A Comparison of Backpropagation and LVQ : a case study of lung sound recognition

Fadhilah Syafria[1], Agus Buono[2] and Bib Paruhum Silalahi[3]
[1,2]Department of Computer Science and [3]Departement of Mathematics
Faculty of Mathematics and Natural Sciences Bogor Agricultural University
Bogor, Indonesia
Email: [1]fadhilahsyafria2102@gmail.com, [2]pudesha@yahoo.co.id , [3]bibparuhum1@yahoo.com

Abstract—One way to evaluate the state of the lungs is by listening to breath sounds using stethoscope. This technique is known as auscultation. This technique is fairly simple and inexpensive, but it has some disadvantage. They are the results of subjective analysis, human hearing is less sensitive to low frequency, environmental noise and pattern of lung sounds that almost similar. Because of these factors, misdiagnosis can occur if procedure of auscultation is not done properly. In this research, will be made a model of lung sound recognition with neural network approach. Artificial neural network method used is Backpropagation (BP) and learning Vector Quantization (LVQ). Comparison of these two methods performed to determine and recommend algorithms which provide better recognition accuracy of speech recognition in the case of lung sounds. In addition to the above two methods, the method of Mel Frequency Cepstrum Coefficient (MFCC) is also used as method of feature extraction. The results show the accuracy of using Backpropagation is 93.17%, while the value of using the LVQ is 86.88%. It can be concluded that the introduction of lung sounds using Backpropagation method gives better performance compared to the LVQ method for speech recognition cases of lung sounds.

I. INTRODUCTION

Lung sounds are part of the respiratory sounds. Respiratory sounds include the sound of the mouth and trachea while lung sounds occur around the chest. Sounds of lungs occurs due to air turbulence when the air enters the respiratory tract during breathing. This turbulence occurs because the air flow from the air duct that is wider to narrower airways or vice versa.

In general, the sound of the lung is divided into two, sounds is the sound of the lungs that are not detectable respiratory abnormality, whereas abnormal lung sounds are sounds lung disorder.

One way to evaluate the state of the lungs is by listening to breath sounds using a stethoscope. This technique is known as auscultation techniques. Auscultation technique is a basic technique that is used by physicians to evaluate breath sounds. This technique is fairly simple and inexpensive, but it has the disadvantage that the results of a subjective analysis [2]. The results of the analysis of breath sounds using auscultation technique relies on the ability, auditory, and experience of the doctor who performed the analysis. In addition, human hearing is less sensitive to low frequency sound is also a problem in this technique, because the sound of breathing occupies a low enough frequency. The next problem in auscultation techniques are the problem of environmental noise and the pattern of sounds that almost similar between one type of breath sounds and the other. Because of these factors, misdiagnosis can occur if the procedure of auscultation is not done properly.

Based on the above problems, we interested in conducting research in performing speech recognition of normal lung and lung sounds were detected interference (abnormal). Sounds of lungs resulting in some cases of the disease showed a specific pattern that can be recognized. This sound pattern can be taken as material for diagnosis [3]. Lung sound will be classified into four classes, namely tracheal, vesicular, crackle and wheeze. Tracheal and vesicular sound indicates nomal lung, while the crackle and wheeze sound indicates abnormalities in the lungs [4][5].

Voice recognition is an effort in order to the voice can be recognized or identified so that it can be used. A voice recognition system is a computational application that is able to identify or verify the sound is automatically. These systems need to be trained in advance to be able to recognize the voice and the voice classifies according to the class. For such purposes, an algorithm is needed to train the speech recognition system to perform more precise.

A system can learn and recognize patterns of sound

can be done with the application of artificial neural networks. Artificial neural network is a computational method that mimics the biological neural network of human that capable trained to solve problems. In general, there are several artificial neural network algorithms that can be applied to pattern recognition. This research will use two artificial neural network algorithms, they are Backpropagation (BP) and Learning Vector Quantization (LVQ). Both methods have advantages and disadvantages of each. Comparison of these two methods performed to determine and recommend algorithms which provide better recognition accuracy of speech recognition in the case of lung sounds.

In addition to using an artificial neural network algorithm in the classification process, it is no less important in voice recognition is the feature extraction. In this study, feature extraction methods are used Mel Frequency Cepstrum Coefficient (MFCC). MFCC is a feature extraction method which gives excellent results in classifying normal lung sounds and wheeze sounds [6].

