# IMPLEMENTATION MEL-FREQUENCY CEPSTRUM COEFFICIENTS (MFCC) METHOD IN IDENTIFICATION OF SOUND BIRDS CHIRPING

Fakhrizal Ahadiat

Informatics Engineering – Indonesia Computer University
Jl. Dipatiukur 112-114 Bandung
E-mail : isal_ada@yahoo.co.id

## ABSTRACT

*Bird is a member of the group of animals with (vertebrates) that had feathers and wings. Birds chirping in Indonesia is one of the most attractive bird species, it can be seen more and more people who began his hobby in maintaining and competitions birds chirping sound of birds frequently and regularly held in each city. Hence the need for a convenience in identifying the type of bird hobbyist birds chirping mainly on the type of birds that have characteristic shapes and colors are almost the same.*

*In a previous research conducted by Risha Annisa, Bambang Hidayat and Inung Wijayanto sound classification of the types of birds using Wavelet Packet Decomposition and Neural Network SOM with a sample number of four species of birds found in the identification accuracy of 83.13%. On the research and preparation of this thesis for identifying the type of bird sounds using Mel-Frequency Cepstrum Coefficients and Neural Network SOM.*

*Based on the results of the test can be concluded that the method MFCC and Neural Network SOM can detect the type of bird sounds with accuracy rate of 85.625% in the case of same bird species, while in the case of bird species who each have sub-types with similar characteristics can detect the sound the accuracy rate of 48.824%, so it can be concluded the system using MFCC as feature extraction can identify the type of bird it is better to cases of bird species similar to previous research conducted*

***Keyword:*** *Birds Chirping, Feature Extraction, MFCC, Neural Network SOM.*

## 1. INTRODUCTION

### 1.1 Background

Bird is a member of the group of animals with (vertebrates) that had feathers and wings. The types of birds are so varied, began from tiny hummingbirds to ostriches much higher than people. It is estimated that the world's bird species there are about 8,800 to 10,200 species, and about 1650 species of birds were found in Indonesia. Birds can be divided into two kinds, namely including ornamental birds and birds chirping. Birds chirping in Indonesia is one of the most attractive bird species, it can be seen more and more people who began his hobby in preserving birds and also competitions the sound of chirping birds frequently and regularly held in each city. Hence the need for a convenience in identifying the type of bird hobbyist birds chirping mainly on the type of birds that have characteristic shapes and color that almost identical.

Voice biometrics is one field that is currently growing. In a previous study conducted by Risha Annisa, Bambang Hidayat and Inung Wijayanto sound classification of the types of birds using Wavelet Packet Decomposition and Neural Network Self Organizing Maps with a sample size of 4 species of birds found in the identification accuracy of 83.13% [1] because it still has shortcomings in the process of identifying the type of bird sounds one of which is not yet able to distinguish the sound characteristics of the two types of birds including bird sounds of cucak ijo and cucak rawa. Hence the need for a comparison between the processing method in order to choose the appropriate method for the extraction process and the identification of types of birds chirping. While other studies in the field of voice biometrics is done by Angga Setiawan, Achmad Hidayatno and R. Rizal Isnanto with the introduction of the title is not dependent reciter text using Linear Predictive Coding (LPC) and Vector Quantization obtained an average percentage of the highest recognition of the testing with record data only reached 68.5% and the highest recognition to the test in real time at 58% [2].

In this research method that will used to build a system of identification of the type birds that use the Mel Frequency Cepstrum Coefficients method (MFCC) for the extraction of sound signals and for identification helped by using classification methods artificial neural network Self Organizing Maps. Mel

Frequency Cepstrum Coefficients method is used because it has several advantages including able to capture important information in a sound signal, generating a minimum of data without eliminating existing information and replicate the human auditory organ in conducting the perception of the sound signal [3]. It was proven in previous research to verify the voice of people using the MFCC obtained the highest accuracy is approximately 93.254% [4].

