simulation method is a top-down prediction method. The future development of the community is simulated through the model, and the result of the system vocal music prediction is allocated to each community. The land-use simulation method has the following advantages: it can make full use of the results of the system's vocal music prediction and predict the spatial distribution of the future distribution network vocal music based on it; at the same time, the land-use simulation model and the vocal music model can be used to analyze in detail the factors affecting the growth of vocal music. It is suitable for multischeme research, can simulate the development of the community, has high prediction accuracy, and can meet the purpose of

long-term planning. But, at the same time, the land-use simulation method requires a lot of spatial data and sufficient computing resources, and the computational workload is relatively large [25–27].

3. RBF Network Parameter Setting

3.1. Analysis of the Radial Basis Equation. The main difference between the various mathematical models of RBF network neurons is that they use different transfer functions, so that neurons have different information processing characteristics. The information processing characteristics of neurons are one of the three major factors that determine the overall performance of the artificial radial basis network.

The transfer function of a neuron reflects the relationship between the output of the neuron and its activation state.

$$L(w, b, a) = \sum \frac{1}{2} W^t W - \sum [\partial y (W^t x + b) - 1]. \quad (1)$$

The nonlinear transfer function is a nondecreasing continuous function from the real number domain R to the closed set of $[0,1]$, which represents the state-continuous neuron model. The change of vocal music has its uncertainty on the one hand.

However, under certain conditions, there is an obvious trend of change. Once the trend of vocal music is found, the future vocal music situation can be judged according to the trend, and the result can be obtained. The main changes in vocal music include horizontal trends, linear trends, and seasonal trends. Their common feature is that they only extrapolate the trend and do not perform statistical processing on the random components.

$$u(P) - [u(x_1), u(x_2), u(x_3), \dots, u(x_{n-1}), u(x_n)] = 0. \quad (2)$$

The neuron model using probabilistic transfer function has an uncertain relationship between input and output, and a random function is needed to describe the probability that the output state is 1 or 0. In competitive learning, the output units of the network compete with each other, and finally only one of the strongest is activated.

The node is also likely to work abnormally or be damaged due to the surrounding environment. The most common situation is that there is a lateral inhibitory connection between the output neurons, so that when there is a certain unit in the original output unit and it is stronger, it will win and suppress other units. In the end, only this strong unit will be active.

$$\begin{cases} y(W^t W + b) \geq 1, \text{ if } (a < b), \\ x(W^t W - a) \geq 1, \text{ if } (a > b). \end{cases} \quad (3)$$

The overlapping range of the membership function should account for at least 25% of the range of a single fuzzy set. The membership function should follow the semantic order and avoid inappropriate overlap. It is worth mentioning that the superposition range between membership functions should not be too large (generally should not

exceed 75%); otherwise, it means that the membership function representing the fuzzy set is too flat, so that, in the actual control process, the change of variables is relatively slow.

3.2. Network Node Evaluation. The RBF network node adopts one-by-one recovery method to check the influence of the 7 samples on the model. First, after preprocessing all samples with a combination of first derivative and SNV, the SPXY algorithm is used to divide 65 calibration set samples and 35 prediction set samples. After analyzing the prediction set samples, it is found that the S-010 sample with a larger residual value is divided into the prediction set. Since it is not determined whether it is an abnormal sample, considering the later model prediction, it needs to be manually divided into the calibration set.

$$\text{if } [abs\{w, t | f(w, t) \in R\} > 0], \sum \min \phi(w) - \frac{1}{2} w^t w = 0. \quad (4)$$

Finally, 66 calibration set samples and 34 prediction set samples are obtained. Then we remove all the tentative abnormal samples from the calibration set and recover them one by one. Finally, Table 1 uses partial least squares regression, RBF network, and radial basis function to establish prediction models, with R , RMSEP, and RPD as the models.

During the process, the overall nature of the particles generated is unchanged, which is a manifestation of the individual's degradation behavior. The continued degradation process of these particles also reflects the individual's future predicted degradation trend.

The parameters of the Gaussian kernel function are usually obtained by the pattern clustering method K-NN. The specific process is as follows: First is initialization, where the parameters of the Gaussian kernel function of each node are determined according to all input samples, including the central value and the initial value of the standardized constant; second is using the K-NN method in the cluster analysis to obtain the parameters of the network kernel function; and third, according to the input sample, we use the algorithm to obtain the weight of the output.

$$\begin{cases} W^t \times x + b \geq 1, \text{ else, } y - 1 = 0, \\ W^t \times x + b \leq 1, \text{ else, } y + 1 = 0. \end{cases} \quad (5)$$

There are two common types of nodes in a radial basis network. One type is to receive input weighted signals from all directions; that is, the collection point of information is called an inner star node; the other type is to send output weighted signals from all directions, that is, the divergence point of information. It is called an alien node. As long as the relationship between input and output is continuously given, in the learning process of the radial basis network, an internal structure representing this relationship will surely be formed within it, and only the speed of the relationship formation can reach a practical value.

TABLE 1: RBF network node evaluation description.

| RBF node | Parameters model 1 | ANN node | Parameters model 2 | KNN node | Parameters model 3 |
|----------|--------------------|----------|--------------------|----------|--------------------|
| 1 | 15 | 1 | 11 | 1 | 12 |
| 2 | 13 | 2 | 34 | 2 | 14 |
| 3 | 40 | 3 | 18 | 3 | 11 |
| 4 | 20 | 4 | 18 | 4 | 24 |
| 5 | 22 | 5 | 53 | 5 | 10 |
| 6 | 54 | 6 | 34 | 6 | 59 |
| 7 | 18 | 7 | 53 | 7 | 17 |

$$g(x) = \left\| \begin{array}{cc} W^t \times x + b & -1 \\ 1 & W^t \times x - b \end{array} \right\| \times \left\| \begin{array}{cc} 1 & -1 \\ -1 & 1 \end{array} \right\|. \quad (6)$$

It can be seen that the radial basis function is only sensitive to the input near the “center.” As the distance from the “center” increases, the output of the function quickly decreases to a very small value, showing a kind of “local sensitivity.” The speed of reduction is determined by the width. Therefore, the energy consumption in the MAC layer network of the wireless sensor network is reduced to a minimum, and the ability to adapt to the frequent changes of the network topology is also required.