II. RESEARCH METHOD

Research method conducted is illustrated in Figure 1. In the image looks that research begins with a literature review of the theories necessary for the completion of this study, such as lung sound theory, MFCC, Backpropagation, LVQ and MATLAB programming.

A. Materials Research

Lung sound data obtained from the repository of respiratory sounds on the Internet, namely Littmann Repository. The sound data consist of 32 lung sound data, which is divided into 8 tracheal sounds, 8 vesicular sounds, 8 crackle sound and 8 wheeze sound.

Tracheal sound is a sound that is audible at the base of the neck and larynx. Tracheal sound is very loud and his pitch is high (Fig. 2a), so this sound is very clear sounding compared to another normal lung. Inspiration and expiratory are relatively equal in length [7].

Vesicular sounds are normal breath sounds that are heard on the side chest and chest near the stomach. The sounds ware soft with a low pitch (Fig. 2b). Inspiration sounds much stronger than expiratory sounds. Often the expiratory process is hardly audible [7].

Crackles sounds are a short burst sound that is discontinuous (Fig. 2c), it is generally more audible sound in the process of inspiration [8]. The conditions that is cause the crackles are ARDS, asthma, bronchiectasis, chronic bronchitis, consolidation, early CHF, nterstitial lung disease dan pulmonary edema [9].
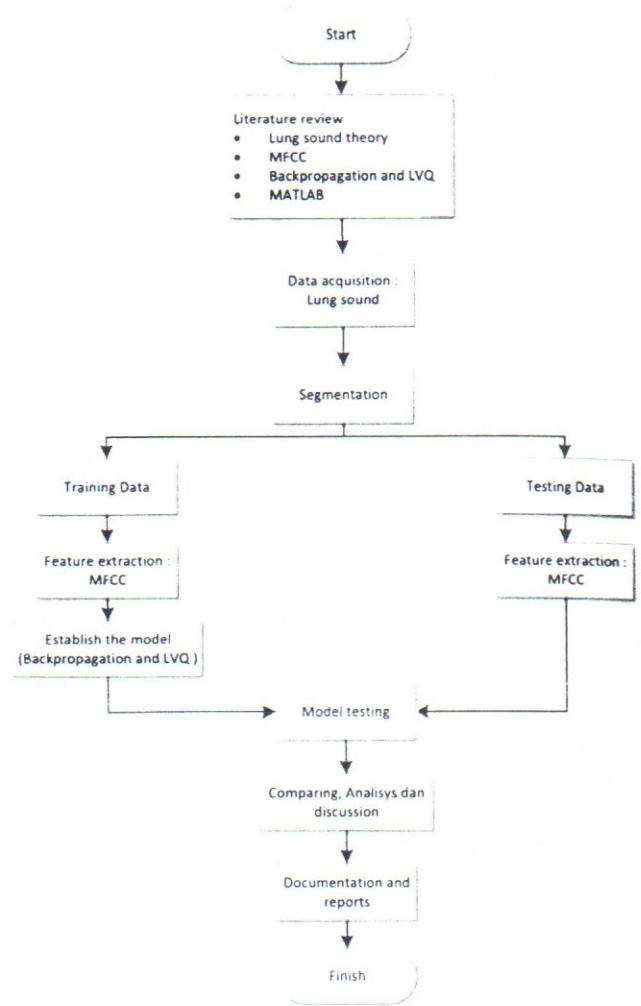

Figure 1. Research Method

Wheezes sound is a type of sound that is continuous, have a high pitch (Fig. 2d), and more often heard on the expiratory. This sound occurs when the flow of air through the narrowed airways due to secretions, foreign body or wounds that preclude [3]. The duration of the process of inspiration longer than the expiratory process. Besides that inspiring sound intensity greater than the expiratory. The conditions that is cause the wheeze are asthma, CHF, chronic bronchitis, COPD dan pulmonary edema [9].
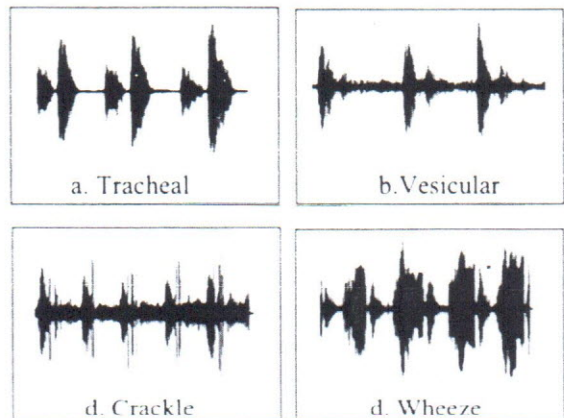