Based on the description above, in this research, by implementing the Mel-Frequency Cepstrum Coefficients (MFCC) method and Artificial Neural Networks Self Organizing Maps method are expected to produce a more accurate and less time on the processing of sound digital to the identification of the type of bird sounds , which can help the people who started the hobby in maintaining the birds to identify the birds by sound with more accurate and faster.

### 1.2 Purpose and Objectives

The purpose of this research is to implement Mel-Frequency Cepstrum Coefficients method in extraction characteristic signal sound of birds chirping and Neural Network Self Organizing Maps method in the identification of bird species.

The objectives to be achieved in this research, are::

1. Implement and measure the level of accuracy of the Mel-Frequency Cepstrum Coefficients method and Neural Network Self Organizing Maps in identification of birds chirping.

## 2. BASIC THEORY

Artificial Intelligence (Artificial Intelligence, or AI) is defined as intelligence demonstrated by an artificial entity. Systems like this are generally considered to be a computer. Intelligence was created and inserted into a machine (computer) in order to do the job as humans do [6].

### 2.1 Relationships Artificial Intelligence with Voice Recognition

Understanding of Speech/Voice (Speech/Voice Understanding), is a technique in order to computer can recognize and understand language for the speech. This process allows one to communicate with the computer by talking to him. The term "voice recognition" means that the ultimate goal is the spoken word without having to know what it means, where part of it is the task of "a sound understanding".

### 2.2 MFCC

MFCC (Mel Frequency Cepstrum Coefficients) is one of method that is widely used in the field of speech technology, both the speaker and the speech recognation. This method is used to perform feature extraction, a process that converts voice signals into several parameters. The advantage of this method is:

1 Able to capture the sound characteristics which is very important for speech recognition. Or in other words, able to capture important information are contained in the voice signal.

2 Generating a minimum of data without losing important information.

3 Adapting the human auditory organ in doing the perception of the sound signal.

MFCC feature extraction is actually an adaptation of the human auditory system where the voice signal will be a linear filter for low frequencies (below 1000 Hz) and logarithmically for high frequency [3]. Here is a block diagram for the process performed on the method of MFCC:

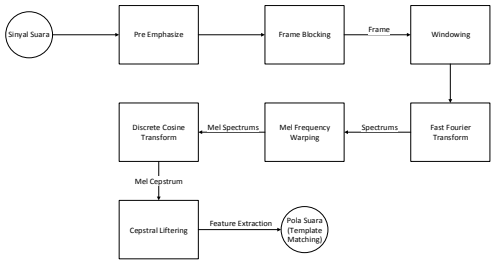



Figure 1 Block Diagram Flow Process MFCC

#### 2.2.1 Pre-emphasize

Pre-emphasis performed to improve the signal from noise nuisance, so it can improve the accuracy of the process of feature extraction. Default value of alpha used in the process of pre-emphasis filtering is 0.97.

$$y[n] = s[n] - \alpha\, s[n-1] \quad (1)$$

Remarks:
*y*[n] = signal from results pre-emphasis filter
*s*[n] = signal before pre-emphasis filter

#### 2.2.2 Frame Blocking

The results of voice recording an analog signal that is in the time domain who are time variant, which is a time-dependent function. Hence the signal should be cut in certain time slots in order could be considered invariant. Each piece of the voice signal is called frame.

To calculate the number of frames used formula:

$$Total\ Frame = ((I - N)/M + 1) \quad (2)$$

Remarks:
I = Sample rate
N = Sample point (*Sample rate* * Frame Time (s))
M = N/2

### 2.2.3 Windowing

Window function are most commonly used in speaker recognition applications is Hamming Window. This function generates sidelobe level is not too high (less than -43dB) in addition to the noise generated is not too large (approximately 1:36 BINS).