The algorithm first arranges the distribution of sampling points through the method of network training and fitting constructs and trains a suitable RBF network according to the characteristics of the sampling points. After training, the prediction performance of the RBF network is just right. The network output can be predicted. The process of genetic algorithm is extreme value optimization.

3.3. Hierarchical Nesting of Radial Basis Networks. This method will cause the particles to lose diversity over time. In order to increase the effectiveness of the algorithm, try to improve the algorithm. Then the remaining life of the RBF network, that is, the remaining time for the system’s fault characteristic value to reach the critical threshold, can be obtained. Due to the uncertainty of the prediction, the remaining life prediction result should be described in the form of probability in Table 2. In order to obtain a point estimate of the remaining life, it is necessary to judge when the degradation prediction estimate exceeds the preset critical threshold.

The first step in designing a fuzzy controller is to select input and output language variables. Too many input and output language variables are selected, and the control process is particularly complicated, and the real-time performance is poor: too few input and output language variables are selected.

Since fuzzy control is a control system that simulates manual operation, we should go back and carefully understand the entire process of manually controlling an object, which may help us to choose the appropriate input and output variables.

$$\lim_{x \rightarrow \infty} [f(x) - f(x-1)] - \lim_{x \rightarrow \infty} \operatorname{sgn}\left\{\sum \gamma \alpha + b\right\} = 0. \quad (7)$$

Local distribution means that when the input falls in a small designated area of the input space, the radial basis

function of the hidden unit makes a meaningful nonzero response. The radial symmetric decay of the center refers to the input with the same radial distance from the center vector c of the basis function.

$$\text{for } (x \rightarrow 1, y \rightarrow x), \frac{\partial \beta \{y[(w \times x) + b]\}}{\partial \beta * x} - 1 = 0. \quad (8)$$

$$\begin{cases} \frac{\partial L(x)}{\partial W} = -\sum \alpha(y+x) = 0, \\ \frac{\partial L(y)}{\partial W} = -\sum \alpha(y-x) = 0, \end{cases} \quad (9)$$

and if the S-type layer has enough neurons, it can train any functional relationship between input and output. In the state-space model in the article, you can try to analyze the influence of different particle numbers on the prediction results. In addition, in order to reduce the impact of particle degradation in the particle filter algorithm, a random resampling method is used.

4. Construction of Multimedia Vocal Singing Automation Evaluation Model

4.1. RBF Network Link. RBF is a commonly used three-layer nonlinear forward radial basis network. The input layer is composed of signal source nodes, the second layer is the hidden layer, and the third layer is the output layer. The transformation from the input space to the hidden space of the RBF neural network is nonlinear, while the transformation from the hidden layer space to the output layer space is linear. Because it simulates the neural network structure of the human brain that is locally adjusted and covers the receiving field (or receptive field), the RBF network is a local approximation network.

$$\frac{\forall L(x, t)}{W(x, t)} \in \sin \theta, \text{ if } (\theta > 0), \sum \frac{\partial L}{\partial W} * w - \sum \alpha y x = 0. \quad (10)$$

It has been proved that it can approximate any continuity with arbitrary precision. For wireless sensor networks, the nodes often rely on batteries, and the batteries are not easy to replace. The RBF network is suitable for the approximation of multivariable functions. As long as the center is selected properly, the RBF network composed of few neurons can obtain better approximation results, and it also has the advantage of the only best approximation point.

Among them, the basis function in the hidden layer produces a localized response to the input excitation; that is,

TABLE 2: Description of radial basis network hierarchy.

| Network step | Input value | Output value |
|--------------|---|--|
| 1 | The main idea of analysis of variance is: | quicksort(srcsum,low,pos-1); |
| 2 | To decompose the total square SST: | void bigRootPile(int * arr, int len) |
| 3 | The regression sum of squares SSR: | int temp = arr(k); |
| 4 | The residual sum of squares SSE: | int pivot = srcsum[0]; |
| 5 | The significance test $M = 2/\sqrt{w^2}$: | int low = 0,high = sizeof(srcsum)/4-1; |
| 6 | The regression equation is $g(x)$: | int srcsum [3] = {101,9,11}; |
| 7 | It is mainly to do analysis of variance: | printf("srcsum(%d) = %d " ,i,srcsum[i]); |
| 8 | Test of the regression coefficient $W^t \times x$: | quicksort(srcsum,pos+1,high); |

each hidden node has a parameter vector called the center, which is used to compare with the network input vector to generate a radially symmetric response. Only when the input

falls in a small designated area, the hidden node will make a meaningful nonzero response. The response value is between 0 and 1.

$$\int L(p, u, v, t^u) dp du dv dt - \int L_{loc}(t^u, v) + L_{cls}(p, u) dp du dv dt = 0. \quad (11)$$

The output unit is generally linear; that is, the output unit performs a linear weighted combination on the hidden node output. After the center and width of the RBF function are selected, the weight coefficients from the hidden layer to the output layer are trained. Since the relationship between the output layers in Figure 1 is a linear equation set, the weight coefficient becomes a linear optimization problem. Various linear optimization algorithms include error correction algorithms based on supervised gradient descent, least squares recursive algorithms, etc.

The adaptation factor value of the island will be continuously changed in the iterative process according to the degree of adaptation of the sample to the island. The population with a relatively large adaptation factor value is retained, other individuals are exchanged, and a new population is generated, avoiding selecting a network from a set of neural networks. When the information stored in the discarded network is lost, the final network after training can be regarded as an optimal solution of the entire solution space.

$$\frac{df(x, y)}{dx dy} - \frac{dsgn\left[\sum_{i=0}^{n-1} a_i \cdot y_i \cdot (x, x_i) + b\right]}{dx dy} = 0. \quad (12)$$

This method is simple due to the calculation process. The speed is faster and it is easy to apply. It has better approximation performance, so the RBF network in this paper adopts this learning method. This method consists of two stages: one is the self-organized learning stage, which is the stage of learning the center and variance of the hidden layer basis function; the other is the supervised learning stage, which is the stage of learning the weights of the output layer.