Figure 2. Type of lung sound [9]

B. Segmentation

Lung sound data that has been collected then made the sound signal segmentation process. This process is done to cut sound into several smaller frames in order to easily processed. Lung sound signal will be cut based on one respiratory cycle. Sound signal segmentation process is done by using the Audacity software. Illustration of signal segmentation process shown in Figure 3. Segmentation process produces 24 tracheal sounds, 24vesicular sounds, 24 crackle sounds and 24 wheeze sounds. So the total sounds obtained from the segmentation process is 96 sounds data.

Figure 3. Illustration of sound signal segmentation

C. Data Distribution of Training and Testing Data

In general, there are two processes that occur in the artificial neural network algorithm, namely training process and testing process. Sound data used in the learning process is called training data, while the sound data used in the testing process is called test data. Therefore, the lung sound data that has been segmented divided into two parts, namely the training data and test data. The number of lung sound data is 96 data. Data lung sounds are then defined as 80 training sounds and 16 testing sounds data.

D. MFCC Feature Extraction

MFCC feature extraction is a technique used to generate a vector which is used as an identifier [10]. The feature is the cepstral coefficient, cepstral coefficients that are used still considering the perception of the human auditory system. MFCC technique can represent a better signal than the LPC, LPCC and the other in speech recognition [11]. This is due to the workings of the MFCC is based on the frequency difference that can be captured by the human ear so that it can represent how people receive sound signals [12]. Figure 4 is a flow diagram of MFCC.
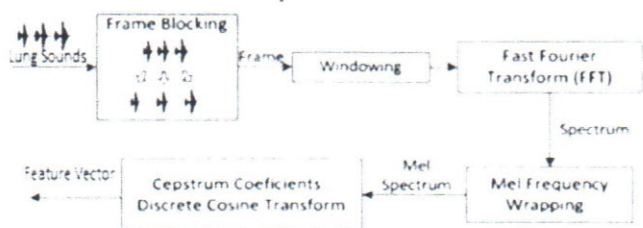

Figure 4. Diagram of MFCC

Feature extraction Using MFCC will not remove any characteristics or important information of each lung sound data. In addition, the size of the lung sound data becomes not too large. Broadly speaking, there are five stages of MFCC, ie frame blocking, windowing, fast fourier transform, mel frequency wrapping and cepstrum coefficient.

Lung sound data that has been segmented then performed 'frame blocking' process. Because the sound data has segmented using audacity software, frame blocking process will read the files that have been previously segmented. Furthermore, windowing process using Hamming Window because of simple formula. The next step is the Fast Fourier Transform (FFT). FFT process is used to convert each frame that has been generated from the previous process from the time domain into the frequency domain, so it can be more easily observed. Once the signal is converted from the time domain into the frequency domain, the next step is the mel frequency wrapping process, this process need the filter, thus will be formed M filter before wrapping process is done. Next is the Discrete Cosine Transform (DCT) to get coeffisient cepstrum. this Coefficient cepstrum that is an output of the MFCC [13].

E. Similarity Measurement (Classification)

Lung sound data that has been obtained their feature vector through feature extraction process will be classified using two methods : bacpropagation and LVQ. The results of preliminary processing of data in the form of voice samples that have been segmented and the extraction process is carried out will be the input of the two methods.

Broadly speaking voice recognition lungs out by using the backpropagation and LVQ has two main parts, namely the stages of learning patterns (training) and the stages of pattern recognition / similarity measurement (testing). Learning outcomes data collected and stored as a learning model that can later be used to measure the similarity of a sound signal entering the lungs so that the subsequent lung sound signals can be recognized. Once the learning phase is complete, the next stage is the stage of pattern recognition (similarity measurements).