Hamming Windows:

$$W_{ham}(n) = 0.52 - 0.46\cos\left[\frac{2\pi n}{(N-1)}\right] 0 \leq n \leq N-1 \quad (3)$$

Remarks:
$W_{ham}(n)$ = Hamming Window
$n$ = Signal to [1 ... n]
$N$ = Sample Point

### 2.2.4 Fast Fourier Transform (FFT)

Windowing process produces the sound spectrum in the time domain, to avoid mistakes in the process of warping path so be converted into time domain spectrum frequency signal by using the Fast Fourier Transform.

FFT (Fast Fourier Transform) is the technique of fast calculation from DFT. FFT is a DFT with rapid calculation technique by utilizing the periodical nature of Fourier transformation. Note the definition of FFT:

$$F(k) = \sum_{n=1}^{N} f(n)e^{-j2\pi knT/N} \quad (4)$$

Or can be written by:

$$F(k) = \sum_{n=1}^{N} f(n)\cos\left(\frac{2\pi nkT}{N}\right) - j\sum_{n=1}^{N} f(n)\sin(2\pi nkT/N) \quad (5)$$

To view the value of the FFT used formula:

$$|f(u)| = [R^2 + I^2]^{1/2} \quad (6)$$

### 2.2.5 Mel Frequency Warping (FilterBank)

Magnitude result of the Fast Fourier Transform (FFT) is then going through a phase Filterbank. Filterbank aims to determine the size of energy on Frequency Bandwidth in the voice signal. The first step taken is to look for the value of the coefficient filterbank.

Here is the formula used in the calculation filterbank.

$$Y[i] = \sum_{j=1}^{N} S[j]H_i[j] \quad (7)$$

Remarks
$N$ = the number of magnitude spectrum
$S[j]$ = magnitude spectrum in the frequency j
$H_i[j]$ = coefficient filterbank on frequency j
$(1 \leq i \leq M)$
$M$ = the number of channels in the filterbank
Where $H_i = \frac{mel\ f}{x_{i/2}}$

### 2.2.6 Discrete Cosine Transform (DCT)

This process is the final step of feature extraction. The results of this DCT are features that are required to process an analysis of the voice recognition. using the formula:

$$T_n = \sum_{k=1}^{K}(log\ S_k)\cos[n(k-\tfrac{1}{2})\tfrac{\pi}{K}] \quad (8)$$

Remarks:
$S_k$ = output of the filter bank on the index $k$
$K$ = number of coefficients expected

### 2.2.7 Cepstral Liftering

Cepstral, the result of the DCT function actually is the end result of the process of feature extraction but it has some drawbacks. Low-order of cepstral coefficients very sensitive to spectral slope, while the high order of his very sensitive to noise. Hence, the cepstral liftering become one of the standard techniques applied to minimize the sensitivity.

Cepstral liftering can be done by implementing the window function to the cepstral features.

$$W[n] = \{1 + \frac{L}{2}\sin(\frac{n\pi}{L})\} \quad (9)$$

Remarks:
$L$ = number of cepstral coefficients
$N$ = index of cepstral coefficients

### 2.3 SOM Neural Network

Kohonen neural network - SOM is one of the ANN models using unsupervised learning method, are studying the distribution of the set of patterns without class information. This network consists of two layers (layer), that the input layer and output layer. Each neuron in the input layer is connected to every neuron in the output layer. Each neuron in the output layer represents a class (cluster) of the a given input. The following illustration ANN architecture SOM:

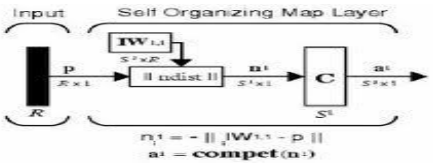



Figure 2 JST-SOM Architecture

In SOM ANN, the neurons target are not placed in a row like another neural network model. Neurons target placed in a two-dimensional shapes/topology can be organized.

There are three kinds of topologies can be formed that is [7]:

1. Gridtop

In the top grid of neurons arranged in a two-dimensional array with a square shape.

