Evaluation of Vocal Music Performance Teaching: A Fuzzy Logic-Based Approach in the Context of Embedded Systems

Qiang Zeng[1,2*]

[1]School of Education, Geely University of China, Chengdu 641423, China
[2]Geely College Brain Science Application and Child Development Research Center, Chengdu 641423, China, zengqiang188@outlook.com

Corresponding Author: Qiang Zeng, zengqiang188@outlook.com

Abstract. Vocal music teaching and learning performance and quality analysis played an important role in music education. The multiple musical theory and intelligence approaches are widely applied to analyze lesson recommendations, design requirements, and operation forms. The intelligence theory utilizes questionnaires to explore student performance and comprehension scores to improve their learning process. However, the existing techniques fail to address the music quality evaluation-related quantitative problems. The research issues maximize the difficulties while analyzing the student learning efficiency. The challenges are overcome by applying the Vocal Music Teaching Method (VMTM) with Fuzzy logic approach. During the analysis, fuzzy rules are generated according to different factors such as expressions, rhythm, and pitch accuracy. The created rules evaluate musical performance and manage the system's flexibility. The vocal performance is explored by collecting data from student feedback. The feedback is continuously examined to improve the overall Vocal Music Teaching Method (VMTM) efficiency in music class. Finally, the experimental results and discussions evaluate the system's efficiency.

Keywords: Music Education, vocal music teaching, quantitative problem, fuzzy logic, fuzzy rules, and student feedback.
DOI: https://doi.org/10.14733/cadaps.2024.S8.117-132

1 INTRODUCTION

In music education [4],[21], teaching should be met the set of principles and objectives by performing continuous training, practice, and information integrity. The teaching concepts are highly applied in the music class to improve the student music learning efficiency [7]. During the teaching process, educators integrate the traditional learning process with the current learning procedure to improve teaching performance in an arts-driven and creative manner. The well-practiced musical techniques must have a few similarities, such as sequential and systematic design, integrity, and authenticity of music (folk music), including the mother-tongue techniques to derive the child's perspective information like timbre, pitch, and rhythm [11]. The gathered information is analyzed using traditional music techniques to identify the student musical expressions. According to the expertise and expressions, music teachers continuously observe the student performance depending on the subjective evaluations [23],[2]. However, the traditional approach faces a quantitative problem while analyzing the numerical data [3]. The quantitative problem occurs during the student performance assessment, teaching method evaluation, music education policy analysis, practice, and curricula. Hence, music education programs require continuous improvement and development processes to maximize student music evaluation efficiency. The statistical research difficulties are resolved by introducing new techniques and computational methods to improve the musical performance evaluation process precisely and objectively [15]. Traditional approaches such as neural models, support vector machines, and K-nearest neighboring approaches [25],[13] are utilized to evaluate musical teaching performance. Among the various computational approaches, fuzzy logic [12] approaches have recently been widely applied to analyzing the various imprecision data. The existing technologies fail to solve the quantitative problems and difficulties in predicting the various notes in musical learning [22]. Therefore, the fuzzy logic approach is applied to resolve the uncertainty and data analysis complexity. Fuzzy logic is an effective mathematical framework investigating inaccurate data with maximum accuracy.

The fuzzy logic approach is effectively utilized in music education [20],[6]. The fuzzy rules maintain the system flexibility while evaluating the musical performance by considering the expressions, rhythm, and pitch accuracy [9]. The fuzzy rules are generated according to the expert knowledge that maximizes the student performance analysis process. During the analysis, the vocal music teaching process was defined with the help of the linguistic variables [8] that consist of tempo, rhythm, pitch, and expressions. These linguistic variables are categorized into fuzzy sets related to the efficiency of each variable in music performance. Considered, the pitch variables are decomposed into various fuzzy sets like acceptable, poor, bad, good, and perfect [19]. The fuzzy sets are created according to the various levels of pitch accuracy. The created fuzzy sets are considered while forming the fuzzy rules because it used to evaluate the system's performance. The rules are developed with the music teachers' help because they play a major role while evaluating their teaching performance. The formed rules evaluate the pitch accuracy that determines the student's performance, whether they belong to good, bad, or perfect performance categories. The student's performance is measured in terms of dynamics, timings, accuracy, and expressions [1]. According to the student's performance, the score value is allocated to the student to improve the student's learning efficiency. Effectively utilizing fuzzy rules and sets is the main reason for choosing this method in this music performance evaluation process.The utilization of fuzzy rules and sets, along with embedded systems, is the key reason for choosing this method for music performance evaluation. Then the primary objective of the study is listed as follows.

- To improve the music student performance analysis efficiency by considering the different factors such as tempo, rhythm, expressions, and pitch accuracy.
- To resolve the quantitative and prediction error problems while analyzing student performance.
- To design the VMTM with Fuzzy logic based music system for resolving the uncertainty and complexity in imprecise data analysis.

Then the remaining paper structure is organized as follows: Section 2 analyzes the various researcher's opinions regarding the student music performance analysis system. Section 3 describes the fuzzy logic-based student music performance evaluation process and the excellency of the system described in Section 4. The conclusion is explained in section 5.

2 RELATED WORKS

This section discusses the various researcher's opinions on evaluating the student's music performance in music education.