1) Backpropagation

Backpropagation is a systematic method of artificial neural network using supervised learning algorithm and is typically used by the perceptron with many layers (input layer, hidden layer and output layer) to change the existing weights in the hidden layer [14]. Backpropagation is training which type controlled using weights adjustment patterns to achieve the minimum value of the error between the output of the prediction results with real output. To get a network error, the forward propagation phase (feedforward) must be done before.

In the feedforward phase, each unit receives input signals and forwards the signal to all units in the layer above it (the hidden layer). Each unit in the hidden layer summing the input signals that is receives. In this step activation function is used to calculate the output signal to be sent to all units on it (the output layer). This step done as much as the number of hidden layers. Then, for each unit of output summing the output signals. In this step activation function is also used to calculate the output signal [15].

After the feedforward phase, the next step is backforward phase. In this phase, each output unit receives a target pattern associated with the training input pattern and then calculate his error information. Then calculate the correction weights and bias correction. This step done as much as the number of hidden layers. Each hidden unit delta summing inputs (from units located in the layer above it). Multiply the delta value by the derivative of activation function to calculate error information, and then calculate the correction weights and bias correction. The final step, for each unit of output fixing weights and bias, as well as the hidden units are also fixing weights and bias [15].

2) Learning Vector Quantization (LVQ)

LVQ algorithm is a method for training competitive layers of the supervised [16]. Competitive layer will automatically learn to classify the given input vector. The approach taken is by classify the input vector based on proximity distance of the input vector to agent vector (euclidean distance method).

LVQ network structure is a two-layer neural network is composed of input layer and output layer. Input layer contains neurons as many as dimensional input, output layer contains neurons as many as the number of classes [17]. These two layers are connected by link between each neuron that has a certain weight is called the agent vector.

LVQ algorithm is a method of artificial neural network based on competition with a mechanism squared euclidean distance in picking winners agent vector to determine the category of the input vector. In LVQ network, the learning process undertaken are a supervised learning, where providing input are accompanied by the expected output information. Network always directed to determine the most appropriate output unit with the target of input vectors through shift position of agent vector. If the vectors of training data are grouped together with agent vector winner, then the agent vector is shifted come near to vector of training data by the equation :

$$W_j(new) = W_j(old) + a(X_i - W_j(old)) \quad (1)$$

If vectors the training data are not grouped together with the agent vector winner, then the agent vector shifted away from the training vector is expressed by the equation :

$$W_j(new) = W_j(old) - a(X_i - W_j(old)) \quad (2)$$

Where :

α = Learning rate
W = Weight of agent vector
X = Weight of input vector

III. RESULTS AND DISCUSSION

A. Testing Results of LVQ Method

Testing at this stage to look at the level of recognition accuracy using LVQ. Testing is done by changing the number of coefficients MFCC and the value of learning rate to see its effect on the level of recognition accuracy. The number of MFCC coefficients used are MFCC 13, MFCC 15, MFCC 20 dan MFCC 30. whereas the learning rate used are 0.1, 0.2, 0.3, 0.4 dan 0.5.

LVQ implementation is done in Matlab environment with `lvqnet` function. Learning functions used are LVQ 2.1 that on Matlab implemented with functions `learnlvq2`.

According to experiment on LVQ by varying the parameter values the learning rate and coefficients MFCC, obtained accuracy levels vary in testing with test data as presented in Table 1. There was no general pattern of the relationship between MFCC coefficient and learning rate, but both parameters affect the accuracy freely. However, the value of the highest accuracy in LVQ obtained with 20 MFCC coefficients and the of 0.2 learning rate with an accuracy 97.02% (marked in bold).

TABLE I
ACCURACY OF TESTING IN LVQ

MFCC Coefficients	Learning Rate				
	0.1	0.2	0.3	0.4	0.5
13	90.18	93.45	94.08	82.44	76.19
15	88.99	93.45	91.37	81.85	82.14
20	89.59	**97.02**	76.19	84.82	86.01
30	92.26	86.01	84.82	90.47	77.08

Overall, testing with LVQ can achieve the highest accuracy of 100%. However, in a minority of experiments obtained an accuracy of 0%, ie no data is correctly identified. This is due to the performance of the LVQ training up and down during the training process iteration or not towards convergent. For example, it was determined that the maximum error of 0.01, however when the training process until the maximum iteration never been obtained the error rate. ven the error becomes higher as the iteration goes though the previous iteration error of low value

B. Testing Results of Backpropagation Method

Testing at this stage to look at the level of recognition accuracy using Backpropagation. Testing is done by changing the number of coefficients MFCC and the value of learning rate to see its effect on the

level of recognition accuracy. The number of MFCC coefficients used are MFCC 13, MFCC 15, MFCC 20 dan MFCC 30. whereas the learning rate used are 0.1, 0.2, 0.3, 0.4 dan 0.5.

