Implementation Of Neural Network Backpropagation Using Audio Feature Extraction For Classification Of Gamelan Notes

Fahri Firdausillah
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
fahri@dsn.dinus.ac.id

Dwiky Gilang Mahendra
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
gilang.mahendra888@gmail.com

Junta Zeniarja
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
junta@dsn.dinus.ac.id

Ardytha Luthfiarta
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
ardytha.luthfiarta@dsn.dinus.ac.id

Heru Agus Santoso
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
heru.agus.santoso@dsn.dinus.ac.id

Adhitya Nugraha
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
adhitya@dsn.dinus.ac.id

Erwin Yudi Hidayat
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
erwin@dsn.dinus.ac.id

Abdul Syukur
Department of Computer Science
Dian Nuswantoro University
Semarang, Indonesia
abah.syukur01@dsn.dinus.ac.id

***Abstract*—Gamelan consists of several musical instruments including kendang, saron, bonang, panerus, kempul, gender, gambang, kethuk, flute, siter, clempung, slenthem, demung, japan, kempyang, peking, and gong. Many people do not know the name of each gamelan instrument and how it sounds. In an effort to increase the popularity and introduction of gamelan as a traditional musical instrument to the community, this research proposes an analysis based on audio classification of gamelan musical instruments. Classification is done by data mining technique using Backpropagation Neural Network (BPNN) method. Audio or sound data from gamelan recordings is conducted by preprocessing using Zero Crossing Rate and Short Time Energy before processed to Backpropagation Neural Network. This research only classifies 4 gamelan musical instruments namely gong, kenong, saron, and gambang. The experimental result shows accuracy at 82.5 %.**

Keywords—Classification, Audio Processing, Neural Network, Backpropagation

I. Introduction

Gamelan is a traditional musical instrument of Indonesia in the region of Java. Its surroundings, made of bronze, is a mixture of tin and copper. Gamelan presented in Javanese arts with a *karawitan* that portray a certain rhythm tone in harmony with the gamelan as a vocal singing accompaniment instrument by *sinden* (singer) [1]. Among the people, the popularity of gamelan is declining [2]. This happens because of the way to play gamelan is not flexible and the purchase price and maintenance of gamelan instruments are very expensive. While the number of gamelan masters is not that much, the popularity of the gamelan may decreasing [3].

Gamelan consists of several musical instruments including kendang, saron, bonang, panerus, kempul, gender, gambang, kethuk, flute, siter, clempung, slenthem, demung, japan, kempyang, peking, and gong [1]. Many people do not know the name of each gamelan instrument and how it sounds. Gamelan is closely related to hearing and feeling. To understand music or sound, a person must be well trained in his or her hearing and feelings in playing the tones. If the hearing is not trained to hear the sound of a tone, the gamelan music player is difficult to determine the tones he or she is hearing. Similarly, the feelings of gamelan music players. If his feelings are not yet integrated with the existing tones played, chances are, the player will play a tone with an incompatible tempo. Along with the development of this very fast era of gamelan as a culture and art of traditional, start to be forgotten.

There are researches about gamelan audio classification using data mining methods, such Naïve Bayes. Primary data is taken directly by recording the sound of gamelan music instrument. The audio feature extraction used in the research are Zero Crossing Rate and Short Time Energy. Based on the results of research methods Naïve Bayes used for the classification of sound gamelan music instrument has a good performance is evidenced with an accuracy of 92.5 % [2].

In an effort to increase the popularity and introduction of gamelan as a traditional musical instrument to the community this research proposes to make an analysis of audio classification of musical instruments gamelan. Each sound generated from each gamelan instrument has different characteristics that vary by sound frequency [4]. Classification is done by data mining technique using Backpropagation Neural Network (BPNN), as Backpropagation outperforms the results of Naive Bayes [5][6]. BPNN is one of the reliable algorithms in classification with various types of data [7].The process of BPNN mimics the workings of the human brain that able to solve a problem by way of learning. BPNN is also able to adapt to learn from the input data provided, thus can map the relationship between input and output [8]. In addition, the ability of BPNN is well known in predicting output based on inputs that have been trained previously. This research uses BPNN as a method of gamelan audio classification and feature extraction of audio use Zero Crossing Rate and Short Time Energy. Classification is based on gamelan audio data that has different frequencies.