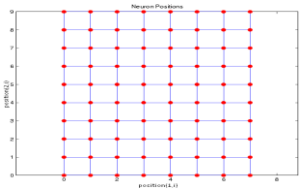


Figure 3 Gridtop Topology

2. Hextop

In hextop neurons are arranged in two-dimensional array with a hexagonal shape.

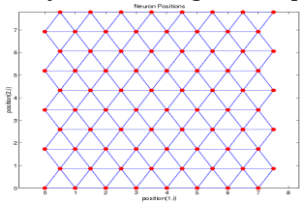


Figure 4 Hextop Topology

3. Randtop

In randtop neurons are arranged in two-dimensional array with random arrangement.

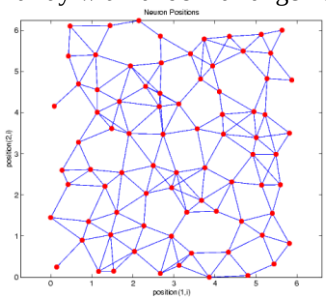


Figure 5 Randtop Topology

# 3 ANALYSIS

## 3.1 General Description System

This section includes a general description of the system to be built and about the method that will applied to the system. In the system of identification of the sounds of birds, there are several stages. The stages of the system can be seen in Figure 6:

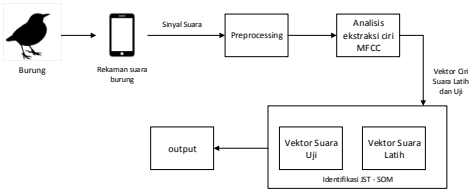



Figure 6 General Description System

Explanation of the the Figure above is as follows:

1. The first phase of this system that is the sound of birds recorded using a mobile phone with voice recording duration 3 seconds later the sound of birds recordings are saved to the computer in order the input sound can be processed by the system.
2. *Pre-processing*
   This stage is to convert analog voice signal into a digital signal with a signal through the stages of conversion, data acquisition, denoising, normalization and Remove Silent/Cropping stages.
3. Feature Extraction MFCC (Mel-Frequency cepstral Coefficients) Analysis
   This stage are the most important stages in the development of an identification system that is the process of extracting a sound signal that has been converted into a digital signal. The process of feature extraction is performed twice: at the stage of train sound and phase sound test. MFCC feature extraction analysis stage consists of the seven processes, namely Pre emphasize, Frame Blocking, Windowing, Fast Fourier Transform, Mel Frequency Warping, Discrete Cosine Transform, and cepstral liftering.
4. Identification of Neural Networks Self Organizing Maps
   Identification of Neural Networks Self Organizing Maps consists of the two parts, namely the training data and test data. In the process of training data process is carried out to determine the grade span of the JST-SOM. The results will be used to identify species of birds that have been adjusted in advance with a class that has been generated by JST-SOM trained.

## 3.2 Algorithm Analysis

This research uses sound feature extraction Mel-Frequency Cepstrum Coefficients (MFCC) and Neural Networks Self Organizing Maps methods in terms of identification. In this system the analysis phase is devoted in the extraction process and the sound birds chirping voice identification, recognition chart as a whole can be seen in Figure 7.

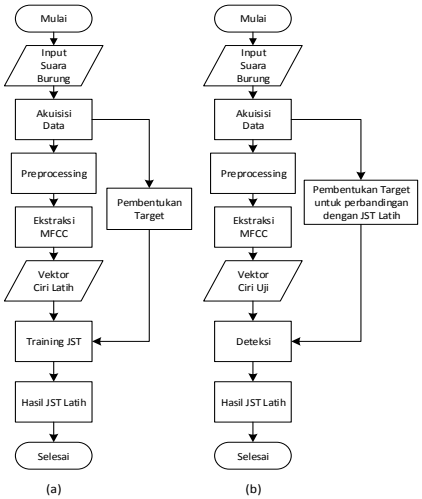



Figure 7 Illustration of Flow Sounds Training Process (a) and Sound Test Process (b)

### 3.3 Pre-processing Analysis

#### 3.3.1 Denoising

This process is necessary to eliminate the unwanted noise that is carried on in the recording process so that the voice data input into the extracted characteristics have a better quality.