Backpropagation implementation is done in Matlab environment with newff. Activation function used is sigmoid (logsig) in the hidden layer and a linear function (purelin) in the output layer. Training algorithm used is gradient descent (traingd).

According to experiment on Backpropagation by varying the parameter values the learning rate and coefficients MFCC, obtained accuracy levels vary in testing with test data as presented in Table 2. Just like the LVQ, not obtained a clear general pattern of the relationship between MFCC coefficient and learning rate. However, from the table it can be seen that the accuracy tends to increase with increasing learning rate with an increase in the fluctuating. In some cases, the value of learning rate that is too high can make a lower accuracy.

the value of the highest accuracy in Backpropagation obtained with 13 MFCC coefficients and the of 0.4 learning rate with an accuracy 97.32% (Table 2). Backpropagation method also can achieve the highest accuracy of 100% as in LVQ, but in some experiments, the accuracy is quite low on a particular class can even reach 0%.

TABLE II
ACCURACY OF TESTING IN BACKPROPAGATION

MFCC Coefficients	Learning Rate				
	0.1	0.2	0.3	0.4	0.5
13	91.96	96.43	94.63	**97.32**	96.73
15	89.88	94.94	94.05	95.83	96.43
20	87.5	91.96	93.75	95.54	94.35
30	80.36	89.58	95.24	91.96	94.94

C. *Effect of Number of MFCC Coefficient Against Accuracy*

From the entire experiment, the average accuracy obtained at each value of the coefficient is shown in Figure 5. From the graph it can be seen that there is a tendency that the higher the coefficient MFCC makes lower accuracy in LVQ and Backpropagation, although the decrease was not significant. Accuracy at MFCC coefficient 13 tend to be higher than the accuracy of the MFCC coefficient greater . This is due to the higher coefficient of MFCC impact on the dimensions of a larger data. Large data dimensions make the generalization capability of ANN is lower so the accuracy is decreased.

D. *Effect of Learning Rate Against Accuracy*

Influence of the learning rate throughout the experiment to the average accuracy is shown in Figure 6. From the graph, in LVQ there is a tendency that

highest accuracy occur on the learning rate is not too low or too high. If the learning rate is too high, accuracy will be decrease because of the LVQ that do the change of network weights with the nature of the competition, ie, increase or decrease the weight of the network, so that the learning rate is too large causing drastic changes at each iteration so that the weight of the network becomes unstable. However, if the learning rate is too low, the learning process will be long.
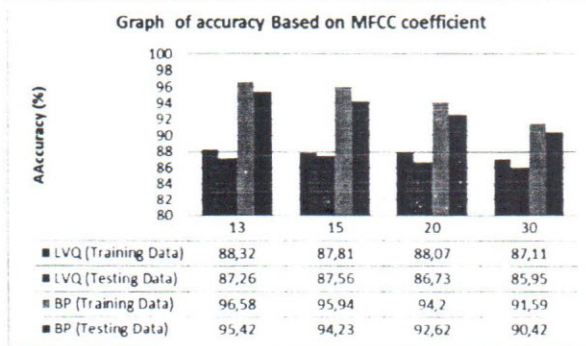

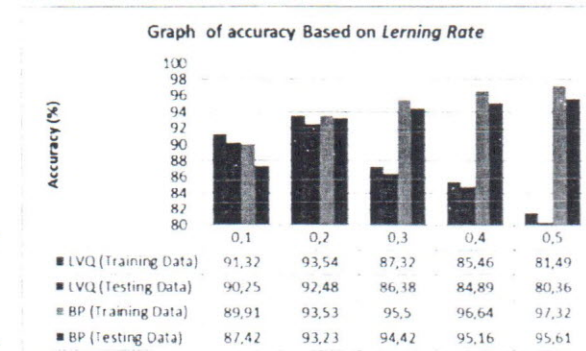Figure 5 Graph of accuracy Based on MFCC coefficient

Figure 6 Graph of accuracy Based on *Learning Rate*

In contrast, in Backpropagation, increased learning rate tends to increase the accuracy. This is due to the algorithm Backpropagation just changing the network weights by summing up the weights (without the nature of the competition) so that the network weight remains stable.