4.2. Multimedia Pattern Recognition. The H learning rule, also known as the least mean square (LMS) error rule, trains the weights, so that it can get faster convergence speed and higher sugar content compared to the perceptron. However, this network protocol does not take energy consumption as

the first consideration, so it cannot be applied to wireless sensor networks.

Using Matlab to modify the parameters in the program before each simulation will waste a lot of time, so I hope that future researchers can integrate several programs into a simulation system, and it can be implemented quickly by modifying different parameters on the control panel.

$$w(x, y) = \begin{cases} w * \sum_{i=0}^{n-1} a_i \cdot x_i \cdot y_i * \sum_{i=0}^{n-1} x_i \cdot y_i, & x > y, \\ w * \sum_{i=0}^{n-1} x_i \cdot y_i, & x \leq y. \end{cases} \quad (13)$$

The RBF network requires a large number of training samples, and, due to the existence of the number of network layers and the number of hidden layer nodes in Figure 2, redundancy makes the convergence speed of the network slow and cannot guarantee that the network error converges to the global minimum; meanwhile the basis function of the hidden layer node of the RBF network adopts it. The activation function is determined by the distance between the center vector of the hidden unit and the input vector.

$$\min_{w,b} 0.5\|w\|^2, \text{ s.t. } y_i(x_i \cdot w + b) > 1 \text{ for } \{i = 1, 2, \dots, n-1, n\}. \quad (14)$$

The response of the RBF network is completely local. Compared with the RBF network, the RBF network not only has a physiological basis but also has a simpler structure and a faster learning speed. It has great flexibility when approaching a limited area. This relies on a large number of hidden units.

4.3. Automated Evaluation Model Weight Distribution. The form of the nonlinear function used in the RBF network is not critical to the automatic evaluation of network performance. The key factor is the determination of nodes and the selection of the data center. However, the selection of

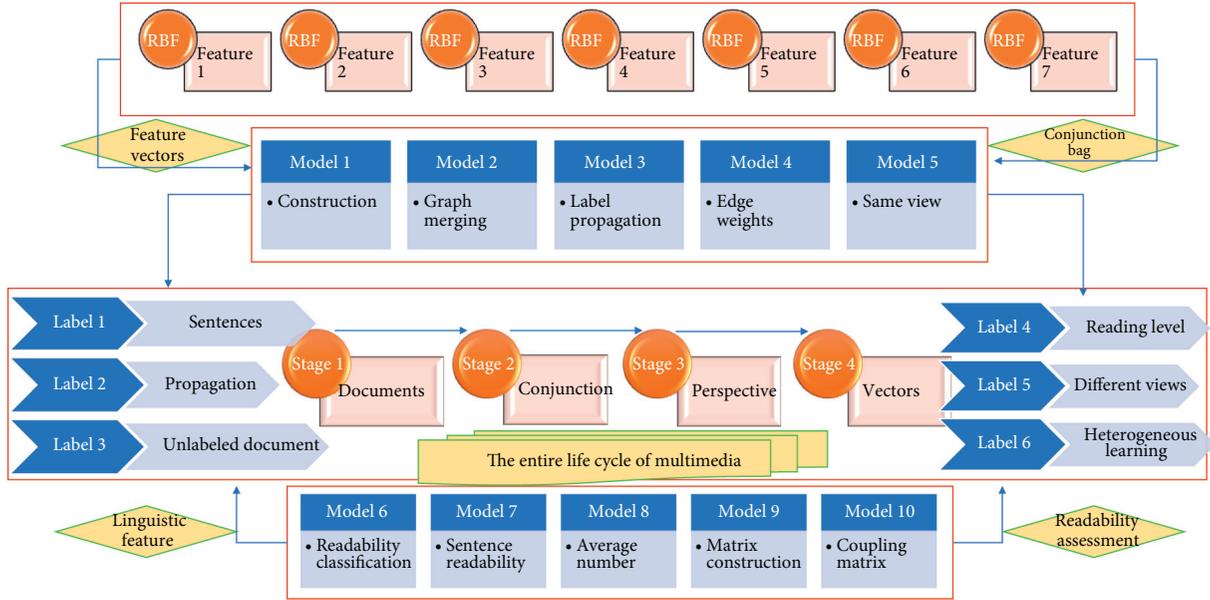


FIGURE 1: RBF network link topology.

points and data centers is particularly difficult to determine, which affects the performance of the RBF network. In view of the problems of RBF networks, many scholars have done in-depth research and proposed new network models. They have made improvements in the determination of data centers and network structures, effectively improving the generalization and recognition capabilities of RBF networks.

$$\begin{cases} \sum x_i \cdot w - \exp(x) * b + 1 & y_i = 1, \\ \sum x_i \cdot w + \exp(x) * b + 1 & y_i = -1. \end{cases} \quad (15)$$

Obviously, if a data point has multiple adjacent data points, the data point has a high density value. Taking the model established by RBF as an example, the analysis data can be obtained: the samples S-010, S-059, S-076, and S-081 with large residual errors are recovered one by one into the calibration set.

The sample S-050 recovered in Figure 3 has significantly improved the performance of the model; that is, the R value is increased from 0.9913 to 0.9914, the RMSEP value is reduced from 16.7741 to 16.6523, and the RPD value is increased from 7.6568 to 7.7128, indicating that the sample has a certain contribution to the model; it is not an abnormal sample, so it is recovered.

After recovering the sample S-001 with an abnormal leverage value, the performance of the model has been improved; that is, the RMSEP is reduced from 16.7741 to 16.7351 and the RPD value is increased from 7.6568 to 7.6746, indicating that the sample has a certain effect on the model.

