2015 IEEE 35th International Conference on Electronics and Nanotechnology (ELNANO)

Right Whale Detection Using Artificial Neural Network and Principal Component Analysis

Kostiantyn Pylypenko
Department of Acoustics and Acoustoelectronics,
National Technical University of Ukraine "KPI",
Kyiv, Ukraine
k.pylypenko@hotmail.com

***Abstract* — In this paper applying PCA algorithm in order to preprocess data in North whale detection problem was proposed. It is shown that reducing number of features obtained from a spectrogram can prevent overfitting and increase performance of the classifier. Proposed classifier was based on a tree layer Neural Network with 50, 100 and 400 input neurons. Proposed technique achieved an F_1 score of 90%.**

Keywords — Artificial Neural Network, Principal Component Analysis, spectrogram, whale detection

I. INTRODUCTION

The North Atlantic right whale is in danger of extinction [1]. The main reason of extinction of whales is high human activity in the areas of their migration.

Monitoring for the occurrences of whales is the one way to reduce whale mortality. Whale species produce many different sounds and detecting these sounds is one of the most popular method of detecting right whale.

The frequency structure of whale sounds may vary widely according to the time of day, season, age, etc. But, analysis of whale calls shows that most of them have duration of 0.3-1.5 s and occupy the frequency band from 30 to 250 Hz [2]-[5]. Fig. 1 shows 12 seconds of recorded sound with whale calls and its spectrogram.

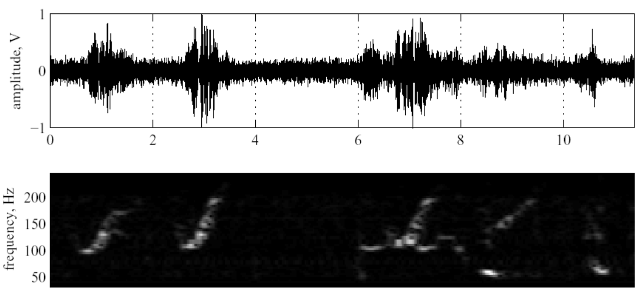

Fig. 1. Example of audio recording of North Atlantic right whale calls and its spectrogram

One of the most popular approach to detect acoustic signals is using energy detector [6], [7]. However, the basic problem with detection whale calls by using energy detector is the lack of a priori information about the signals and high level of the noise, which means that incoming signals should be properly denoised [8]. Moreover, sounds from other species could be detected as whale calls.

There are several another approaches to this task [9], [10], some of them are based on neural networks [11] and some use a tree classifier as the train model [12].

In this paper Neural Network approach was improved by using Principal Component Analysis to preprocess the data.

II. DATASET

The dataset was provided by the Marinexplore and Cornell University Whale Detection team. The dataset consist of 28000 clips for training and 19000 for test. Each sound clip is 2-second long signal with a sample rate of 2kHz. It is should be noticed that signal-to-noise ratio (SNR) for different signals was from 1 to 20 dB and samples contained various amount of acoustic clutter, including sounds from other species.

III. METHODS

A. Feature extraction

In order to obtain the spectrograms, a short time Fourier transform with window size of 256~ms (Hamming window, 512 samples, 85% overlap) was applied to each sample. After obtaining a spectrogram, the lower and upper frequencies of spectrogram were hard-limited at 30 and 250 Hz, respectively. As a result of previous steps each example in the dataset became a 56 pixel by 56 pixel grayscale image of spectrogram. The 56 by 56 grid of pixels was "unrolled" into a 3136-dimensional vector.

There are 16 spectrograms from the dataset with whale calls are presented in the Fig. 2, Fig. 3 shows 16 spectrograms without whale calls.

As one can see the dimension of the dataset is huge and it has to be preprocessed in order to reduce its dimensionality. Reducing dimensionality has computational benefits and can also reduce complexity of the hypothesis class in classification procedure and help to avoid overfitting.

