Dog Cough Sound Classification Using Artificial Neural Network and the Selected Relevant Features from Discrete Wavelet Transform

Panchana Kakabutr, Kullanan Sae Chen, Viewpaka Wangvisavawit,
Praisan Padungweang and Olarn Rojanapornpun
School of Information Technology
King Mongkut's University of Technology Thonburi, Bangkok, Thailand

Abstract—**Coughing is one of the important signs of several diseases in dogs. There are two types of dog cough: dry cough and productive cough. The latter is most often associated with an infectious condition. It is difficult to differentiate between the two types even by experienced practitioners. In this paper, an automatic cough sound classification using neural network is introduced. A discrete wavelet transform is employed to decompose the cough sound into low frequency and high frequency components. The statistical features of these components are used as the sound features. The discrimination power of these features in classification are evaluated. Finally, the artificial neural network is used to classify dog cough sound using a subset of discriminant features. The experimental results show that classifying dog cough sounds needs only one fourth of all features and an average accuracy as high as 90% is achievable.**

Keywords-classification; discrete wavelet transform; dog cough sound; feature selection; neural network

I. INTRODUCTION

Despite of a dog owner's love, many terminally ill dogs are not treated in a timely manner. Heart disease is the most common cause of sudden and unexpected death [1]. Incidentally, coughing is associated with many respiratory and cardiovascular system diseases, including heart disease. It could be classified as dry or productive. Dry cough is an indicator of allergic bronchitis, respiratory system diseases, heartworms diseases and circulatory system diseases [2], [3]. Whereas a productive cough is an indicator of infectious tracheobronchitis, tonsillitis and pneumonia [4], [5]. Unfortunately, in some cases, it is difficult to identify whether the cough is dry or productive. Improvement in the ability to identify the class of dog coughs could lead to better diagnosis and treatment.

In the past decade, there have been some studies on animal health analysis such as classification of wasting diseases in pigs [6], classification of healthy and mastitis in Murrah buffaloes [7], machine intelligence for dog health pre-diagnosis [8] and sound identification of tracheal collapse and laryngeal paralysis in dogs [9]. Currently, there are no studies on analysis and classification of dog coughs.

In 2008, Knockikova et al. published their study on the analysis of voluntary cough sound in patients with respiratory diseases [10]. Göğüş et al. later reported their study on classification of asthmatic breath sounds in 2015 [11]. These studies use wavelet transform which is a signal processing technique widely used on non-stationary signals [12]. Another popular signal processing technique is Fourier transform [13]. However, it is ineffective on non-stationary signals such as a cough sound [14].

There are many classification techniques, for example, Support Vector Machines (SVM), K-Nearest neighbors (KNN) and Artificial Neural Network (ANN). However, in sound classification, speaker identification and voice recognition applications, ANN is more popularly used [11], [15]–[17]. It provides very good results in the task of speech/music/noise classifications [18].

Since dog coughs and human coughs are different, the objective of this study is to explore the feasibility of using discrete wavelet transform and neural networks to classify dog cough sound. Furthermore, the suitable cough features and neural networks configurations are also studied.

This paper is organized into five sections. Section II covers the overview of related techniques. Section III presents the materials and methods used in this study. The results of this study are shown in Section IV. Finally, Section V concludes with our findings and suggestions.

II. RELATED TECHNIQUES

Three techniques are used in this study to classify dog cough sound. They are feature extraction based on discrete wavelet transform, feature selection based on discriminant analysis, and artificial neural network.

A. Feature Extraction based on Discrete Wavelet Transform

Discrete wavelet transform (DWT) is a signal processing technique to transform raw signal in time domain to time-frequency domain. It is used to extract relevant information from a signal and is suitable for analysis of non-stationary

978-1-4673-9077-4/17/$31.00 ©2017 IEEE

signals [12]. The signal is analyzed in different scales by decomposing it into sub-bands, comprised of low frequency band (approximation coefficients) and high frequency band (detail coefficients). This process is counted as one level of decomposition. The low frequency band can be used for further decomposition.

Fig. 1 shows three levels of decomposition where A is approximation coefficients and D is detail coefficients. In this example, wavelet coefficients of D1–D3 and A3 can then be used in the extraction process of signal features.