After recovering the sample S-095 with abnormal leverage and residuals in Table 3, R increased from 0.9913 to 0.9915, the RMSEP value decreased from 0.7741 to 0.5668, and the RPD value increased from 0.6568 to 0.6568.0.7526; the model performance is comprehensively improved, and it needs to be recycled. Similarly, analyzing PLS model and

RBF neural network can get abnormal samples that need to be eliminated, while S-001 and S-095 are pairs of models. All have a certain contribution and need to be recycled.

IEEE 802.11 network protocol cares more about the network bandwidth, network delay, and the fairness of the entire network. As can be seen from the above description, the clustering algorithm selects the data point with the highest density index as the center of the RBF network based on the distance between the data points. If the two data points are very close, they are very representative at the same time. In this case, using the clustering algorithm to select the center vector may miss some meaningful data points in the data set, and the clustering algorithm generally needs to set the number of clusters, and the order of center point selection will also affect the clustering results.

4.4. Vocal Singing Evaluation Factor. In order to obtain enough sample values for vocal performance evaluation, to improve the accuracy of the regression equation in the later stage, and to better ensure that the errors of the uniform design table are sufficiently reasonable, usually the number of levels is chosen to be about times the number of factors. Taking into account the factors selected in this experiment, it is the automatic evaluation of the main multimedia vocal performances that affects product defects, so the average is 17, and, within the range of values, 17 levels are uniformly selected.

$$\frac{\sum Y [i | i = 0, 1, 2, 3, j - 1, j]_j}{\sum_{i=1}^N N \times W_j} - \frac{\sum_{i=1}^N X_i \times W_j}{N} = 0. \quad (16)$$

Here, the degree of freedom is $n-s-1$ distribution. Variables that are not significant after testing are secondary variables. At this time, the variable with the smallest partial regression sum of squares can be removed to form a new

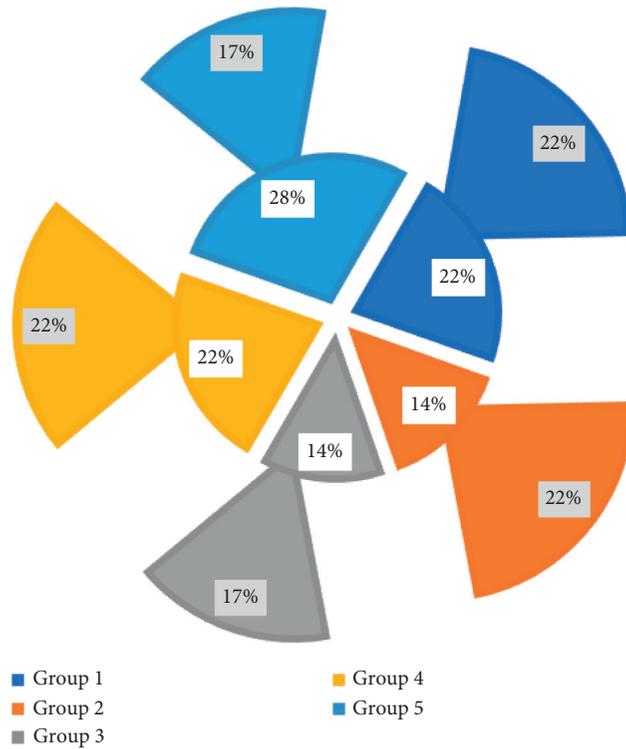


FIGURE 2: Distribution of hidden nodes in RBF network.

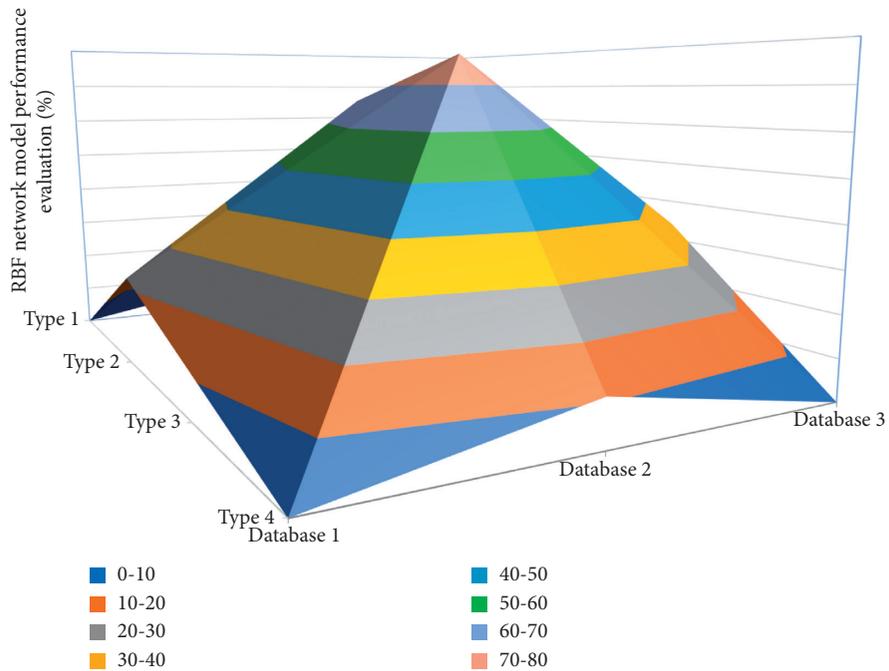


FIGURE 3: RBF network model performance evaluation distribution.

regression equation and test again until the final regression equation is significant and only contains important variable.

The sensor node only listens to the wireless channel when it needs to send and receive data and tries to enter the dormant state when there is no data to send and receive. Because the above calculation process is relatively

complicated, statistical software is often used for auxiliary calculation.