Authors	*The intention of the work*	*Methods utilized in this work*	*Advantages*	*Disadvantages*	*Findings*
Wang, Xiaolu, and Yao Chen et al., 2021 [18]	*Developed the FPGA and neural network-based music teaching environment*	*FPGA-based hardware accelerator along with neural feedback model*	*High processing speed, minimizes power consumption and latency*	*Limited for music components and difficult to operate by several users*	*Predicting the student music performance with minimum latency.*
Chen Xu et al., 2021 [19]	*Analyzed and assessing the music teaching ability*	*Compensated Fuzzy Neural Network (CFNN)*	*Maximum accuracy, ability to handle the uncertain data*	*Requires the huge volume of data for training and complexity to understand the system process*	*Accuracy-89.5%*
Wang Xiaolu, Qun Wang, and Yao Chen 2020 [20]	*To explore the online music training curriculum to improve the music education*	*Backpropagation Neural Model (BNM)*	*Maximum accuracy, ability to handle large data, and optimize the course schedule.*	*The system fails to manage the scalability and flexibility*	*Promising accuracy*
Yang Qing et al., 2021 [21]	*Created the musical teaching database in the cloud*	*Neural Network with open Cloud design*	*Maintain scalability, flexibility, and maximum accuracy*	*Need technical knowledge to implement*	*Maximum accuracy*
Wang Weiqing, Jin Pan, Hua Yi, Zhanmei Song, and Ming Li 2021 [22]	*Developed the musical performance evaluation system for analyzing the audio-based piano learners' efficiency*	*Convolution Neural Model with Attention Mechanism (CNM-AM)*	*Providing personalized feedback for learners and being able to analyze the huge volume of data*	*Requires the optimization model to reduce the deviation between the computed outputs*	*Accuracy - 85.1%*

Zhang Xiaoxing 2021 [23]	*To create the system for analyzing preschool student visualization and signing performance*	*Music-related information is collected from 72 students via surveys and questionnaires. The gathered information is processed with the help of multi-sensory methods.*	*Maximize the learning experience and increases creativity & imagination*	*Lack of control because of the huge variables and restricted generalizability due to the small data.*	*Promising results*
Liu, Xi, and Chengyu Yang 2021 [24]	*Applied cloud systems and embedding platforms to analyze the student music performance in the remote music classroom.*	*Neural Networks approach applied on the cloud systems to identify the student learning performance*	*Convenient and provides the low-cost based remote classroom learning*	*Limitations in instrument quality and capability*	*The overall satisfaction score is 4.4. out of 5.*
Seshadri, Pavan, and Alexander Lerch 2021 [25]	*To developed the music performance assessment system.*	*Contrastive Learning With The Neural Model (CL-NM) is applied to assessing the student's musical performance*	*Able to handle the tempo and dynamic factors related to variabilities in music training*	*Limitation of the diversity of dataset information*	*Focusing on various music characteristics such as pitch, accuracy, and other factors leads to maximizing the prediction accuracy*

Various researchers evaluate music performance using machine learning, neural model, contrastive learning, and other prediction approaches. The existing methods can analyze the large volume of musical information successfully. However, they fail to manage the quantitative and prediction error rate problem. These problems are addressed with the help of the fuzzy logic approach that solves uncertainty and imprecise data analysis difficulties. The detailed working process of the fuzzy logic related to Vocal Music Teaching Method-based music performance analysis is discussed below.

3 MUSIC PERFORMANCE ANALYSIS USING THE VOCAL MUSIC TEACHING METHOD (VMTM)

The main intention of this work is to measure and evaluate student music performance using the Vocal Music Teaching Method (VMTM). The objective is achieved by deriving the most relevant information from the student's audio recordings. During the analysis, fuzzy logic is incorporated to evaluate the student efficiency according to different parameters such as volume, time, pitch accuracy, tempo, and rhythm. The derived music features are analyzed with the help of the fuzzy rules used to evaluate student learning and performance efficiency. The VMTM approach-based music performance analysis consists of several steps: data collection, feature extraction, fuzzy logic-based evaluation, feedback, and improvement. The detailed working process of the VMTM-based music performance analysis process is illustrated in Figure 1.

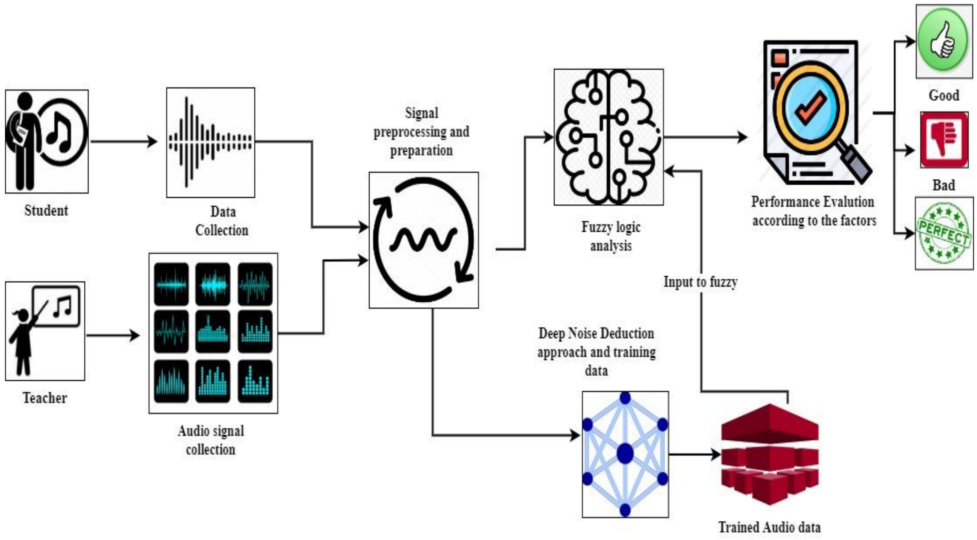


Figure 1: Overall Working Process of VMTM.