E. *Introductory Rate of Each Lung Sounds*

Average of accuracy of the entire experiment based on the type of lung sounds is shown in Figure 7. From the graph it can be seen that the type of tracheal sound has the highest accuracy, meaning that most can be distinguished from other types of sound. It is, as mentioned previously (Figure 7) that a tracheal sound signals has the form the most different than the other three types of sound. Meanwhile, the sound of vesicular type, crackle, and wheeze has little resemblance so that its accuracy is lower, meaning that the sound of vesiculer is sometimes identified as a crackle type or wheeze and vice versa.
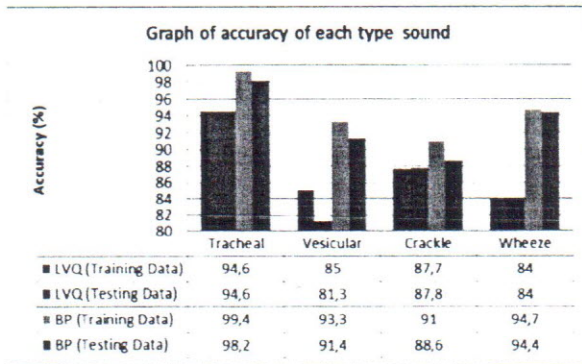
Figure 7 Graph of accuracy in Each Type Sound

	Tracheal	Vesicular	Crackle	Wheeze
■ LVQ (Training Data)	94,6	85	87,7	84
■ LVQ (Testing Data)	94,6	81,3	87,8	84
■ BP (Training Data)	99,4	93,3	91	94,7
■ BP (Testing Data)	98,2	91,4	88,6	94,4

F. Comparison of Accuracy Rate Backpropagation and LVQ Method

Comparison of the accuracy of LVQ and Backpropagation method in identifying lung sounds is presented in Figure 8. Overall, backpropagation method gives better accuracy results than LVQ method, with an overall average accuracy of 94.58% for training data and 93.17% for test data, while LVQ only offer results of 87.83% for training data and 86.88% for test data. Testing the training data provide greater accuracy results because the data is also used in the training process, while the test data were not included during training.

Comparison of the accuracy of LVQ method and BP

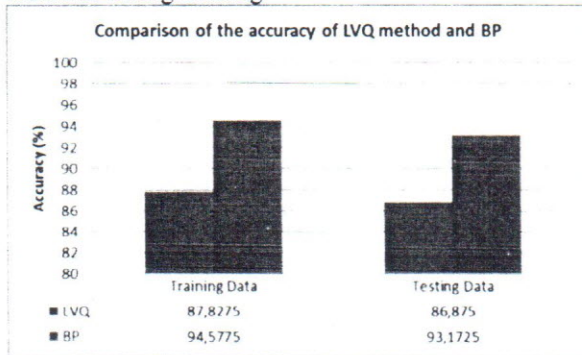|	Training Data	Testing Data
■ LVQ	87,8275	86,875
■ BP	94,5775	93,1725

Figure 8 Comparison of Accuracy graphs of LVQ and Backpropagation

Although LVQ method gives lower accuracy, but the method is simpler than computing aspects so that training time is faster than the Backpropagation method. This is consistent with studies conducted Hawickhorst *et al.* (1995) that Backpropagation gives better accuracy results than LVQ, but higher computational complexity [18].

IV. CONCLUSION

From the research that has been done, it can be concluded that it has made a model of lung sound recognition with MFCC feature extraction. Classification with LVQ and Backpropagation method can be applied for the identification of lung sounds, by adjust parameters MFCC coefficient and learning rate that maximizes the accuracy. Backpropagation method has better accuracy than LVQ with an average accuracy on the test data was 93.17%, while 86.88%

of LVQ. In addition, tracheal sound can be recognized better than the lung sound types vesicular, Crackle, and wheeze.