Figure 4 uses time to replace the factors that affect vocal music. According to the past statistical data of vocal music, the law of its changes over time is found, and a time series model is established to infer future vocal values. The basic

TABLE 3: Model performance data description.

| Data group | RPD 1 | RPD 2 | RPD 3 | RPD 4 | RPD 5 |
|------------|----------|---------|---------|----------|---------|
| 1 | 1.82087 | 0.09899 | 0.43799 | -0.20306 | 0.53818 |
| 2 | -0.27033 | 0.12907 | 0.1874 | -4.18358 | 0.34569 |
| 3 | 9.99147 | 0.14925 | 0.53636 | 0.41295 | 0.53928 |
| 4 | 0.20987 | 0.01689 | 0.55578 | -0.78639 | 0.22905 |
| 5 | -1.19597 | 0.24574 | 0.50713 | 1.82087 | 0.09899 |
| 6 | 1.4026 | 0.10238 | 0.01055 | -0.27033 | 0.12907 |

hypothesis is as follows: the changing law of vocal music in the past will continue into the future; that is, the future is a continuation of the past. The main mathematical models include autoregressive (AR) model, moving average (MA) model, and autoregressive moving average (ARMA) model.

In the experiment, 100 samples were tested by lever value and residual T to eliminate 4 abnormal samples, and 96 samples remained. This study uses SPXY algorithm to divide the 96 samples into two parts: correction set and prediction set. The correction set is used to build the prediction model, and the prediction set is used to test and evaluate the performance of the model. When dividing, we divide the set with a ratio of 2 : 1 and select 60 representative samples from 96 samples as calibration set samples and the remaining 36 samples as prediction set samples for the correlation coefficient index of the sample distribution of the prediction set.

$$\sum s(i, j) - s(i-1, j-1) - \sum_m \sum_n w(m, n) \cdot x(i+m, j+n) = 0. \quad (17)$$

According to different application backgrounds, wireless sensor networks are generally used for data collection, detection, emergency alarms, and target tracking. The main purpose of the significance test of the regression coefficient is to eliminate those dispensable variables, which is conducive to better prediction and control. Here, the partial regression sum of squares is mainly used to measure the degree of influence of each variable in the regression and then to judge whether the variable is important or not.

5. Application and Analysis of Multimedia Vocal Singing Automation Evaluation Model

5.1. RBF Network Data Preprocessing. The idea of RBF radial basis network integration and optimization treats the entire network system as an integration of multiple single networks for training. Each single network can be regarded as an island. Each network adopts the same structure and inputs the same training samples.

During the training process, strategies such as training times or error expectations can be used to control, and an adaptation factor is set for each island network. Let the adaptation factors of different networks migrate to each other, and carry out crossover, mutation, and selection.

$$\lim_{x \rightarrow \infty} \frac{\sum_m \sum_n w(m, n) \times x(i+m, j+n)}{\sum_m \sum_n w(m, n)} = 1. \quad (18)$$

The uniform design mainly considers that the test points are evenly distributed within the test range. Compared with the orthogonal design, it has a unique way of distributing points, which is mainly manifested in the following aspects: (1) Each level of each factor is performed twice and only one test is done. (2) The test points of any two factors can be marked according to the increase in the number of levels.

Uniform design has the advantages of arranging less test times and evenly spreading test points within the test range. This considers the use of uniform design to arrange sampling points. In the uniform design in Figure 5, the level of each factor directly affects the accuracy of the fitting model. At the same time, these tests do not have the problem of interaction combination of various areas in the horizontal soil where they are located:

$$s(t) = \left\langle \begin{array}{c} \sum_a w(a) \cdot x(t-a) \\ x(t-a) \end{array} \middle| \begin{array}{c} -x(t-a) \\ \sum_a w(a) \end{array} \right\rangle. \quad (19)$$

The OLS method is derived from the residuals through the orthogonalization of the regression operator. We learn to choose the appropriate regression operator vector and its number, so that the network output meets the quadratic performance requirements: the regression operator is directly constituted by the RBF.

5.2. Simulation of the Automatic Evaluation Model of Vocal Music Singing. Orthogonal least squares algorithm can combine the selection of RBF center with the learning of output weights, so it has been widely used in the automatic evaluation of actual vocal music. However, if the learning data contains noise or the stopping threshold of the learning algorithm is not selected properly, it will cause the OLS algorithm to overfit and reduce the generalization ability of the network. Table 4 provides the relevant radial basis network toolbox in the MATLAB Radial Basis Network Toolbox using the recursive orthogonal least squares (ROLS) method.

The data proves that the uniform design has better robustness than the traditional experimental design. At the same time, the uniform design can arrange an experimental design scheme with more factors and fewer experiments, which can reflect the main laws of changes in things.

Therefore, the uniform experimental design method mainly includes the following steps: firstly we select the test factors and determine the factors; secondly, we select the appropriate uniform design table, arrange the factor levels, and arrange the test data according to the test plan and