Figure 1 illustrated that the working process of vocal music teaching model based student learning performance analysis. The system uses the different audio information like vocal range, energy factor, spectral analysis, tempo and volume parameters. These parameters are evaluated using the fuzzy logic system that recognize the performance level such as good, bad and perfect. From the predicted output value, the feedbacks are created which is transferred to the teachers to improve the overall music learning process. The system uses the Free Music Archive (FMA) [8] dataset information to evaluate the VMTM efficiency. The dataset was utilized to perform the musical information retrieval. The dataset has 106574 audio tracks that are collected over 343 days. During the data collection, 14854 albums and 16341 artists are involved. In addition, 161 genres of hierarchical taxonomies are utilized. The collected information is processed by a feature extraction step that covers the preprocessing and post-processing of the data.

3.1 Musical Data Feature Extraction

The feature extraction process was crucial in assessing the student's music performance. The collected audio information is frequently analyzed to derive the relevant details that help to examine the performance. Initially, the audio information is integrated with the help of the preprocessing technique to eliminate irrelevant artifacts and noise information. The noise removal process removes the noise information and normalizes the signal, enhancing the signal quality. This work utilizes the Deep Noise Destruction (DND) approach to maximize the signal quality. The DND approach uses the deep learning concept while eliminating the noise in the audio. The detailed working process of musical data feature extraction is illustrated in Figure 2.

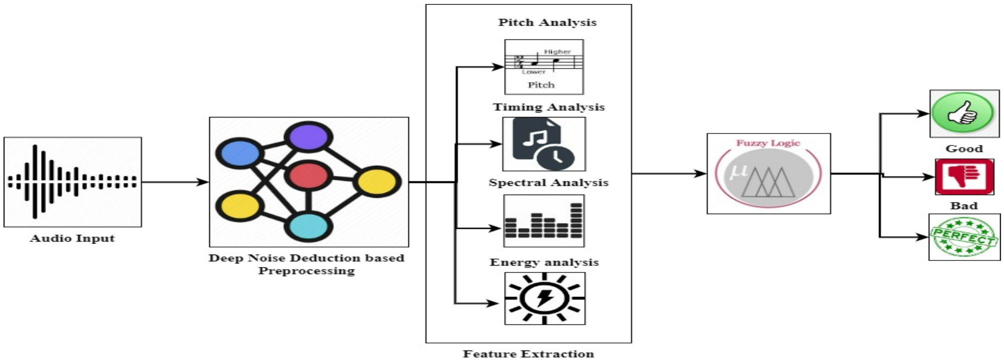


Figure 2: VMTM Feature Extraction Analysis.

The data has been collected in any environment; the noise can be occurred because of the data collection environment. Therefore the collected audio signals are arranged in the active phase to eliminate irrelevant information. The DND has three layers, input, hidden, and output, that continuously process the inputs to eliminate irrelevant information. The student musical audio signal is ay, and this work aims to get the clean signal cx. Therefore, the DND approach has been framed as using equation (1)

$$\mathrm{ay} = \mathrm{ax} + \mathrm{n} \tag{1}$$

In equation (1), n is denoted as the noise in the audio signal. The DND approach has multiple layers that learn every feature in the input, and the noise information is eliminated with the help of equation (2).

$$\left.\begin{aligned} \mathrm{n_{hat}} &= \mathrm{f(ay)} \\ \mathrm{x_{hat}} &= \mathrm{ay} - \mathrm{n_{hat}} \end{aligned}\right\} \tag{2}$$

In equation (2), DND utilized activation function is denoted as the f. After eliminating the noise information, it has been passed to the high pass filter for removing the low-frequency noise impacts. During this process, Perceptual linear prediction (PLP) coefficients are utilized. The pre-emphasis process of PLP is defined using equation (3).

$$\mathrm{ay[n]} = \mathrm{ax[n]} - \alpha * \mathrm{ax[n-1]} \tag{3}$$

In equation (3), the input audio signal is denoted as $ax[n]$, pre-emphasis filter output is denoted as $ay[n]$ and the emphasis coefficient component is denoted as α. After performing the emphasis, the frame has to be segmented from the original input. The signal is split into frames 20 to 30 ms with 50% overlap between the frames. Then the frame segmenting process is defined using equation (4)

$$\mathrm{ax_m[n]} = \mathrm{ax[n + m * H]} \tag{4}$$

According to equation (4), the input audio signal $ax[n]$ is divided into the mth frame with 50% of hop size (H) for the mth frame index. After that, the windowing function is applied to every audio frame to minimize spectral leakage. Here, hamming window function is applied to reduce the frame truncation-based spectral leakage.The hamming window function is defined using equation (5)

$$w[n] = 0.54 - 0.46 * \cos\left(\frac{2*\pi*n}{N}\right) \tag{5}$$

In equation (5), the window function is defined as $w[n]$ that is applied on window length N for n sample index. The frames are further explored with the help of the fourier transform to get the frequency spectrum computed using equation (6).

$$\mathrm{X_m[k]} = \sum \mathrm{x_m[n]} * \mathrm{e}^{\left(\frac{-\mathrm{j}*2*\pi*\mathrm{k}*\mathrm{n}}{\mathrm{N}}\right)} \tag{6}$$

In equation (6), the mth frame spectrum frequency is represented as $X_m[k]$, the windowed frame is defined as $x_m[n]$. The $X_m[k]$ is computed for N frame length, j imaginary unit, and k frequency index. Then the extracted frequency spectrum is processed with the help of mel filters that identifies the frequency representation via the mel scale. This process compresses the frequency representation and the mathematical illustration in equation (7).

$$H_m[k] = \min\left(\frac{(X_m[k]-F_{m-1})}{(F_m-F_{m-1})}, \frac{(F_{m-1}-X_m[k])}{(F_{m+1}-F_m)}\right) \tag{7}$$