978-1-4673-6534-5/15/$31.00 ©2015 IEEE

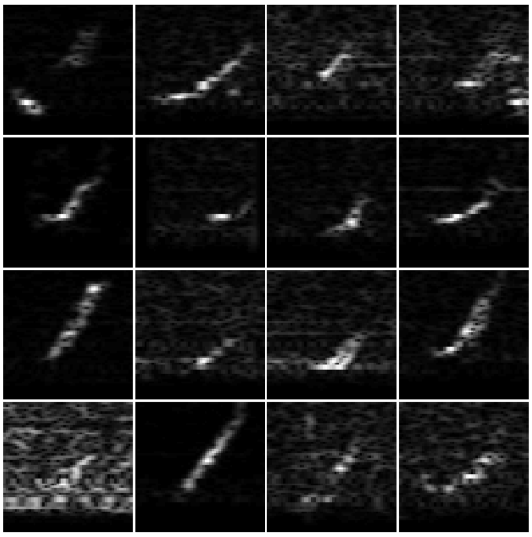

Fig. 2. Examples from the dataset with right whale call

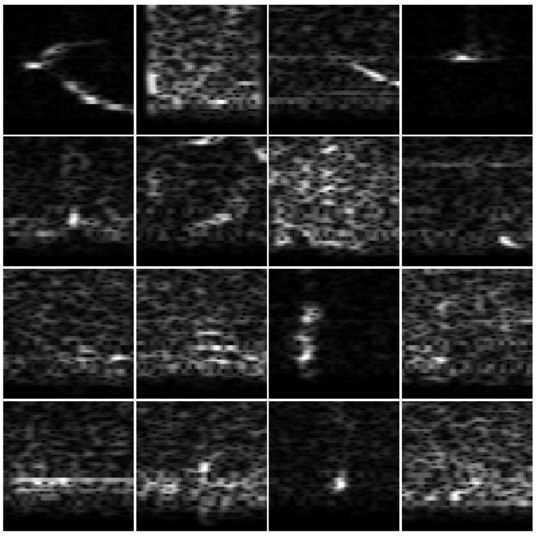

Fig. 3. Examples from the dataset without right whale call

B. Principal Component Analysis (PCA)

PCA is a variable reduction procedure. It is useful when obtained data on a number of variables has some redundancy in those variables. In this case, redundancy means that some of the variables are correlated with one another, possibly because they are measuring the same construct. Because of this redundancy, it should be possible to reduce the observed variables into a smaller number of principal components (artificial variables) that will account for most of the variance in the observed variables.

Before using PCA, it is important to normalize data and scale each dimension so that variables are in the same range. First, one should compute the covariance matrix of the data

$$\Sigma = \frac{1}{m} X^T X ,$$

where X is the data matrix with examples in rows, and m is the number of examples.

After computing the covariance matrix the eigenvectors u of Σ should be founded. Top k eigenvectors $u_1, u_2, \ldots, u_k$ are called k principal components of the data. This components form a new orthogonal basis that data can be projected into by computing corresponding vector

$$y_i = \begin{bmatrix} u_1^T x_i \\ u_2^T x_i \\ \vdots \\ u_k^T x_i \end{bmatrix} ,$$

where x_i is the i -th row of the data matrix X .

Projection original data set into k -dimensional subspace allows reducing number of dimensions of presented data from 3136 to k .

Fig. 4 shows the first 25 principal components $u_1, u_2, \ldots, u_{25}$ that were rolled up into the 56 by 56 grid of pixels and describe the largest variation.

The first component extracted in a principal component analysis accounts for a maximal amount of total variance in the observed variables. Under typical conditions, this means that the first component is correlated with at least some of the observed variables. It may be correlated with many. The second component extracted has two important characteristics. First, this component accounts for a maximal amount of variance in the data set that was not accounted for by the first component. Again under typical conditions, this means that the second component is correlated with some of the observed variables that did not display strong correlations with component 1.