#### 3.3.2 Normalization

This process aims to determine the amplitude of the each bird sounds recorded by the system so as to be in the range of -1 and +1. Equalization is done because the maximum amplitude or power levels generated sample subject on each recording is not definitely same. So as to minimize the difference in power used the process of normalization amplitude.

#### 3.3.3 *Remove Silence*

Remove Silence process performed to eliminate the silence area of the sound signal to improve the accuracy of the system. This process is done by finding the standard deviation of the signal sound. The standard deviation is calculated by the following formula:

$$std = \left(\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{1/2} \quad (10)$$

For the standard deviation value set are 0:01

Signal value smaller than the value of the standard deviation of the signal noise will be considered a silent area. Having obtained the silent area, sought the maximum and minimum index of the the silent region with the aim of cutting the silent area of the the index the minimum to the maximum area. Thus, the silent region is an area that is cut beginning and end of the sound signal. Here's an overview of the concept of Pre-processing:

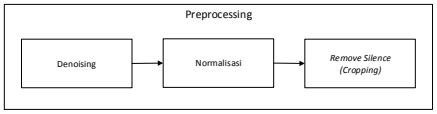



Figure 8 Concept Preprocessing

### 3.4 MFCC Analysis

Methods Mel-Frequency Cepstrum Coefficients (MFCC) is using some parameters that will play a vital role in determining the success rate of sound signal recognition. Here is the entire process of MFCC feature extraction:

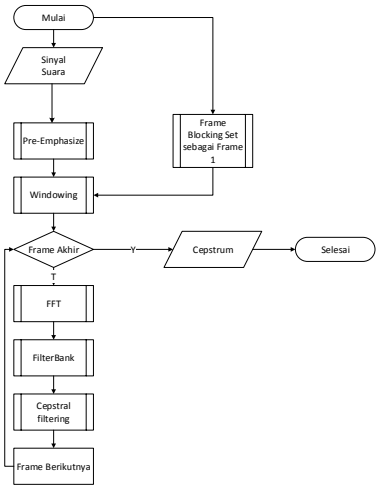



Figure 9 Mel Frequency Cepstrum Coefficients methods process

### 3.5 Identification of Neural Networks Self Organizing Maps Analysis

To be able to identify the type of birds chirping, the new incoming data signals are matched with the data that has been trained by JST-SOM earlier. The parameters affecting the JST-SOM topology practice is used, the function of the distance and repetitions each feature vector or epoch. Topology is used for the identification of species of birds in this research is using topology Gridtop, Hextop and Randtop. Distance function used is Euclidean Distance, Distance and Manhattan Distance Box with epoch number are 150, 200, 250 and 300.

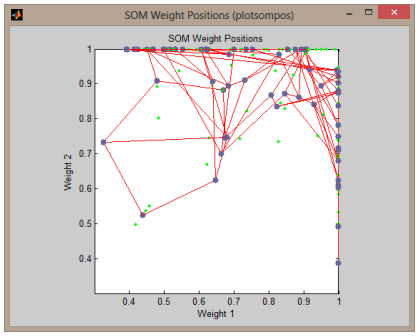



Figure 10 Weight Position Train JST-SOM Sounds Bird

Figure above is an example of the result of the practicing JST-SOM extraction epoch sound of birds with the number 150, topology gridtop and distance function used is euclidean distances. From these results to matching the extracted in the testing process are taken by taking the value of the test results closer to one neuron training results. When one neuron to train selected as the neurons that best matches the value of the test resultsso neurons are disposed trainer.

# 4 IMPLEMENTATION AND TESTING

## 4.1 Hardware Implementation

The hardware used to implement the system can be seen in Table 1.