In equation (7), mth Mel filter output is defined as $H_m[k]$, mth frame frequency spectrum defined as $X_m[k]$, mel filter center frequency is F_m, F_{m+1} and F_{m-1} is defined as the adjacent mel filters center frequency. The computed mel frequency output (L_m) is further compressed with the help of the logarithmic function, which is computed using equation (8)

$$L_m = log\left(\sum H_m[k]^2\right) \tag{8}$$

At last, the PLP coefficients are obtained from the mel-filterbank that is obtained with the help of equation (9)

$$P_i = \sum l_m * \cos\left(\frac{\pi*i*(m-0.5)}{M}\right) \tag{9}$$

In equation (9), the PLP coefficient is represented as P_i, log mel filterbank coefficient is denoted asl_m with m index and filter, counts are represented as M. The derived mel frequency components and PLP coefficients are trained with the help of the deep neural model. The neural Network uses the sigmoid function as an activation function to identify and extract the features from the audio signal. The extracted features are applied to the neural model, and the training process is initiated to minimize the deviations. Here, the neural model reduces the difference between the ground truth audio and predicted clean audio. According to the above features, noise is continuously examined, and irrelevant information is removed from the original audio input. The discussed process successfully removes the reverberation, microphone, and background noise.

3.2 Pitch Detection

The pitch factor played an important role in analyzing the vocal performance. The pitch characteristics are denoted as the fundamental sound frequency, which helps to determine the accuracy of the music performed by hitting the right notes. The vocal performance pitch regarding cepstral analysis and autocorrelation has been analyzed. The autocorrelation examines the relationship between the delayed and signal version. For signal $ax(n)$, the autocorrelation function $R(k)$ is estimated using equation (10).

$$R(k) = \sum ax(n) * ax(n-k) \tag{10}$$

In equation (10), k is the lag value computed by taking the distance between the timing signals. From the computed autocorrelation value, the signal obtained a high correlation value at a certain period is considered the pitch period. In addition, cepstral analysis is performed on computed spectrum value to identify the pitch value. The cepstral analysis uses the fourier transform to compute the signal power spectrum value defined in equation (11).

$$\mathrm{X(f) = |F(x(n))|^2} \tag{11}$$

In equation (11), the fourier transform is denoted as F, and the magnitude of the signal is represented as |.|. From the computed cepstral analysis, cepstrum is estimated as $cepstrum = F^{(-1)\,[\log(X(f))]}$. The computed cepstrum values belong to the signal pitch value. The selected autocorrelation and cepstrum values are computed continuously to identify the pitch value of the student input audio signal.

3.3 Timing Analysis

The next important factor is timing analysis used to evaluate the student's vocal performance efficiency. The timing-related information is extracted with the help of the tempo estimation process. This factor is used to measure the performance timing that helps to identify the beats in their performance. The signal time value is obtained by taking the difference between successive peaks presented in the onset function of audio signals. The onset function derives the audio signal activities at every frame. The tempo value is estimated from the computed autocorrelation value and lag value. The tempo value is computed using equation (12)

$$\mathrm{Tempo} = \frac{60}{(\mathrm{lag} * \Delta)} \tag{12}$$

In equation (12), timing signal successive frames related time is denoted as Δ (constant value). After identifying the audio signal timing value, spectral information must be analyzed.

3.4 Spectral Analysis

The preprocessed audio signal is decomposed into sub-signals according to the constituent frequency. This process is done by applying a discrete fourier transform to get the signal spectral information. The extracted spectral information explores the music tonality and timbre details. Considered the signal $ax(n)$ with length N and the discrete fourier transform is defined using equation (13).

$$X(k) = \sum \left[x(n) * \exp\frac{(-j*2\pi nk)}{N}\right] \tag{13}$$

In equation (13), the frequency bin index is defined as k, ranges from 0 to N-1, and the imaginary unit is denoted as j. The framed fourier transform analyzes the association between the frequency components. During the analysis, the window function is utilized to derive the spectral content that minimizes the spectral leakages, maximizing the spectral information prediction and resolution. Here, hamming window function is utilized to get the spectral information that is computed using equation (14)

$$W(n) = 0.54 - 0.46 * \cos\left(\frac{2\pi n}{N-1}\right) \tag{14}$$

In equation (14), N is defined as window length, n is the sample index, and the windowed signal is defined using equation (15).

$$xw(n) = x(n) * w(n) \tag{15}$$

The spectral information is obtained from the windowed signal, and a spectrogram is drawn in which time is defined in the x-axis and frequency is denoted in the y-axis. In the spectrogram, the color denoted the spectral content magnitude. The obtained spectral analysis-related frequency content was used to evaluate the vocal music performance.

3.5 Energy Analysis

The next important feature is energy analysis, in which audio signal amplitude is measured. The extracted energy feature relates more to the audio dynamic and volume-related information. The energy factor is computed with the help of the Root Mean Square Amplitude (RMSA), which is estimated using equation (16).

$$\mathbf{RMSA} = \sqrt{\frac{1}{\mathbf{N}}\sum(\mathbf{ax}(\mathbf{n})^2)} \tag{16}$$

In equation (16), the total number of samples is denoted as N for input signal $ax(n)$. The RMSA value is computed for every frame in the input signal, and the signal is segmented from 10 to 50 ms duration. From the computation, the input signal volume is high if the signal has a high RMSA value. In addition, the envelope detection approach is applied to derive the input audio signal amplitude value that identifies the attack and decay of every note. The extracted audio signal features are investigated with the help of the fuzzy logic approach to minimize the uncertainty and imprecise data analysis difficulties. The detailed working process is described in the below section.