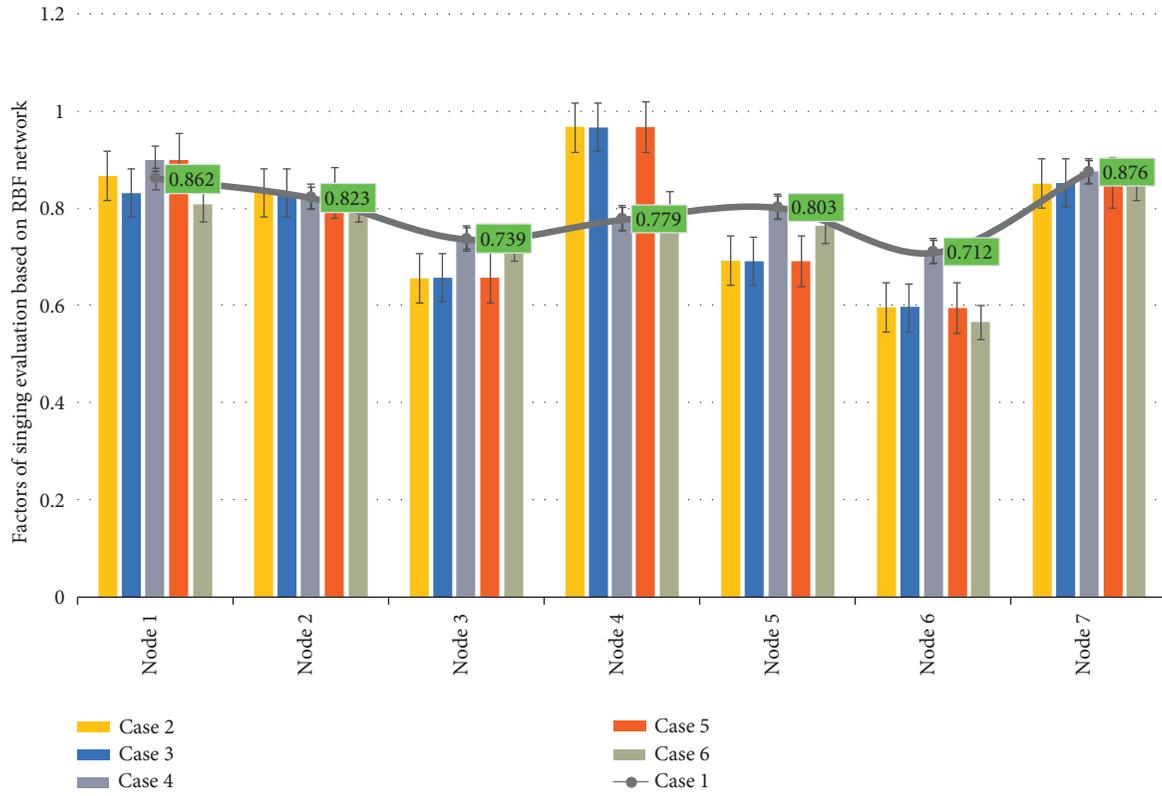


FIGURE 4: RBF network vocal music performance evaluation factor distribution.

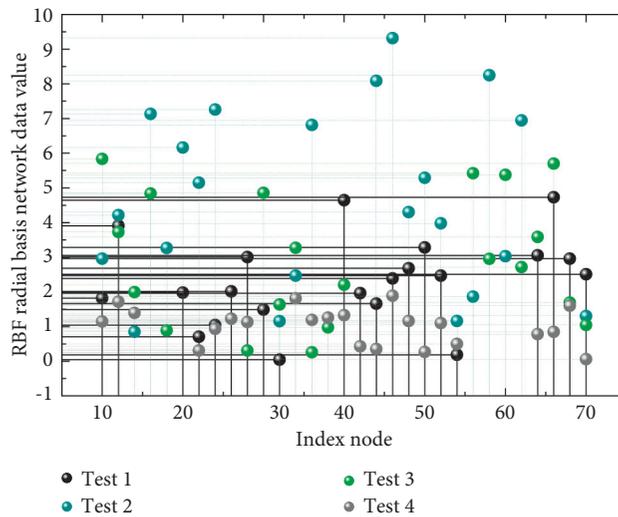


FIGURE 5: Scattered distribution of RBF network data.

conduct the test operation. Finally, according to the uniformly designed test data that has been arranged, combined with regression analysis, the validity of the model is judged by using the significance test and other indicators, and the importance of the influencing factors is analyzed.

$$\begin{cases} \sum u(P) - u(P - 1) = 0, \\ \sum [u(x_1), u(x_2), u(x_3), \dots, u(x_{n-1}), u(x_n)] = 1. \end{cases} \quad (20)$$

The design of fuzzy controller relies on practical experience. Insufficient understanding of the control process, or failure to sum up a complete experience, will affect the control effect. In addition, even if the control rules are perfect, if the control process continues to change, the control will not match the actual situation. Aiming at the above problems, a self-organizing, self-adapting, fuzzy controller can be used, which can automatically modify and adjust the control during operation, so that the system

TABLE 4: Automated evaluation process of vocal music singing.

| Algorithm steps | The selection of RBF center | The generalization of the network |
|-----------------|--|------------------------------------|
| 1 | This is mainly divided into it: | if(temp ≥ arr[i])break; |
| 2 | The significance test of $\partial y(W^t x + b)$: | arr(k) = arr[i]; |
| 3 | The regression equation $\beta(n \in t)$: | for(int i = 2*k; i ≤ len; i * = 2) |
| 4 | The significance of $\sum (a * x + b)$: | for(int i = low; i ≤ high; i++) |
| 5 | The uniform experimental design: | arr(k) = temp; |
| 6 | Method mainly includes the following steps: | while(i ≤ mid) |
| 7 | We select the test factors $L(w, b, a)$: | if((i < len)&&(arr[i] < arr[i+1])) |
| 8 | We determine the factors $\lim_{x \rightarrow \infty} \frac{1}{2} w^t w$: | for(int i = 0; i ≤ len; i++) |
| 9 | For uniform design table: | void bigsort(int * arr, int len) |
| 10 | Secondly, we select the appropriate: | int temp = arr[0]; |
| 11 | The validity of the model: | arr[0] = arr[len-i]; |
| 12 | By using $\min \phi(w)$: | arr[len-i] = temp; |
| 13 | Significance test and other indicators: | adjustdownmy(arr,0,len-i-1); |