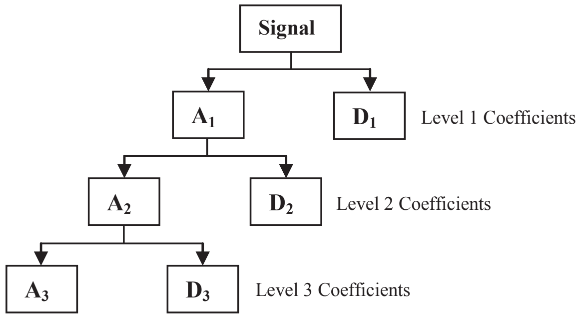

Figure 1. Discrete wavelet transform with three levels of decomposition

Since the dimensions of these coefficients are still too high for classification, basic statistical analysis is often used to extract the signal features further. The common statistical analysis used with discrete wavelet transform are mean, mode, median, minimum, maximum, and variance. These are applied to the amplitudes, energy, and frequencies of the sub-bands coefficients [19]–[21].

B. Feature Selection based on Discriminant Analysis

Creating a classification model with high-dimensional data is time consuming. It might also converge to a local solution due to the huge amount of search space. Feature selection can find a subset of relevant features to improve the performance of the model building [22]. On the other hand, selecting the truly relevant features can simplify the model which improve both time and performance.

In this paper, Fisher score was chosen as a technique to reduce the dimensions of the data. Fisher score is a simple and effective technique to select the most relevant features from the data set [23]. It calculates the score of each feature based on a ratio of between-class variance and within-class variance. The Fisher score of the i^{th} feature can be computed as follows [24]:

$$F(i) = \frac{\left(\overline{x}_i^p - \overline{x}_i\right)^2 + \left(\overline{x}_i^q - \overline{x}_i\right)^2}{\frac{1}{n_p - 1}\sum_{k=1}^{n_p}\left(x_{k,i}^p - \overline{x}_i^p\right)^2 + \frac{1}{n_q - 1}\sum_{k=1}^{n_q}\left(x_{k,i}^q - \overline{x}_i^q\right)^2} \tag{1}$$

Where $\overline{x}_i$, $\overline{x}_i^p$, $\overline{x}_i^q$ are the average of the i^{th} feature of the whole, dry cough, and productive cough data sets, respectively; $x_{k,i}^p$ is the i^{th} feature of the k^{th} dry cough instance, and $x_{k,i}^q$ is the i^{th} feature of the k^{th} productive cough instance. The numerator indicates the discrimination between the dry cough and productive cough sets, and the denominator indicates the one within each of the two sets.

The features with high Fisher scores are important for classification. These features separate the two classes of cough more than others. Therefore, the Fisher score of each feature can be used as a criterion for feature selection.

C. Artificial Neural Network

Artificial neural network (ANN) is widely used in classification. The basic concept of ANN is the simulation of biological neural network like the human brain structure with an interconnected cluster of neurons. Fig. 2 shows the neural network architecture that commonly has three layers: input layer, hidden layer and output layer. This is also known as Multi-layer perceptron (MLP) [11] which is a widely used feed-forward only neural network topology. The $\sum$ sign refers to the weighted summation of the signals from the previous layer. The letter F refers to the activation function (transfer function) for transferring the appropriate output value to be an input signal of the next layer.

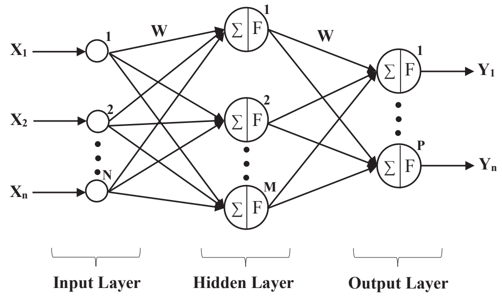

Figure 2. A feed-forward artificial neural network architecture

The amount of nodes in the input layer coincides to the dimensions of features and additional one bias node. In supervised learning, a target is specified as output in the model. The backpropagation algorithm, i.e., Levenberg-Marquardt algorithm [25], is used for updating weight and bias values. It is designed to approach second-order training speed using an approximated Hessian matrix. For each layer, the weighted summation in each node will be passed through the activation function before forwarding to the next layer. There are many activation functions known as logistic, linear and hyperbolic tangent transfer function [26] as shown in equation (2)–equation (4) respectively.