REFERENCES

[1] A. Cohen, D. Landsberg, "Analysis and automatic classification of breath sounds", *Biomedical Engineering, IEEE Transactions on*, (9), 585-590, 1984.

[2] H. Kiyokawa, MDM. Greenberg, K. Shirota, H. Pasterkamp, "Auditory Detection of Simulated Crackels in Breath Sounds", CHEST, 119(6) : 1886-1892, 2013

[3] A. Rizal, MD. Samudra, I. Iwut, V. Suryani, "Pengenalan Suara Paru Menggunakan Spektogram dan K-Mean Clustering", *Proceeding SITIA2010*, Februari 2010

[4] AA. Abaza, JB. Day, JS. Reynolds, AM. Mahmoud, WT. Goldsmith, WG. McKinney, EL. Petsonk, DG. Frazer, "Classification of voluntary cough sound and airflow patterns for detecting abnormal pulmonary function", Cough. 5(8), 2009.

[5] A. Gurung, CG. Scrafford, JM. Tielsch, OS. Levine, W. Checkley W, "Computerized lung sound analysis as diagnostic aid for the detection of abnormal lung sounds: a systematic review and meta-analysis", Respir Med, 105(9): 1396-1403, 2011.

[6] M. Bahoura, "Pattern Recognition Methods Applied to Respiratory Sounds Classification into Normal and Wheeze Classes", Computers and Biology and Medicine, 39(9):824-843, 2009.

[7] A. Rizal, L. Anggraeni, V. Suryani, "Pengenalan Suara Paru-Paru Normal Menggunakan LPC dan Jaringan Syaraf Tiruan Back-Propagation', *Proceeding EECCIS2006*, Mei 2006.

[8] T. Katila, P. Piirila, K. Kallio, E. Paajanen, T. Rosqvist, AR. Sovijarvi, "Original waveform of lung sound crackles: a case study of the effect of high-pass filtration" *Journal of Applied Physiology*, 71(6), 2173-2177, 1991.

[9] MZ. Ramadhan, "Perancangan Sistem Instrumentasi untuk Identifikasi dan Analisis Suara Paru-Paru Menggunakan DSP TMS320C6416T [Essay]", Depok[ID]: Universitas Indonesia, 2012.

[10] K. Patel, RK. Prasad, "Speech Recognition and Verification Using MFCC & VQ", Interntional Journal, 2013.

[11] A. Buono, "Representasi nilai hos dan model MFCC sebagai ekstraksi ciri pada sistem indentifikasi pembicara di lingkungan ber-noisemenggunakan HMM [dissertation]". Depok: Program Studi Ilmu Komputer, Universitas Indonesia, 2009.

[12] L. Muda, M. Begam, I. Elamvazuthi, "Voice Recognition Algorithms Using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques", Journal of Computing, Volume 2, Issue 3, 2010.

[13] M. Slaney, "Auditory Toolbox, Interval Research Corporation", Tech. Rep, 10, 1998.

[14] F. Suhandi, "Prediksi Harga Saham dengan Pendekatan Artificial Neural Network menggunakan Algoritma Backpropagation", 2009.

[15] L. Fausett, "Fundamentals of neural networks architectures, algorithms and applications", *Florida Latitude of Technology*, 1994.

[16] S. Kusumadewi, S. Hartati, "Integrasi Sistem Fuzzy dan Jaringan Syaraf", Yogyakarta. Graha Ilmu, 2010.

[17] L. Rahadianti, "Pengembangan Algoritma Pembelajaran Berbasis Dimensi serta Komparasinya terhadao Pembelajaran Berbasiskan Vektor pada *Fuzzy-Neuro Learning VectorQuantization* untuk Pengenalan Citra Wajah Frontal [Essay]", Depok [ID] : Fasilkom Universitas Indonesia, 2009.

[18] BA. Hawickhorst, SA. Zahorian, R. Rajagopal, "A Comparison of Three Neural Network Architectures for Automatic Speech Recognition" Di dalam: Dagli CH, *et al* editor. *Proceedings of the Artificial Neural Networks in Engineering, ANNIE'95*, New York: ASME Press, hlm 221-6.

[19]