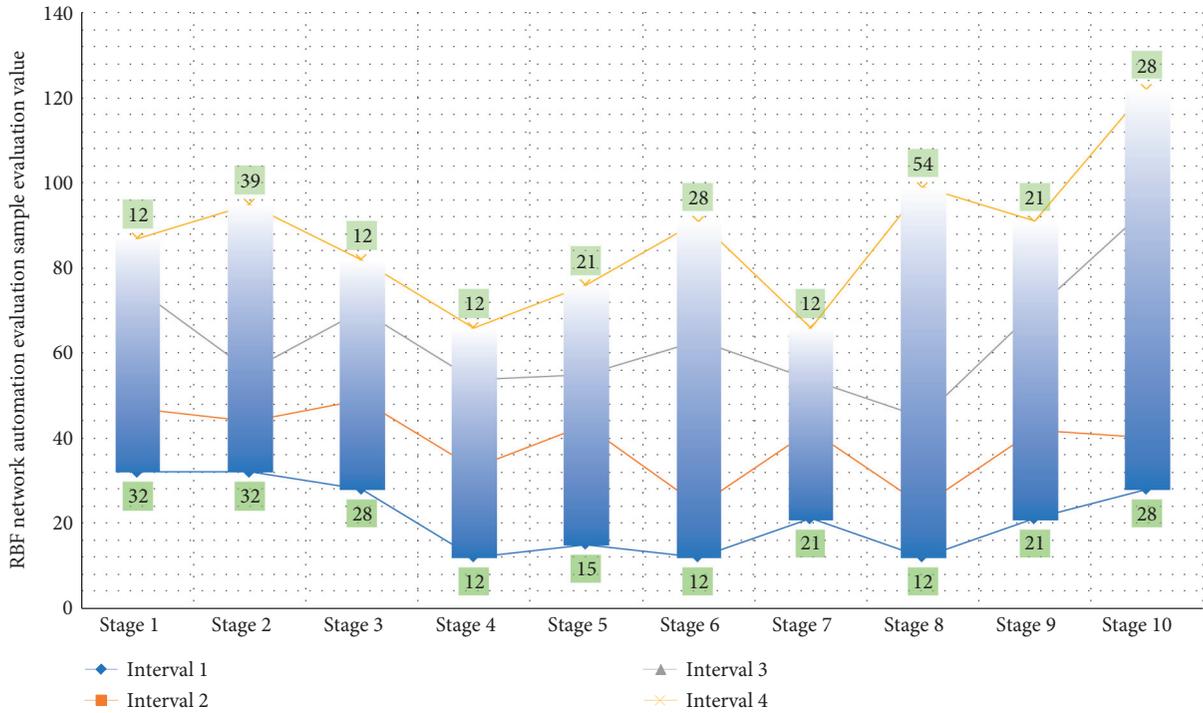


FIGURE 6: Distribution of evaluation values of RBF network automation evaluation samples.

performance can be continuously improved until the pre-determined effect is achieved.

5.3. Example Application and Analysis. There are two types of vocal score results from Moldflow analysis. One is a graphical result showing the possible shapes and sizes of the vocal score; the other is an XML file, where the coordinates of the vocal score are mostly extracted through XML. The file of the Moldflow vocal score result includes two parts: a file header and a data segment. The file header includes the stored file name, path, and other information.

$$\frac{\partial H(P)}{\partial P} = \frac{\partial [h(x_1), h(x_2), h(x_3), \dots, h(x_{n-1}), h(x_n)]}{\partial h(x_n)}. \quad (21)$$

The core part of fuzzy control is the fuzzy controller, which plays an important role in the fuzzy control system. Therefore, in the fuzzy control system, the work of designing and adjusting the fuzzy controller is very important. The design process generally has the following steps: (1) the current error and error rate of change are calculated based on the control setting and actual output; (2) the exact value of the error and error rate of change are calculated; (3) the control rule (namely, the synthesis algorithm) calculates the fuzzy control quantity; (4) the calculated fuzzy control quantity is defuzzified and added to the control object.

The evaluation values of 30 groups of experts on the new model in the database in Figure 6 are taken as the training samples of the first-level evaluation network, in which the factor set is the input samples of the first-level evaluation

network; the rest are the output samples of the first-level evaluation network and are used as the second-level evaluation.

Therefore, the wireless sensor network MAC protocol is different from the traditional MAC protocol. In order to

$$\frac{\tan(\alpha(i, j))}{\tan(\alpha(i-1, j-1))} = \frac{\cos(T_i(i-1, j-1)/T_j(i-1, j-1))\sin(T_i(i-1, j-1)/T_j(i-1, j-1))}{\cos(T_i(i, j)/T_j(i, j))\sin(T_i(i, j)/T_j(i, j))} \quad (22)$$

The experimental results show that the input and output space of vocal music prediction is divided into unequal fuzzy sets, such as high, medium, and low, and then a set of fuzzy rules are used to express the relationship between the input and output of vocal prediction. The antecedent of the relationship is the combination of different fuzzy sets of the input variables, the latter is the expression of the output variable, and the output variable is expressed as a linear combination of the input variables.

6. Conclusion

Based on RBF network multiobjective optimization and fuzzy mathematics evaluation method, this paper selects the appropriate membership function and weighting coefficient to improve the original network and establishes the automatic evaluation of multimedia vocal singing with the network comprehensive score as the output value. The RBF network realizes the evaluation of multimedia vocal music quality and finds the best evaluation combination of multimedia vocal singing automation evaluation through optimization. First, a criterion layer with 4 elements and a scheme layer with 5 prediction models with relatively strong predictive ability are established. The combined multiobjective decision-making method has the characteristics of human thinking analysis, judgment, and synthesis. Second, the combination prediction method and the RBF method are combined to predict the total value of the categorized vocal music value, the increase amount, and the corresponding growth rate, so that the inherent law of the categorized vocal music value can be predicted more comprehensively. At the same time, compared with a single indicator as a single prediction and analysis object, the stability of prediction accuracy is better. An accurate and reliable CAE simulation model is established through experiments, and a uniform test design is arranged to simulate and analyze the multimedia vocal forming process. Finally, according to the analysis results, the warping area regression equations are established respectively, and the significance test and analysis of the regression coefficients are used to obtain the degree of influence of each multimedia vocal singing automation evaluation on each defect. The simulation results show that the hierarchical form reflects the decision-making characteristics of the complex system well. The sampling points are arranged uniformly for many times, and the network factor is used as the network input to construct a

reduce energy consumption. The input samples of the network are used for the final comprehensive evaluation and the evaluation interval is set. The general evaluation system can be divided into very good, good, good, fair, poor, and very poor according to habits.

high-precision RBF network and find the optimal solution for the automated evaluation system.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by Shanghai Young Teachers Training Program.