$$f(z) = \frac{1}{(1+e^{-z})} \tag{2}$$

$$f(z) = z \tag{3}$$

$$f(z) = \frac{(e^z - e^{-z})}{(e^z + e^{-z})} \tag{4}$$

Where z refers to the weighted summation function, $\sum$ function in the Fig. 2. Logistic function transforms $\sum$ into the range 0 and 1. Next, the linear function does not change $\sum$. The last one, the hyperbolic tangent function transforms $\sum$ into the range -1 and 1.

III. MATERIALS AND METHODS

In this study, dog cough samples were collected and transformed to extract the features of dog cough. The dimensions of the extracted features were then reduced by feature selection. Finally, performance of neural network models with various configurations were studied.

A. Acquisition of Dog Cough Samples

Dog cough sounds were recorded in an examination room at the Faculty of Veterinary Medicine, Kasetsart University. The types of coughs were specified by a veterinarian based on the diagnostics of the dogs. These cough sounds were collected from six dogs, five of which had dry coughs. Since there were very few samples collected on-site, additional samples were collected from the internet. Two veterinarians were asked to identify the types of coughs independently. Only the samples with total agreement were used in this study. The selected clips were recorded from eleven dogs, seven were identified as dry coughs. These recordings were cut into 165 individual cough sound segments: 115 dry cough segments and 50 productive cough segments.

B. Discrete Wavelet Transform and Feature Extraction Method

In this study, Daubechies 3 was selected as a mother wavelet. It is found that the maximum level of decomposition of each sound segment is different and each has the value of at least 10 levels. Thus each sound segment was decomposed to level 10 and the DWT coefficients are obtained from D1–D10 of high frequency bands and A10 of low frequency band. Eight statistical features are extracted from each DWT coefficients: mean, variance, maximum energy, minimum energy, maximum amplitude, minimum amplitude, average frequency, and half point of the energy. Note that in this study, variance and mean energy have the same values and thus only one of them is selected.

C. Feature Selection using Fisher Score

Fisher score computes the score of each feature based on the idea of discriminant analysis. The extracted features are evaluated and ranked according to their discrimination score in a non-decreasing order. In particular, the most relevant feature subset of size m consists of the m top ranked features. Four subsets of extracted features are selected based on their ranking. Each subset consists of 11, 22, and 44 top ranked features respectively and the last one is the full 88 extracted features.

D. Classification using Artificial Neural Network

Different configurations that affect performance of neural network models such as the number of nodes in the hidden layer, activation functions in the hidden layer and the output layer are considered. MATLAB-R2016B with neural network toolbox is used for the creation of neural network models and performance evaluations. Firstly, the number of nodes in the input layer are selected based on the number of features in each feature subset. Secondly, neural network structures are created with various numbers of neurons in the hidden layer which are 5, 10, 20, 40, 80 and 160 nodes. Next, two types of dog cough sounds are encoded with binary vectors, dry cough is encoded as [1 0] and productive cough is encoded as [0 1]. Therefore, there are two nodes in the output layer. Finally, the three activation functions are tested in the hidden layer and two activation functions (the three without linear function) are also tested in the output layer. These settings make a combination of six different configurations for each structure. Therefore, there are 36 different configurations of neural network models for each selected features subset.

A neural network model can overfit to the training data performing poor predictive performance on the unseen data. In order to avoid overfitting, it is necessary to use additional techniques for model validation. We use the 5-fold cross validation strategy for dividing data into three subsets: training, validation and test subset. The training subset is used for updating weight and bias values. The validation subset is used for stopping learning when the network begins to overfit or prediction error increases. Test subset is used as the unseen data for model selection. The data of each class are randomly and equally partitioned into five subsets (folds). Each fold contains both data subset from dry cough and productive cough. We perform the following steps 10 times for each features subset.

1) Use three subsets (60%) as training data, one as validation (20%) and the rest as test data (20%).
2) Train the 36 models with the training data using the performance of predicting the validation data as a stopping criterion. The weight and bias values are randomly initialized.
3) Measure the classification accuracy of each model on the test data.
4) Repeat steps 1–3 until each of five subsets is used exactly once as the test set.