II. Related Study

There are several studies that serve as the study review in this study. Some related researches on gamelan and the use of neural network are described as follows.

In [2], the author discusses the use of data mining methods naïve bayes for gamelan sound classification. Primary data is taken directly by recording the gamelan music instrument. Gamelan instruments consist of kendang, saron, bonang, panerus, kempul, gender, gambang, kethuk, flute, siter, clempung, slenthem, demung, japan, kempyang, peking, and gong. According to the authors, many people do not know the name of each instrument of gamelan and how his voice. The sound of gamelan has different frequency in every gamelan instrument. Based on the results of research methods naïve bayes used for the classification of sound gamelan music instrument has a good performance with an accuracy of 92.5 %.

In [9], the author uses a radial basis function network for the identification of Javanese gamelan instruments. The method used for radial base network function training is the extended Kalman filter. Gamelan sound data are obtained from the recording of the three instruments of the saron group. There is no standard frequency in the ladder system of gamelan, as in modern musical instruments. This research was identifying Javanese gamelan using radial base function network. Radial basis function network is a multilayer feed-forward network whose training is hybrid. The results showed that the identification of Javanese gamelan instruments using radial basis function network trained using extended Kalman filters resulted in high accuracy of over 90 %. Thus the radial basis function network with training using extended Kalman filters can be used for the identification of Javanese gamelan instruments.

In [10], a system for voice recognition that focuses on numbers by applying MFCC and SOM Neural Network training is developed. Primary data is as input for training and testing steps. The system designed in this paper is a voice recognition application using the Mel method cepstral coefficients method to extract features and self-organizing maps to recognize sound. This application can recognize a voice that focuses on numbers. Based on the results of research the accuracy obtained by the proposed method is Mel frequency cepstral coefficients for extracting features and self-organizing maps to recognize notes of 82 percent.

Discussion of the classification of human voices is proposed by [11]. The algorithm used is Levenberg-Marquadt and Backpropagation Neural Network. Data is taken directly in the form of human voice recordings. The current speech recognition technology has improved especially in terms of speech processing. Speech processing is a way to extract the desired information from a sound signal. This study discusses the male and female voice classification system. Extracting the characteristics of the sound signals of each frame in the time zone and the frequency region is helpful to simplify and speed up the calculations. The features for voice or audio are Short Time Energy, Zero Crossing Rate, Spectral Centroid and others. The results of the system test show that the classification of human voice by using artificial neural network backpropagation and Levenberg-Marquadt algorithm for weight matrix changes, very good and fast because of the complexity of calculation that is not too high. Using 40 voices sample database with 5 test data, the result shows a classification that has been identified by a similarity value> = 0.5 as male and <0.5 as female. Tests using artificial neural networks resulted in average success rate in sound classification is 91%.

From the related study above, this research purposed a model to classify the sound of gamelan instruments using Backpropagation Neural Network and measuring its accuracy. To perform the audio classification of gamelan, Matlab is used. The data is the recording of the sound of gamelan instruments. The classification method is the Backpropagation Neural Network using audio extraction feature Short Time Energy and Zero Crossing Rate.

III. Experimental Result

A. Feature Extraction

The preprocessing stage is the processing of data from voice data into numerical values in order to be processed by BPNN method. The feature extraction stage consists as follows:

1. Short Time Energy(STE): marks the violence of the voice at a short time whos taken from

$$\frac{1}{N}\sum_{i=1}^{N} X(n)^2$$

- N: Number of samples / length of frequency
- X (n): The value of the signal from the sample

2. Zero Crossing Rate (ZC): denotes a sequential sample on a digital signal has a different sign, the size of a signal's noise in the domain feature whos taken from

$$ZC = \frac{\sum_{n=1}^{N} |sgn\ x(x)\text{-}sgn\ x(n-1)|}{2N}$$

- sgn x (n): the value of x (n), is 1 if x (n) is positive, -1 if x (n) is negative
- N: Number of Samples

Table 1 shows the feature extraction result of the gambang instrument. The table shows the file name, STE (Short Time Energy), ZC (Zero Crossing Rate) and Label.