References

- [1] L. Cai and Q. Cai, "Music creation and emotional recognition using neural network analysis," *Journal of Ambient Intelligence and Humanized Computing*, pp. 8–10, 2019.
- [2] S. Rajesh and N. J. Nalini, "Musical instrument emotion recognition using deep recurrent neural network," *Procedia Computer Science*, vol. 167, pp. 16–25, 2020.
- [3] X. Wang, "Research on the improved method of fundamental frequency extraction for music automatic recognition of piano music," *Journal of Intelligent and Fuzzy Systems*, vol. 35, no. 3, pp. 2777–2783, 2018.
- [4] J. Böhm, F. Eyben, M. Schmitt, H. Kosch, and B. Schuller, "Seeking the superstar: automatic assessment of perceived singing quality," in *Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 1560–1569, IEEE, Anchorage, AK, USA, May 2017.
- [5] I. Shahin, A. B. Nassif, and N. Hindawi, "Speaker identification in stressful talking environments based on convolutional neural network," *International Journal of Speech Technology*, vol. 24, no. 4, pp. 1055–1066, 2021.
- [6] M. A. Kızrak and B. Bolat, "A musical information retrieval system for Classical Turkish Music makams," *SIMULATION*, vol. 93, no. 9, pp. 749–757, 2017.
- [7] R. Sarkar, S. Choudhury, S. Dutta, A. Roy, and S. Kumar Saha, "Recognition of emotion in music based on deep convolutional neural network," *Multimedia Tools and Applications*, vol. 79, no. 1, pp. 765–783, 2020.
- [8] J. Zhang, Y. Zhou, and Y. Liu, "EEG-based emotion recognition using an improved radial basis function neural network," *Journal of Ambient Intelligence and Humanized Computing*, pp. 10–12, 2020.

- [9] A. Amjad, L. Khan, and H.-T. Chang, "Effect on speech emotion classification of a feature selection approach using a convolutional neural network," *PeerJ Computer Science*, vol. 7, p. e766, 2021.
- [10] G. K. Birajdar and M. D. Patil, "Speech/music classification using visual and spectral chromagram features," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 1, pp. 329–347, 2020.
- [11] N. Reljin and D. Pokrajac, "Music performers classification by using multifractal features: a case study," *Archives of Acoustics*, vol. 42, no. 2, pp. 223–233, 2017.
- [12] Y. Hu, J. Zhang, W. Jiang, and R. Sun, "Chinese pop music emotion classification based on FA-SVM," in *Proceedings of the 2018 International Conference on Control, Automation and Information Sciences (ICCAIS)*, pp. 233–237, IEEE, Hangzhou, China, Oct. 2018.
- [13] D. Y. Loni and S. Subbaraman, "Robust singer identification of Indian playback singers," *EURASIP Journal on Audio Speech and Music Processing*, vol. 9, no. 1, pp. 12–14, 2019.
- [14] Y. R. Pandeya, B. Bhattarai, and J. Lee, "Deep-learning-based multimodal emotion classification for music videos," *Sensors*, vol. 21, no. 14, p. 4927, 2021.
- [15] N. A. Hindawi, I. Shahin, and A. B. Nassif, "Speaker identification for disguised voices based on modified SVM classifier," in *Proceedings of the Multi-Conference on Systems, Signals & Devices (SSD)*, pp. 687–691, IEEE, Monastir, Tunisia, March 2021.
- [16] M. Lataifeh, A. Elnagar, I. Shahin, and A. B. Nassif, "Arabic audio clips: identification and discrimination of authentic Cantillations from imitations," *Neurocomputing*, vol. 418, pp. 162–177, 2020.
- [17] N. Priyadarshani, S. Marsland, and I. Castro, "Automated birdsong recognition in complex acoustic environments: a review," *Journal of Avian Biology*, vol. 49, no. 5, pp. jav-01447, 2018.
- [18] A. Rosner and B. Kostek, "Automatic music genre classification based on musical instrument track separation," *Journal of Intelligent Information Systems*, vol. 50, no. 2, pp. 363–384, 2018.
- [19] Z.-H. Michalopoulou, P. Gerstoft, B. Kostek, and M. A. Roch, "Introduction to the special issue on machine learning in acoustics," *Journal of the Acoustical Society of America*, vol. 150, no. 4, pp. 3204–3210, 2021.
- [20] M. Sajjad and S. Kwon, "Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM," *IEEE Access*, vol. 8, pp. 79861–79875, 2020.
- [21] M. Premoli, D. Baggi, M. Bianchetti et al., "Automatic classification of mice vocalizations using machine learning techniques and convolutional neural networks," *PloS one*, vol. 16, no. 1, p. e0244636, 2021.
- [22] S. Rajib, C. Sombuddha, D. Saikat, and A. Roy, "Recognition of emotion in music based on deep convolutional neural network," *Multimedia Tools and Applications*, vol. 79, no. 1-2, pp. 765–783, 2020.
- [23] F. Daneshfar, S. J. Kabudian, and A. Neekabadi, "Speech emotion recognition using hybrid spectral-prosodic features of speech signal/glottal waveform, metaheuristic-based dimensionality reduction, and Gaussian elliptical basis function network classifier," *Applied Acoustics*, vol. 166, p. 107360, 2020.
- [24] J. Bai, K. Luo, J. Peng et al., "Music emotions recognition by machine learning with cognitive classification methodologies," *International Journal of Cognitive Informatics and Natural Intelligence*, vol. 11, no. 4, pp. 80–92, 2017.
- [25] H. Mukherjee, A. Dhar, S. M. Obaidullah, K. C. Santosh, S. Phadikar, and K. Roy, "Segregating musical chords for automatic music transcription: a LSTM-RNN approach," *Lecture Notes in Computer Science*, vol. 11942, pp. 427–435, 2019.
- [26] A. Christy, S. Vaithyasubramanian, A. Jesudoss, and M. D. A. Praveena, "Multimodal speech emotion recognition and classification using convolutional neural network techniques," *International Journal of Speech Technology*, vol. 23, no. 2, pp. 381–388, 2020.
- [27] G. Muhammad, M. Masud, S. U. Amin, R. Alrobaea, and M. F. Alhamid, "Automatic seizure detection in a mobile multimedia framework," *IEEE Access*, vol. 6, pp. 45372–45383, 2018.