The average accuracy results from 50 runs of each model-feature subset are used for appropriate model-feature subset selection.

IV. RESULTS

Features of 165 cough sound segments were extracted by DWT and statistical calculations. The Fisher score of each statistical feature was then calculated and ranked in non-decreasing order. The 88 ranked features were divided into four nonoverlapping parts, where the first part consists of 11 top ranked features, the second part consists of the next 11 ranked features, the third part consists of the next 22 ranked features and the fourth part consists of the last 44 features. Table I shows the results of dividing 88 extracted features into four parts. The subset of features S_i is the union of part 1 to part i. For instance, S_3 consists of 44 features from part 1 to part 3. The total Fisher scores of each selected features subset is accounted to 45%, 67%, 90%, and 100% of the total Fisher scores of the extracted features, respectively.

TABLE I. FISHER SCORE RANKING OF EXTRACTED FEATURES

Statistical Feature	**Coefficients**										
	A 10	***D 1***	***D 2***	***D 3***	***D 4***	***D 5***	***D 6***	***D 7***	***D 8***	***D 9***	***D 10***
Mean	4	4	4	3	4	4	1	4	4	4	4
Variance	2	4	2	2	1	4	1	3	2	3	3
Max Energy	1	4	4	4	3	4	1	4	2	3	3
Min Energy	4	3	4	4	4	3	4	3	4	4	4
Max Amp.	2	4	4	4	3	4	1	4	2	3	3
Min Amp.	1	4	4	3	3	4	1	3	3	2	3
Avg. freq.	4	2	3	4	3	2	1	4	2	4	4
HaPo.	3	4	1	4	1	4	4	4	4	3	4

The results from feature selection have shown that the variance of amplitude is one of the most relevant statistical features in dog cough sound classification and detail coefficients of decompositions at level 6 and 8 contain many relevant features. On the other hand, minimum energy does not provide relevant features and detail coefficients of decompositions at level 7 and 10 can be discarded.

The training and validation data sets are used in the process of neural network model building. There are some experiments that the model can only predict a single class of dog cough in the validation data set, even though the model does not overfit to the training set. It is possible that the training got stuck at an improper local solution. We tried to avoid this predicament by setting an acceptable accuracy threshold of the validation performance. Therefore, the validation data set has not only been used for avoiding overfitting but also avoiding some improper local solution. However, to ensure fair comparison, the test set is still held out as the unseen data for model evaluation. The average accuracy of predicting the test set are shown in Fig. 3 and Fig. 4.

Fig. 3 compares the average accuracy results of neural network models using different activation functions. It can be seen that the performances of neural networks using logistic function as the activation function in the hidden nodes and hyperbolic tangent function as the one in the output nodes (logistic-tanh models for short) outperform the performances of the other activation function configurations. The average accuracy results of the logistic-tanh models using a different

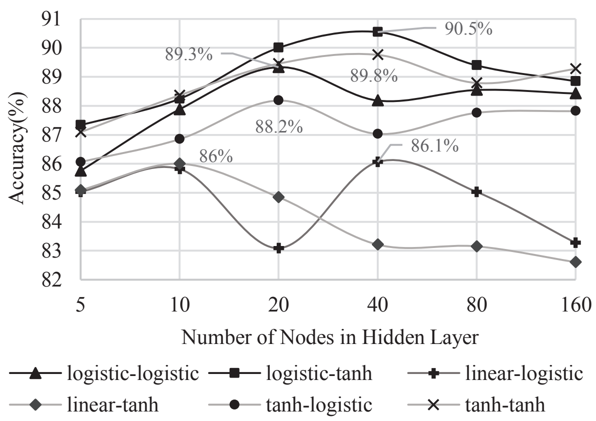

Figure 3. Classification accuracy of ANN on 22 features with different activation functions in the hidden/output layer