Table 1: Dataset Extraction Features of Gambang

file	STE	ZC	label
gmb1 1.wav	0.0062	0.0293	gambang
gmb1 2.wav	0.052746	0.002318	gambang
gmb1 3.wav	0.052558	0.002319	Gambang
…	….	….	….
gmb9 5.wav	0.037425	0.010657	gambang

Table 2 shows the feature extraction result of the gong instrument. The table shows the file name, STE (Short Time Energy), ZC (Zero Crossing Rate) and Label.

Table 2: Dataset Extraction Features of Gong

file	STE	ZC	label
gong sedang 1.wav	0.005763	0.077479	gong
gong sedang 2.wav	0.005264	0.085212	gong
gong sedang 3.wav	0.005279	0.07921	gong
gong sedang 4.wav	0.005232	0.090106	gong
gong sedang 5.wav	0.005502	0.087265	gong
….	….	….	….
gs5 5.wav	0.00911	0.063908	gong

Table 3 shows the feature extraction result of the Kenong instrument. The table shows the file name, STE (Short Time Energy), ZC (Zero Crossing Rate) and Label.

Table 3: Dataset Extraction Features Data Kenong

File	STE	ZC	label
knp1 1.wav	0.028715	0.042177	kenong
knp1 2.wav	0.029188	0.041141	kenong
knp1 3.wav	0.028351	0.05063	kenong
knp1 4.wav	0.026956	0.044046	kenong
knp1 5.wav	0.027862	0.046402	kenong
…	….	….	….
….	….	….	….
kns6 5.wav	0.021367	0.057533	kenong

Table 4 shows the feature extraction result of the saron instrument. The table shows the file name, STE (Short Time Energy), ZC (Zero Crossing Rate) and Label.

Table 4: Dataset Extraction Features of Saron

file	STE	ZC	Label
sk1 1.wav	0.109531	0.0777	Saron
sk1 2.wav	0.109128	0.038031	Saron
sk1 3.wav	0.109064	0.04372	Saron
sk1 4.wav	0.107416	0.051157	Saron
sk1 5.wav	0.109092	0.06176	Saron
…	….	….	….
….	….	….	….
ss5 5.wav	0.041064	0.074181	Saron

B. Backpropagation Neural Network

1) BPNN Training

Research conducted in backpropagation network training using neural network parameters in Table 5. The amount of parameters depicted in the table are used due to the most optimal value regarding to an accuracy percentage and computing power, in such lower value will result lower accuracy but higher require more computing power without any significant improvement in accuracy.

Table 5: Backpropagation Neural Network Training Parameters

Parameter	**Value**
Epoch	10000
Learning Rate	0.01
Target Error	0.00000001
Hidden Layer	100 Neuron

Based on network parameters in Table 6 training is done with the process that can be seen in Fig 1.

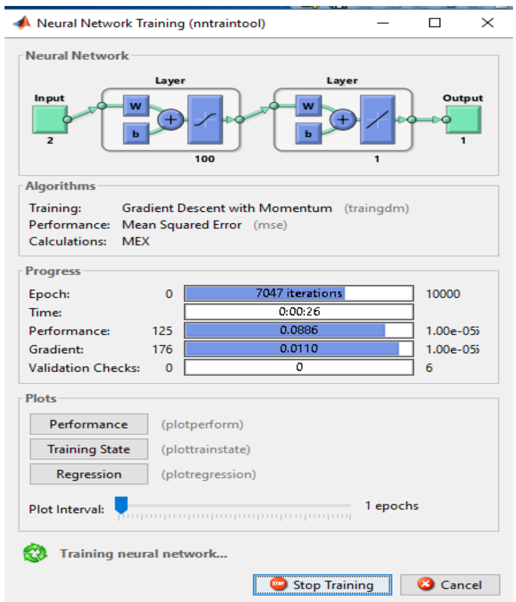

Fig 1. Backpropagation Neural Network Training

2) Implementation

This chapter describes the implementation of BPNN in the GUI using Matlab. Here are the steps of the classification that have been applied to the GUI.

a) Interface of Application

When the source code is executed, the initial view is shown in Figure 2. In the program view there are several components such as file formats and feature extraction results on the file.