Table 1 Hardware Specifications

| No. | Hardware | Specifications |
|---|---|---|
| 1. | Prosesor | intel i3 2.30 GHz |
| 2 | Monitor | Monitor 14" 1366 x 768 (32Bit) (60Hz) |
| 3 | VGA | 512Mb |
| 4 | Memory (RAM) | 2GB |
| 5 | Keyboard and Mouse | Standard |
| 6 | Speaker | Standard |

## 4.2 Software Implementation

The software used to build and implement the system can be seen in table 2.

Table 2 Software Specifications

| No. | Software | Specifications |
|---|---|---|
| 1 | Operating System | Windows 7, 8, dan 8.1 |
| 2 | Tools | Matlab R2009 and then |
| 3 | Programming Language | Matlab |

## 4.3 Performance Tests

In the classification SOM ANN-influencing parameters are topology, the function of the distance, and the repetition of each feature vector or epoch is used. Topology used in the classification is Gridtop, Hextop, and Randtop. The distance function used is Euclidean Distance, Distance Box, Manhattan Distance, and Link Distance. While the number of epochs used were 150, 200, 250, and 300. Here are the parameters used in the calculation accuracy of the system:

a. Accuracy

Accuracy is a measure of the accuracy of the system to recognize input is given so as to produce the correct output. Mathematically, the equation can be written:

$$Akurasi = \frac{The_amount_data_correctly}{Total_data} \times 100\%$$

b. Error

Error is the system error rate in terms of recognizing a given input to the system overall amount of data. Mathematically, the equation can be written:

$$Error = \frac{the_amount_data_incorrectly}{Total_data} \times 100\%$$

### 4.3.1 Case Testing with Total Sample 240

Here is a test based on the type birds and the same amount of data with previous research conducted using 4 kinds of sounds birds. sound recordings used are doves, canaries sound, the sound of cucak ijo and the sound of cucak rawa. In testing the feature extraction system used is MFCC and classifications are JST-SOM. ANN-SOM trained first by using 20 training data on each individual birds so that the amount of voice data train consisting of the 80 data. For the test data used are 40 different test data with training data on each birds so that the amount of data tested to be 160 data.

**1. Results of Previous Research Testing**

Table 3 The Results of Testing Accuracy On Previous Research

| Type | Gridtop | | | | Hextop | | | | Randtop | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link |
| Perkutut | 33 | 38 | 39 | 38 | 22 | 39 | 38 | 38 | 34 | 38 | 39 | 30 |
| Kenari | 18 | 34 | 32 | 32 | 17 | 35 | 34 | 36 | 29 | 36 | 36 | 38 |
| Cucak Ijo | 18 | 16 | 26 | 28 | 12 | 22 | 18 | 21 | 14 | 29 | 15 | 9 |
| Cucak Rawa | 29 | 24 | 30 | 35 | 27 | 26 | 26 | 23 | 22 | 30 | 23 | 34 |
| Total | 98 | 112 | 127 | 133 | 78 | 122 | 116 | 118 | 99 | 133 | 113 | 111 |
| Akurasi (%) | 61,25 | 70 | 79,38 | 83,13 | 48,75 | 76,25 | 72,5 | 73,75 | 61,88 | 83,13 | 70,63 | 69,38 |
| Error (%) | 38,75 | 30 | 20,63 | 16,88 | 51,25 | 23,75 | 27,5 | 26,25 | 38,13 | 16,88 | 29,38 | 30,63 |

The accuracy results obtained by previous research that is equal to 83.13% occurred in the epoch number 200, topology Randtop and function of the Box distance.