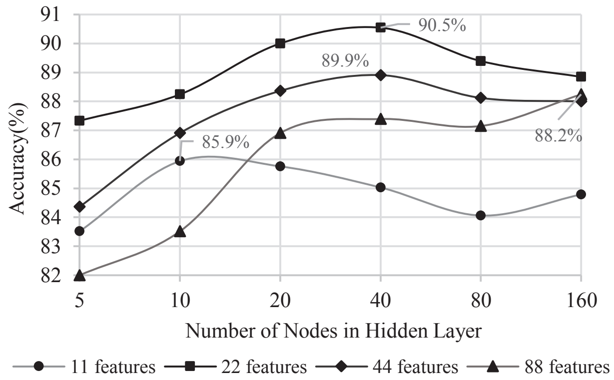

Figure 4. Classification accuracy of logistic-tanh ANN models with different number of features

number of features are also shown in Fig. 4. The results from the 22 features input data outperform the results from the others. Overall, the results of the logistic-tanh models consisting of 40 nodes in the hidden layer performs the best with 90.5% accuracy.

V. DISCUSSION AND CONCLUSION

The purpose of this study has been to evaluate whether an artificial neural network with the selected relevant features from discrete wavelet transform can be effectively applied to classify dog cough sound and demonstrate the promising model that provides the highest accuracy.

Dog cough sound classification was carried out by using discrete wavelet transform to transform a raw signal from time domain into time frequency domain and extract relevant information from the signal. At the beginning, the appropriate level of DWT decomposition was not known. We can conclude from the results from feature selection and classification that 8 levels of decomposition is sufficient to classify dog cough sound, with level 6 and level 8 as the most significant sub-bands.

Based on our samples and experiments, the best performing neural network model consists of 40 nodes in the hidden layer and uses logistic function as the activation function in the hidden nodes and hyperbolic tangent function in the output nodes. The model with 22 features consistently outperforms other models with different data dimensions This means that there are more than 11 relevant features and many of our extracted features are irrelevant.

This study shows that cough sound classification using artificial neural network and the selected relevant features from discrete wavelet transform has a potential to be a facilitative tool for preliminary diagnosis of a dog's illness. However, further work should be conducted to include a larger data set of cough sound with a variety of dog breeds to enhance the model's ability to handle broad breeds in real-life scenarios. Moreover, redundancy of the selected features may be studied to reduce the number of features further.

ACKNOWLEDGMENT

We would like to thank the Faculty of Veterinary Medicine at Kasetsart University for their assistance in data collection. We would also like to thank Bangmod Veterinary Hospital and Ban Suan Thon Veterinary Clinic for their advice in this research.

REFERENCES

[1] T. F. Olsen and A. L. Allen, "Causes of sudden and unexpected death in dogs: A 10-year retrospective study," *The Canadian Veterinary J.*, vol 41, pp 873-875, Nov. 2000.

[2] E. Ward, VCA Animal Hospital. *Heartworm disease in dogs* [Online]. Available: http://www.vcahospitals.com/main/pet-health-information/article/animal-health/heartworm-disease-in-dogs/773.

[3] PetWave (2015, July 16). *Treating asthma (allergic bronchitis)* [Online]. Available: http://www.petwave.com/Dogs/Health/Asthma/Treatment.aspx

[4] K. Becker. (2014, Jan 20). *Coughing: This Seemingly harmless symptom could foretell a deadly disease* [Online]. Available: http://healthypets.mercola.com/sites/healthypets/archive/2014/01/20/dog-coughing.aspx

[5] E. Ward, VCA Animal Hospital. *Tonsillitis in dogs* [Online]. Available: http://www.vcahospitals.com/main/pet-health-information/article/animal-health/tonsillitis-in-dogs/726

[6] Y. Chung, S. Oh, J. Lee, D. Park, H. H. Chang and S. Kim, "Automatic detection and recognition of pig wasting diseases using sound data in audio surveillance systems," *Sensors*, vol. 13, no. 10, pp. 12929-12942, Sep. 2013.

[7] I. Panchala, I. K. Sawhneya, A. K. Sharmab, A. K. Dangc, "Classification of healthy and mastitis Murrah buffaloes by application of neural network models using yield and milk quality parameters," *Comput. and Electron. in Agriculture*, vol. 127, pp. 242-248, Sep. 2016.

[8] K. B. Kim, D. H. Song and Y. W. Woo, "Machine intelligence can guide pet dog health pre-diagnosis for casual owner: A Neural Network Approach," *Int. J. of Bio-Science and Bio-Technology*, vol 6, no.2, pp. 83-90, 2014.