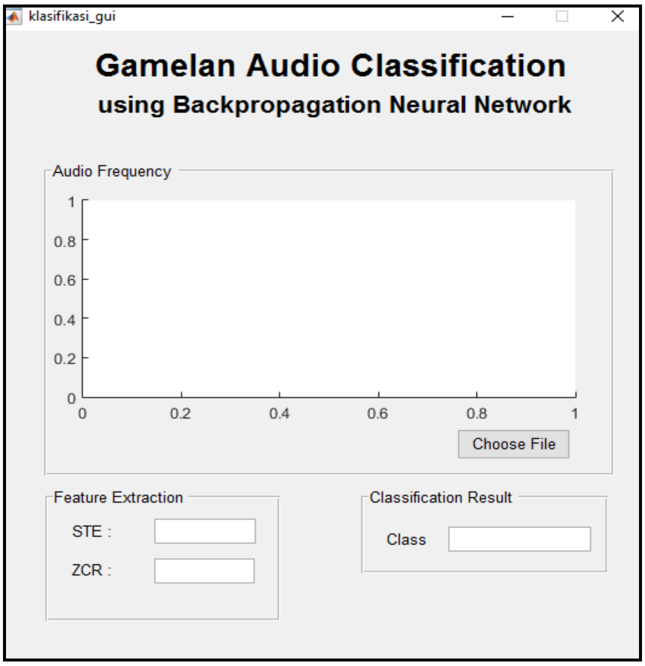

Fig. 2. Interface of Application

b) Audio File Selection

To be able to classify the initial stage is to select the file by pressing the file select button. Then the system will display a file selection dialog that is only in the filter file existence .wav

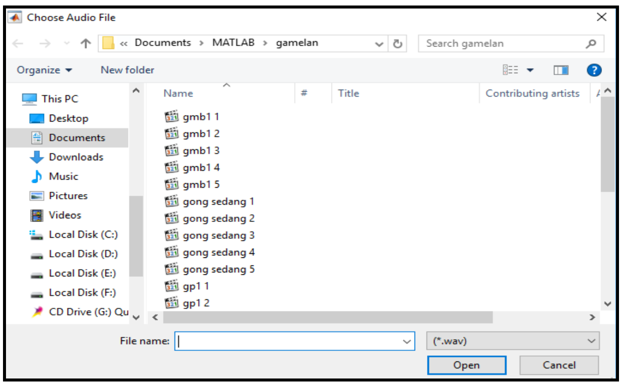

Fig.3. Select Audio File

c) Classification Results

After selecting the file, the system will automatically display all results starting from the feature extraction and the result of the classification of the selected audio file.