3.6 Fuzzy Logic-based Music Performance Analysis

The extracted musical features are applied to the fuzzy logic to analyze the student's musical performance. The analysis system uses fuzzy rules and principles to identify the actual performance. The derived features such as energy, spectral, timing, pitch, and volume details are given as input to the fuzzy systems. The features are analyzed using fuzzy rules to evaluate student performance. The fuzzy system has three components: input, output, and set of rules. The extracted audio features are treated as input, and the score value is the output variable. During the analysis, a set of rules is utilized to map the input variable to output variables according to the linguistic rules, which are declared with the help of fuzzy logic. The fuzzy system uses the membership function with values between 0 and 1. According to the degree of membership value, the input variables are mapped to the output variables. Let's assume "pitch efficiency "is the input variable that consumes the value from 0 and 1. If the system attains zero as the output value, it denotes poor pitch efficiency, and one means perfect pitch. Therefore, the fuzzy rules are generated with the help of expert knowledge that maximizes the overall vocal performance analysis efficiency. According to the discussion, the fuzzy rule is "if pitch efficiency is high and good timing, then the student score is excellent. Therefore, the fuzzy rules are generated by combining the fuzzy union and intersection operators. The generated fuzzy rules and sets predict the overall efficiency score; hence, defuzzification is performed to get the single numerical score. Then the fuzzy sample rules are listed as follows:

1) IF pitch efficiency is minimum, then the overall performance score is very poor.
2) IF pitch efficiency is medium, good timing and the overall performance score is fair.
3) IF pitch efficiency is maximum, good timing and the overall performance score is good.

4) IF pitch efficiency is high, excellent timing and the overall performance score are excellent.

5) Then the overall working VMTM with Fuzzy logic systems working steps are defined in table 1.

Step 1: collect the audio signals from the students
Step 2: removing noise from the audio input by using the deep noise deduction approach
Step 3: analyzing the audio frame, windowing and segment the irrelevant portion
Step 4: Extracting the pitch, volume, tempo, energy and spectral related features.
Step 5: analyzing the features and create the fuzzy rules according to the membership value.
Step 6: Explore the quality and performance of the music from the audio input.
Step 7: categorize the input into good, bad and perfect.

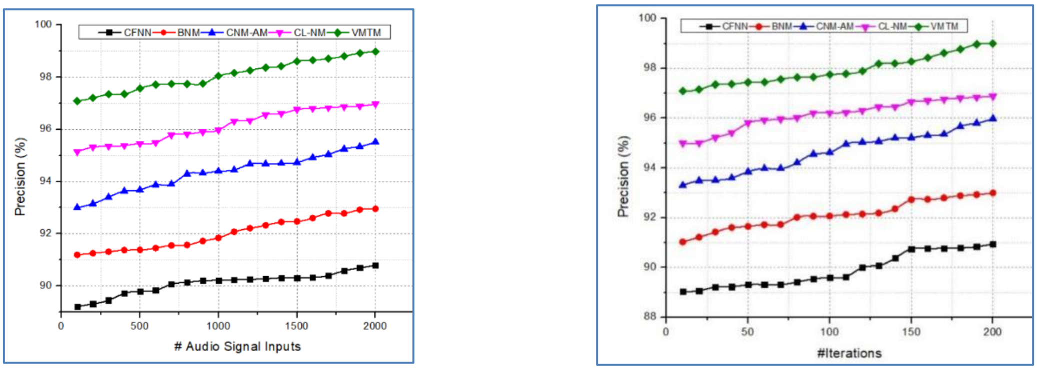
Table 1: Algorithm Steps for VMTM with Fuzzy Logic-Based Music Performance Analysis.

Thus, the fuzzy sets of pitch efficiency, timing, and performance score could be defined using triangular or trapezoidal membership functions. Fuzzy intersection and union would merge the fuzzy input sets according to the rules. The fuzzy rules-based generated output is given as feedback to the teacher and student. The feedback identifies the student's and teachers' strengths and weaknesses. Students' efficiency is determined using dynamics, volume, timing, and pitch efficiency. The generated output value-related performance score is obtained and represented in the graphic format to improve the overall data analysis. The teacher can create a student-specific training plan using feedback. Students can improve their musical performance with a customized training regimen. The teacher can provide pitch-training exercises if the learner needs to increase pitch accuracy. The teacher can also provide rhythm and timing activities for the students. The fuzzy logic vocal music education method requires feedback and improvement. It allows students to obtain constructive criticism and assistance for growth and teachers to create a personalized training plan to help students reach their maximum potential. The teacher can create a student-specific training plan using feedback. Students can improve their musical performance with a customized training regimen. The teacher can provide pitch-training exercises if the learner needs to increase pitch accuracy. The teacher can create a student-specific training plan using feedback. Students can improve their musical performance with a customized training regimen. The teacher can provide pitch-training exercises if the learner needs to increase pitch accuracy. The teacher can also provide rhythm and timing activities for the students. The fuzzy logic vocal music education method requires feedback and improvement. It allows students to obtain constructive criticism and assistance for growth and teachers to create a personalized training plan to help students reach their maximum potential. Then the system's efficiency is evaluated using the experimental study, which is discussed in section 4.

4 RESULTS AND DISCUSSIONS

This section discusses the Vocal Music Teaching Method (VMTM) efficiency and how the system improves the perception of music education. The VMTM system has several phases audio signal collection, noise removal, feature derivation, and fuzzy logic-based performance analysis. These steps use the Free Music Archive (FMA) dataset information to evaluate the system excellently. The collected audio signal is processed according to the above discussions and different factors such as timing, pitch, volume, and spectral features. These features are investigated with the help of the fuzzy logic system that identifies the student performance over the period. The fuzzy system uses audio features and a set of fuzzy rules to improve user-friendliness and usability for teachers and students. In addition, predicted score values are given as feedback to the teachers and students to enhance their overall learning and performance. The discussed VMTM approach is evaluated using different metrics such as accuracy, error rate, precision, recall, and Matthew Correlation Coefficient (MCC). These metrics are used to analyze how accurately the VMTM approach identifies student performance without compromising sound quality.