[9] S. C. Yeon, H. C. Lee and H. H. Chang and H. J. Lee, "Sound signature for identification of tracheal collapse and laryngeal paralysis in dogs," *J. of Veterinary Medical Sci.*, vol. 67, no. 1, pp. 91-95, 2005.

[10] J. Knocikova, J. Korpas and M. Vrabec and M. Javorka, "Wavelet analysis of voluntary cough sound in patients with respiratory diseases," *J. of Physiology and Pharmacology*, vol. 59, pp. 331-340, 2008.

[11] F. Z. Göğüş, B. Karlık, and G. Harman, "Classification of asthmatic breath sounds by using wavelet transforms and neural networks," *Int. J. of Signal Process. Syst.*, vol. 3, no. 2, pp. 106-111, Dec. 2015.

[12] O. Rioul and M. Vetterli, "Wavelets and signal processing," *IEEE Signal Processing Mag.*, vol. 8, pp. 14-38, Aug. 2002.

[13] H. Shatkay, "The Fourier Transform - a Primer," Dept. Comput. Sci., Brown Univ., Tech. Rep. CS-95-37, Nov. 1995.

[14] M. Sifuzzaman, M. R. Islam and M. Z. Ali, "Application of wavelet transform and its advantages compared to fourier transform," *J. Physical Sci.*, vol. 13, pp. 121-134, 2009.

[15] K. J. Piczak, "Environmental sound classification with convolutional neural networks," *IEEE International Workshop on Machine Learning for Signal Processing (MLSP)*, Boston, USA, 2015.

[16] P. Sangpeth and S. Nitsuwat, "Speaker Identification using Artificial Neural Network," *The 8th National Conference on Computing and Information Technology,* Pattaya City, Thailand, 2012, pp. 966-972.

[17] K. Rongsawad and W. Chatwiriya, "Classification of Thai Expressive Speech for Requesting and Ordering using Neural Networks," *The Tenth National Conference on Computing and Information Technology*, Phuket, Thailand, 2014.

[18] L. Alvarez, C. Llerena and E. Alexandre, "Application of neural networks to speech/music/noise classification in digital hearing aids," *GAVTASC'11 Proceedings of the 11th WSEAS International Conference on Signal Proceedings, computational geometry and artificial vision, and Proceedings of the 11th WSEAS International Conferenceon Systems theory and scientific computation*, Florence, Italy, 2011, pp. 97-102.

[19] M. M. Butt, U. Akram and S. A. Khan, "Classifying normal sinus rhythm and cardiac arrhythmias in ECG signals using statistical features in temporal domain," *9th Asia Modelling Symposium (AMS),* Kuala Lumpur, 2015, pp. 28-31.

[20] K. M. Priya, S. Kavitha and B. Bharathi, "Brain tumor types and grades classification based on statistical feature set using support vector machine," *10th International Conference on Intelligent Systems and Control (ISCO)*, Coimbatore, India, 2016.

[21] I. Romero and L. Serrano, "ECG frequency domain features extraction: A new characteristic for arrhythmias classification," *23rd Annual International Conference of the IEEE on Engineering in Medicine and Biology Society*, Istanbul, Turkey, 2001, pp. 2006-2008.

[22] Y. He, K. Fataliyev and L. Wang, "Feature selection for stock market analysis," *International Conference on Neural Information Processing*, Daegu, 2013, pp 737-744.

[23] K. Arunasakthi, L. KamatchiPriya and A. Askerunisa, "Fisher score dimensionality reduction for SVM classification," *International Conference on Innovations in Engineering and Technology*, Madurai, 2014, pp. 1900-1904.

[24] Y. W. Chen and C. J. Lin, "Combining SVMs with various feature selection strategies," *Studies in Fuzziness and Soft Computing*, vol 207, pp 315-324.

[25] Hagan, M. T., and M. Menhaj, "Training feed-forward networks with the Marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, pp. 989-993, 1994.

[26] Vogl, T.P., J.K. Mangis, A.K. Rigler, W. T. Zink, and D. L. Alkon, "Accelerating the convergence of the backpropagation method," *Biological Cybernetics*, vol. 59, pp. 257-263, 1988.