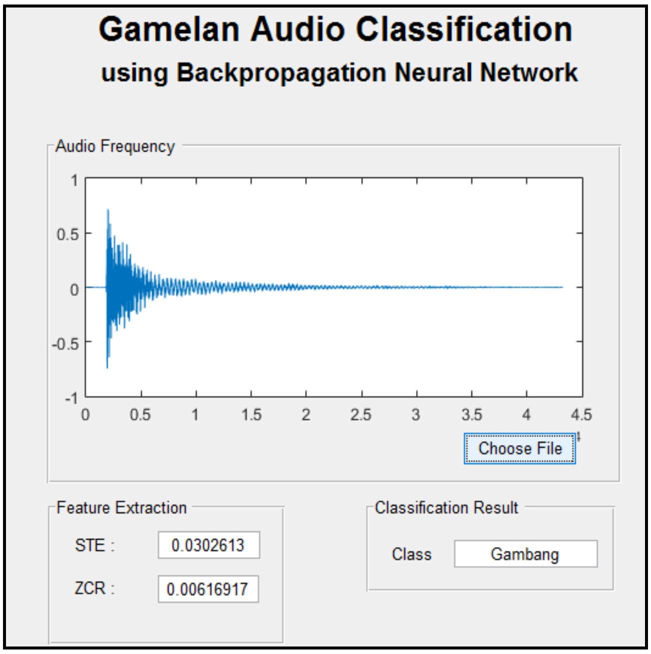

Fig. 4. Classification Result of Gambang

3) Evaluation

Once the Backpropagation Neural Network method has been successfully applied to the classification, tests are performed to determine the performance of the proposed method. The evaluation is done 40 times by using cross validation where training data is checked towards Backpropagation Neural Network model which has been made. After, the result of the test value is compared to available dataset.

Based on the test table results, confusion matrix is arranged to calculate the accuracy of the proposed classification.

Table 7. Evaluation Model using Confusion Matrix

		Class			
		Gambang	Gong	Kenong	Saron
Predicted Class	Gambang	9	0	1	0
	Gong	0	9	0	1
	Kenong	1	0	9	2
	Saron	0	1	1	6

$$Accuracy = \frac{True_{gambang}+True_{gong}+True_{kenong}+True_{saron}}{gambang+gong+kenong+saron} \quad (1)$$

$$Accuracy = \frac{9+9+9+6}{10+10+10+10} = \frac{33}{40}$$

$$Accuracy = 82.5\%$$

Based on the experimental results that have been done in the initial process to produce the model. The model is then evaluated using confusion matrix to calculate the accuracy value. As shown in above calculation, the accuracy value of the evaluation reached 82.5%.

IV. Conclusion And Future Work

The proposed data mining method of Backpropagation Neural Network has been successfully applied for gamelan sound classification. The application of Backpropagation Neural Network method used for the classification in this study has an accuracy of 82.5 %.

Further research can apply and perform additional features used for classification so as to add the data calculation for the classification process to improve its performance. More in-depth research and other variations of methods for classification such as c45, Decision Tree, SVM, as well as optimization methods such as Genetic Algorithms and Particle Swarm Optimization are also another option.

References

[1] J. Lindsay, Javanese Gamelan, Oxford : Oxford University Press, 1St Edition edition, 1992.

[2] I. NUGROHO, "Analisa Dan Klasifikasi Suara Gamelan Menggunakan Algoritma Naïve Bayes," 2017.

[3] Y. T. C. Pramudi, F. Budiman and S. , "Desain Virtual Gamelan Jawa Sebagai Media Pembelajaran," in Seminar Nasional Aplikasi Teknologi Informasi 2010 (SNATI 2010), Yogyakarta, 2010.

[4] A. Rizky, "Identifikasi Instrumen Gamelan Jawa Menggunakan Jaringan Fungsi Basis Radial Dengan Metode Pelatihan Extended Kalman Filter," in Prosiding Seminar Nasional Matematika, 2014.

[5] M. Simfukwe, "Comparing Naive Bayes Method and Artificial Neural Network for Semen Quality Categorization," in International Journal of Innovative Science, Engineering & Technology, 2015.

[6] A. . A. S. "Heart Disease Prediction System using Multilayered Feed Forward Neural Network and Back Propagation Neural Network," in International Journal of Computer Applications, 2017.

[7] David, "Penerapan Algoritma Levenberg-Marquadt dan Backpropagation Neural Network Untuk Klasifikasi Suara Manusia," 2013.

[8] A. Hamid, "Fast speaker adaptation of hybrid NN/HMM model for speech recognition based on discriminative learning of speaker code," in IEEE International Conference on, 2013.

[9] A. Riski, "Identifikasi Instrumen Gamelan Jawa Menggunakan Jaringan Fungsi Basis Radial Dengan Metode Pelatihan Extended Kalman Filter," in Prosiding Seminar Nasional Matematika, 2014.

[10] M. Puspitasari, "Pengenalan Suara Menggunakan Mel Frequency Cepstral Coefficients Dan Self Organizing Maps," 2014.

[11] David, "Penerapan Algoritma Levenberg-Marquadt dan Backpropagation Neural Network Untuk Klasifikasi Suara Manusia," 2017.