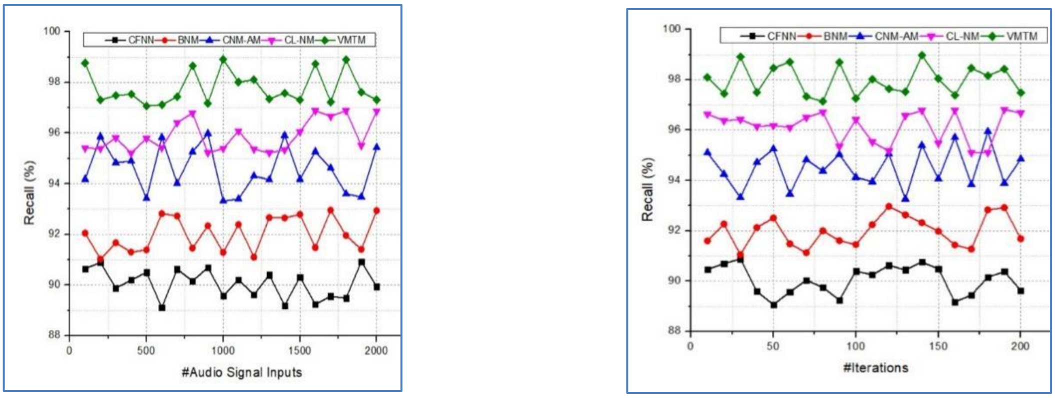
Along with this, task completion time, user satisfaction, and error counts are related details evaluated via user testing or survey. This process helps to maintain user-friendliness and usability in music education. The discussed VMTM approach is evaluated and compared with different methods, such as Compensated Fuzzy Neural Network (CFNN) [20], Backpropagation Neural Model (BNM) [21], Convolution Neural Model with Attention Mechanism (CNM-AM) [23] and Contrastive Learning with The Neural Model (CL-NM) [26]. For the comparison, [20, 21, 23, and 26] references are utilized because it effectively utilizes the neural model working process to reduce the deviation between the computed output values. Then these methods provide promising results and manage the system flexibility and scalability factors.



(a) Inputs (b) Iterations

Figure 3: Precision Analysis of VMTM.

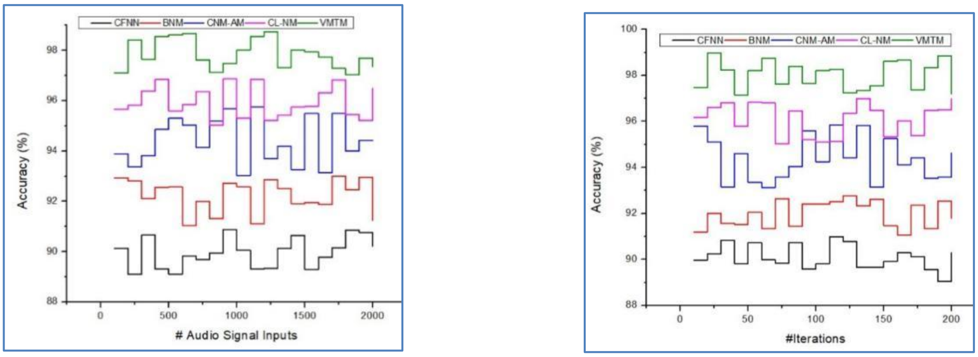
Figure 3 illustrates the precision analysis of the Vocal Music Teaching Method (VMTM) approach. The VMTM efficiency is compared with the existing methods discussed in Section 2. The VMTM approach uses fuzzy logic and fuzzy rules to identify and evaluate students' musical performances. The fuzzy logic system uses pitch information, vocal status, volume, and timing; these factors help to improve the overall performance analysis efficiency. The system uses the PLP (equation 3) coefficient, tempo (equation 12), and energy (equation 16) to form the fuzzy rules that maximize the total system efficiency. The accountability of multiple audio factors enhances the performance quality. The rules are generated depending on pitch and timing, reframing the feedback and improving the music performance. In summary, effectively utilizing and combining various factors and fuzzy sets maximize the precision values. The high precision value indicates that the VMTM approach effectively assesses student performance.



(a) Inputs (b) Iterations

Figure 4: Recall Analysis.

Figure 4 illustrates, the VMTM approach-based music performance analysis system efficiency is evaluated using the recall metrics. The metric used to predict how effectively the VMTM approach identifies the overall instances of specific classes such as (good, bad, and perfect). From the analysis, the VMTM approach attains the maximum recall value compared to the other prediction models defined in Section 2. Here, the VMTM approach uses factors such as energy, spectra, timing, tempo, and pitch to predict the student's vocal performance. Combining these features explores the bad and good vocal performance with high recognition accuracy. Considered, the pitch features are computed according to the vocal performance-related fundamental frequency value. Therefore, the pitch feature is used to predict whether the singer sings the right note. The timing feature is identified depending on the tempo or music beat value. The tempo helps predict whether the singer sings the song slowly or fast. The energy value is analyzed based on the sound dynamic and volume that measures the singer singing the song loud or soft. Therefore the above features are student performance is analyzed successfully (bad and good performance) with maximum recall value. Thus, VMTM obtains high recall values and allows vocal performance development.