**2. The Result of Testing MFCC Methods**

Table 4 Results of Testing Accuracy with using MFCC method

| Type | Gridtop | | | | Hextop | | | | Randtop | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link |
| Perkutut | 40 | 34 | 35 | 39 | 39 | 34 | 35 | 38 | 40 | 40 | 40 | 40 |
| Kenari | 31 | 26 | 32 | 34 | 31 | 33 | 34 | 30 | 39 | 39 | 31 | 34 |

| Type | Gridtop | | | | Hextop | | | | Randtop | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link |
| Cucak Ijo | 9 | 9 | 13 | 10 | 9 | 15 | 13 | 9 | 18 | 15 | 18 | 15 |
| Cucak Rawa | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 |
| Total | 120 | 109 | 120 | 123 | 119 | 122 | 122 | 117 | 137 | 134 | 129 | 129 |
| Akurasi (%) | 75 | 68,125 | 75 | 76,875 | 74,375 | 76,25 | 76,25 | 73,125 | 85,625 | 83,75 | 80,625 | 80,625 |
| Error (%) | 25 | 31,875 | 25 | 23,125 | 25,625 | 23,75 | 23,75 | 26,875 | 14,375 | 16,25 | 19,375 | 19,375 |

The accuracy results obtained by using the MFCC that is equal to 85.625% occurred in the epoch number 150, topology Randtop and function of the Euclidean distance.

### 4.3.2 Case Testing with Total Sample 510

Here is a test using seven kinds of sounds birds, each birds has a sub species so that the total number of bird species that is 17 subtypes. Footage used is sub-type of the sound perkutut, kenari sound, sound of jalak, gelatik bird sounds, sound of a kolibri, the sounds of kacer and the sound of a murai batu. In testing the feature extraction system used is MFCC and classifications are JST-SOM. ANN-SOM trained first with 10 training data on each individual birds so that amount of sound data train consisting of 170 data. For the test data used are 20 different test data with training data on each birds so that amount of data tested are the data 340.

Tabel 5 Accuracy Testing Results by Using 510 Sample

| Type | | Gridtop | | | | Hextop | | | | Randtop | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link | Euclidean | Box | Manhattan | Link |
| Gelatik | Jawa | 8 | 10 | 13 | 9 | 13 | 10 | 10 | 9 | 9 | 11 | 10 | 10 |
| | Wingko | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 4 | 0 |
| Jalak | Kerbau | 10 | 12 | 10 | 10 | 10 | 10 | 12 | 7 | 10 | 12 | 10 | 10 |
| | Nias | 8 | 13 | 8 | 9 | 10 | 14 | 9 | 7 | 8 | 12 | 9 | 11 |
| Kacer | Jawa | 7 | 7 | 6 | 7 | 5 | 7 | 5 | 7 | 7 | 5 | 5 | 7 |
| | Madagaskar | 13 | 15 | 13 | 7 | 13 | 8 | 16 | 15 | 13 | 12 | 15 | 12 |
| | Poci | 5 | 4 | 5 | 7 | 5 | 4 | 5 | 5 | 3 | 3 | 4 | 5 |
| Kenari | Border | 17 | 17 | 18 | 17 | 19 | 17 | 17 | 18 | 18 | 18 | 16 | 17 |
| | Norwich | 4 | 2 | 1 | 4 | 1 | 0 | 0 | 2 | 3 | 4 | 1 | 0 |
| | Yorkshire | 10 | 8 | 10 | 10 | 9 | 9 | 9 | 9 | 10 | 10 | 9 | 10 |
| Kolibri | Manggar | 20 | 19 | 19 | 18 | 19 | 19 | 20 | 18 | 20 | 20 | 20 | 19 |
| | Sriganti | 2 | 5 | 0 | 2 | 2 | 2 | 5 | 11 | 0 | 0 | 4 | 8 |
| Murai Batu | Aceh | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 2 | 1 | 2 |
| | Borneo | 4 | 12 | 5 | 8 | 6 | 3 | 11 | 10 | 5 | 3 | 10 | 3 |
| | Medan | 1 | 0 | 0 | 1 | 0 | 0 | 2 | 4 | 0 | 0 | 0 | 0 |
| Perkutut | Bangkok | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| | Jawa | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 | 20 |
| Total | | 152 | 166 | 150 | 151 | 154 | 145 | 164 | 164 | 148 | 152 | 158 | 154 |
| Akurasi (%) | | 44,706 | 48,824 | 44,118 | 44,412 | 45,294 | 42,647 | 48,235 | 48,24 | 43,53 | 44,706 | 46,471 | 45,294 |
| Error (%) | | 55,294 | 51,176 | 55,882 | 55,588 | 54,706 | 57,353 | 51,765 | 51,76 | 56,47 | 55,294 | 53,529 | 54,706 |