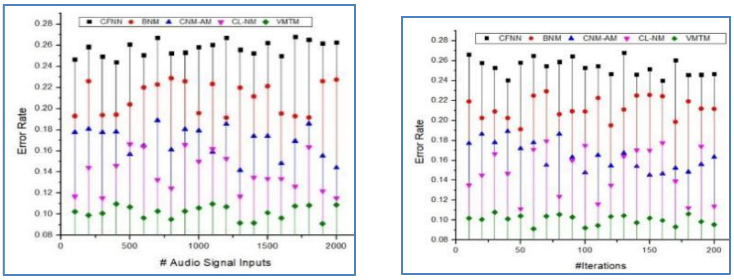


(a) Inputs (b) Iterations

Figure 5: Accuracy Analysis of VMTM.

Figure 5 illustrates the accuracy analysis of the VMTM bases student vocal performance analysis system. The above figure 5 clearly shows that the VMTM approach attains the maximum accuracy

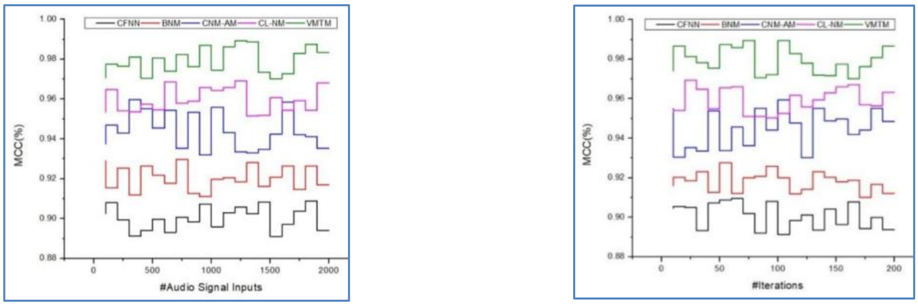
(97.91% for inputs and 97.93% for iterations) compared to the other methods. The VMTM approach analyzes student performance by collating the actual score to the predicted value. In addition, fuzzy logic systems are utilized to make the quantitative assessment of student performance. The quantitative assessment process is performed to the teacher's subjective assessment. The VMTM approach considered various factors such as autocorrelation between the notes (pitch), tempo, timing, spectral and energy features. These features maximize the overall quantitative assessment procedure of student performance. During the analysis, the fuzzy system uses the expert's knowledge to generate the fuzzy rules that maintain the system's adaptability, flexibility, and scalability. In addition, the feedback systems are enabled with the help of the performance score that personalizes the training process and improves their overall learning and performance in the future.



(a) Inputs (b) Iterations

Figure 6: Error Rate Analysis of VMTM.

Figure 6 shows the error rate analysis of the VMTM approach, which is computed by taking the ratio between the number of incorrect evaluations and the total number of evaluations. The computation gives the minimum error value, which shows that the VMTM approach attains maximum accuracy while analyzing the student's musical performance. The system ensures the minimum error rate because the VMTM approach can derive multiple features such as pitch, tempo, timing, and energy features. These features-based generated fuzzy rules are more adaptable to the environment and analyze the student performance with a maximum recognition rate. In addition, the deep noise deduction process eliminates the irrelevant information from the audio signal, reducing the deviation between the computed outputs. More ever, every pitch-related correlation value is frequently examined to improve the overall system efficiency.



(a)Inputs (b) Iterations

Figure 7: Matthew Correlation Coefficient (MCC) Analysis of VMTM.

The MCC can be utilized to assess how well the fuzzy logic-based assessment system assessed the student's vocal performance. If the MCC is high, the system can reliably determine whether the performance is good or terrible. From Figure 7, the VMTM method's MCC performance can be evaluated against other prediction models in which VMTM attains the maximum MCC value (0.97% for inputs and 0.975% for iterations). According to the results and analysis, the VMTM approach recognizes the student performance up to 4.21% of various inputs and 4.83% for iterations. The VMTM approach uses the deep noise deduction model of different layers that minimize the deviations between the outputs up to 48.97% for various inputs and iterations. Thus the Vocal Music Training Model (VMTM) successfully uses the deep noise deduction approach and fuzzy logic systems to explore the student vocal performance with maximum accuracy and minimum error rate compared to the other prediction models.

5 CONCLUSION

Thus the paper describes the Vocal Music Training Model (VMTM) using fuzzy logic-based student vocal performance analysis. Initially, the audio inputs are collected from the Free Music Archive (FMA), which is continuously processed by different phases of the preprocessing approach. The method analyzes the audio signal using frame fragmentation, windowing, decomposition, mel-filter, and PLP components. These steps normalize the audio signal to improve the overall performance analysis. Further, the VMTM process examines the input signal to derive the pitch, tempo, timing, spectral, and energy features because it maximizes the vocal performance. According to these features, fuzzy logic is created to determine the good, bad, and perfect performance with maximum accuracy (97.91% for inputs and 97.93% for iterations) and minimum error rate (0.097 for inputs and 0.100 for iterations). In the future, optimized techniques will be incorporated with this process to fine-tune the factors to minimize the computation time.

Qiang Zeng, https://orcid.org/0009-0000-4890-1252

REFERENCES

[1] Akkermans, J.; Renee, S.; Daniel, M.; Kelly, J.; Daniel, S.; David, B.; Veronika, B.: et al. Decoding Emotions in Expressive Music Performances: A Multi-Lab Replication and Extension Study, Cognition and Emotion, 33(6), 2019, 1099-1118. https://doi.org/10.1080/02699931.2018.1541312

[2] Breinholt, A.; Meier Jæger,M: How Does Cultural Capital Affect Educational Performance: Signals or Skills?, The British Journal of Sociology, 71(1), 2020, 28-46. https://doi.org/10.1111/1468-4446.12711

[3] Bresler, L.: What Formative Research Can Do for Music Education: A Tool for Informed Change, Visions of Research in Music Education, 16(5), 2021, 24.

[4] Camlin, D.A.; Tania, L.: The Digital 'Turn'in Music Education, Music Education Research 23(2), 2021, 129-138. https://doi.org/10.1080/14613808.2021.1908792

[5] Chen, X.: Compensated Fuzzy Neural Network-Based Music Teaching Ability Assessment Model, Computational Intelligence and Neuroscience, 2021, 2021. https://doi.org/10.1155/2021/3865190

[6] Dong, X.; Kang, X.; Xiaolei, D.: Influence and Analysis of Music Teaching Environment Monitoring on Students Mental Health Using Data Mining Technology, Journal of Environmental and Public Health 2022, 2022. https://doi.org/10.1155/2022/1120156

[7] Gorgoretti, B.: The Use of technology in Music Education in North Cyprus According to Student Music Teachers, South African Journal of Education, 39(1), 2019. https://doi.org/10.15700/saje.v39n1a1436

[8] Jia, H.: Design and Implementation of Personalized Teaching System for Ethnic Vocal Music Learning Resources Based on Computer Vision. 2023.
[9] Lee, K.; ChihMing, H.: Research on the Development of Music Information Retrieval and Fuzzy Search, Scientific and Social Research 4, (4), 2022, 1-10. https://doi.org/10.26689/ssr.v4i4.3771
[10] Liu, X.; Chengyu, Y.: Remote Music Teaching Classroom Based on Embedded System and Cloud Platform, Microprocessors and Microsystems 82, 2021, 103844. https://doi.org/10.1016/j.micpro.2021.103844
[11] Miksza, P.; Julia T. S.;Richerme,L.K; Phillip M.H.; Donald A. H.: Music Education Research: An Introduction, Oxford University Press, 2023. https://doi.org/10.1093/oso/9780197639757.001.0001
[12] Mittal,K.; Amita, J.; Kunwar S. V.; Oscar, C.; Kacprzyk.J: A comprehensive Review on Type 2 Fuzzy Logic Applications: Past, Present and Future, Engineering Applications of Artificial Intelligence 95, 2020, 103916. https://doi.org/10.1016/j.engappai.2020.103916
[13] Poulose, A.; Chinthala, S.R.; Shatarupa, D.; Bharat Jyoti Ranjan Sahu. Music Recommender System via Deep Learning, Journal of Information and Optimization Sciences, 43(5), 2022, 1081-1088. https://doi.org/10.1080/02522667.2022.2094546
[14] Seshadri, P.; Alexander, L.: Improving Music Performance Assessment with Contrastive Learning, Arxiv Preprint Arxiv:2108, 2021, 01711 .
[15] Tabuena, A.C.: Effectiveness of Classroom Assessment Techniques in Improving Performance of Students in Music and Piano, Global Researchers Journal 6, (1), 2019, 68-78. https://doi.org/10.2139/ssrn.3719451
[16] Wang, W.;Pan.J; Hua, Y.; Zhanmei, S.; Ming, L.: Audio-Based Piano Performance Evaluation for Beginners with Convolutional Neural Network and Attention Mechanism, IEEE/ACM Transactions on Audio, Speech, and Language Processing 29, 2021, 1119-1133. https://doi.org/10.1109/TASLP.2021.3061267
[17] Wang, X.;Yao, C.: Music Teaching Platform Based on FPGA and Neural Network, Microprocessors and Microsystems 80, 2021, 103337. https://doi.org/10.1016/j.micpro.2020.103337
[18] Wang, X.; Qun,W.;Yao, C.: Analysis of Music Online Teaching Curriculum Arrangement Based on BP Neural Network Model, In Journal of Physics: Conference Series, vol. 1648, (3), IOP Publishing, 2020, 032-101. https://doi.org/10.1088/1742-6596/1648/3/032101
[19] Wang, X.: Design of Vocal Music Teaching System Platform for Music Majors Based on Artificial Intelligence, Wireless Communications and Mobile Computing 2022, 2022, 1-11. https://doi.org/10.1155/2022/5503834
[20] Wang, Y.: The Intervention of Music Education on Students Mental Health Based on Fuzzy Computing, Mathematical Problems in Engineering 2022, 2022. https://doi.org/10.1155/2022/5632481
[21] Williamon, A.; Jane, G.; Rosie, P.; George, W.: Performing Music Research: Methods in Music Education, Psychology, and Performance Science, Oxford University Press, 2021. https://doi.org/10.1093/oso/9780198714545.001.0001
[22] Wu, Y.; Berlin, C.; Li, S.: Multi-instrument Automatic Music Transcription with Self-Attention-Based Instance Segmentation, IEEE/ACM Transactions on Audio, Speech, and Language Processing 28, 2020, 2796-2809. https://doi.org/10.1109/TASLP.2020.3030482
[23] Xin, X.; Yu, S.; Pang, N.; Dou, C.; Li, D.: Review on A big Data-Based Innovative Knowledge Teaching Evaluation System in Universities, Journal of Innovation & Knowledge 7, (3), 2022, 100197. https://doi.org/10.1016/j.jik.2022.100197
[24] Yang, Q.: Cloud Music Teaching Database Based on Opencl Design and Neural Network, Microprocessors and Microsystems 82, 2021, 103897. https://doi.org/10.1016/j.micpro.2021.103897

[25] Zhang, H.; Xuejun, C.; Tangyun, L.; Yijie, L.; Junwei, Wa.; Xiaomei, H.: A Hybrid Data Mining Method for EEG Analysis in Online Education Enhanced by Background Music, In 2020 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2020, 718-723. IEEE. https://doi.org/10.1109/WIIAT50758.2020.00110
[26] Zhang, X.: Multi-sensory Research of Singing visualization in Pre-School Music Education, Interactive Learning Environments, 2021, 1-11. https://doi.org/10.1080/10494820.2021.1912107