The accuracy of the results obtained by using the MFCC that is equal to 48.824% occurred in the epoch amount 150, Gridtop topology and function of the Box distance.

# 5 CLOSING

## 5.1 Conclusion

Based on the results of the implementation and testing obtained in the research and have been adjusted for purposes of research, the following conclusion:

1. Application of the MFCC (Mel-Frequency Cepstrum Coefficients) method and SOM neural network classification was implemented successfully in the identification of birds chirping and showed an accuracy of 85.625% for the same case of bird species, while in the case of birds, each of which has a subtype with similar characteristics can detect the sound with an accuracy rate of 48.824%, so we can conclude the system using MFCC as feature extraction method can better identify the type birds for bird species and sample cases similar to previous research conducted such perkutut birds, kenari birds, cucak ijo birds and cucak rawa birds.

## 5.2 Suggestions

The following suggestions can be made to the development of the research that has been done:

1. Develop an application identification of bird sounds in real time.
2. Increase the number of sound samples training and testing in order obtain a more optimal result accuracy.
3. Using the hardware that has better specs in order get maximum results in using MFCC (Mel-Frequency Cepstrum Coefficients) and Artificial Neural Networks Self Organizing Maps.

4. To further investigations may use methods of feature extraction and other classification methods that can distinguish sound of each sub-species of birds.

## BIBLIOGRAPHY

[1] R. Annisa, B. Hidayat dan I. Wijayanto, "KLASIFIKASI JENIS BURUNG BERDASARKAN SUARA KICAU BURUNG MENGGUNAKAN WAVELET PACKET DECOMPOSITION DAN JARINGAN SYARAF TIRUAN SELF ORGANIZING MAP," 2012.

[2] F. AN, "PENGENALAN PENGUCAP TAK BERGANTUNG TEKS DENGAN METODE VEKTOR QUANTIZATION (VQ) MELALUI EKSTRAKSI LINEAR PREDICTIVE CODING," 2004.

[3] A. D. Andriana, "PERANGKAT LUNAK UNTUK MEMBUKA APLIKASI PADA KOMPUTER DENGAN PERINTAH SUARA MENGGUNAKAN METODE MEL FREQUENCY CEPSTRUM COEFFICIENTS," *Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)*, vol. 2, no. 1, pp. 21-26, 2013.

[4] D. Putra dan A. Resmawan, "Verifikasi Biometrika Suara Menggunakan Metode MFCC dan DTW," *Lontar Komputer*, vol. 2, no. 1, pp. 8-21, 2011.

[5] I. Sommerville, Software Engineering (9th Edition), USA: Pearson Education, 2011.

[6] S. Kusumadewi, Artificial Intelligent (Teknik dan Aplikasinya), Yogyakarta: Graha Ilmu, 2003.

[7] A. Hermawan, Jaringan Saraf Tiruan Teori dan Aplikasi, Yogyakarta: Andi, 2006.

[8] R. A.S dan M. Shalahuddin, Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek, Bandung: Informatika, 2013.

[9] G. A. Away, The Shortcut of MATLAB Programming, Bandung: Informatika Bandung, 2006.

[10] R. A. S dan M. Shalahuddin, Rekayasa Perangkat Lunak Terstruktur dan Berorientasi Objek, Bandung: Informatika, 2013, p. 271.