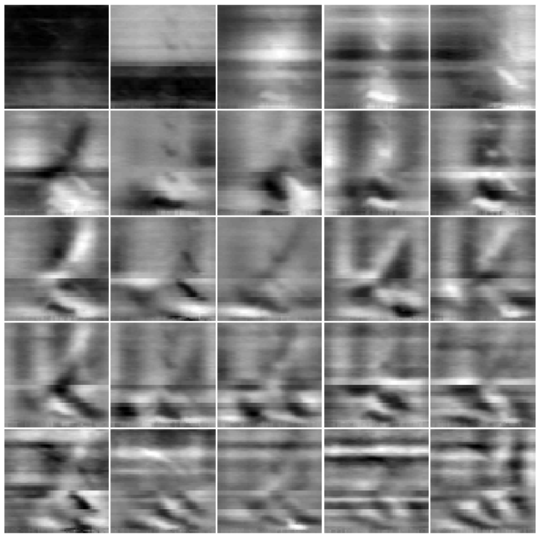

Fig. 4. Principal components on the dataset

The second characteristic of the second component is that it is uncorrelated with the first component. So, if one computes the correlation between components 1 and 2, that correlation would be zero.

The remaining components that are extracted in the analysis display the same two characteristics: each component accounts for a maximal amount of variance in the observed variables that was not accounted for by the preceding components, and is uncorrelated with all of the preceding components.

In order to understand what is lost in the dimension reduction, the original data were recovered using only the projected dataset. Fig. 5 and Fig. 6 show data reconstructed from 400 principal components and Fig. 7 and Fig. 8 show spectrograms obtained from top 50 principal components. As can be observed from reconstructions, the general structure and appearance of the whale calls are kept while some details are lost.

In case of using only 50 principal components it is a remarkable reduction (more than 60 times) in the dataset size which means that learning algorithm can be significantly speeded up.

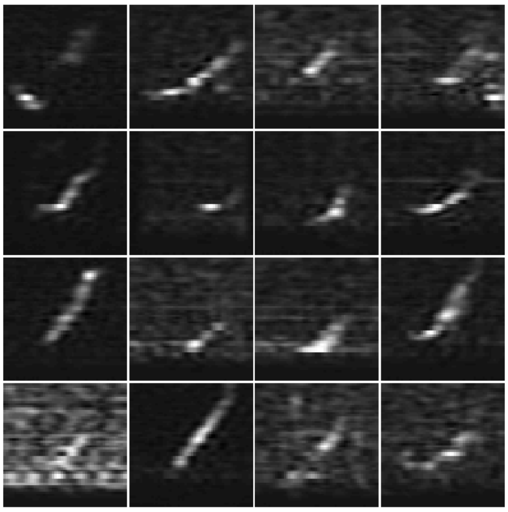

Fig. 5. Examples with right whale call reconstructed from top 400 principal components

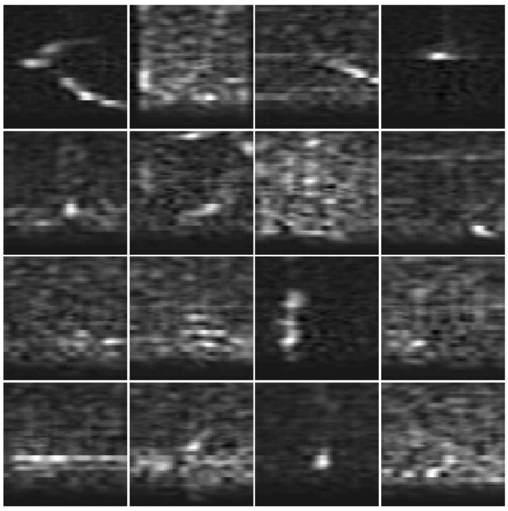

Fig. 6. Examples without right whale call reconstructed from top 400 principal components

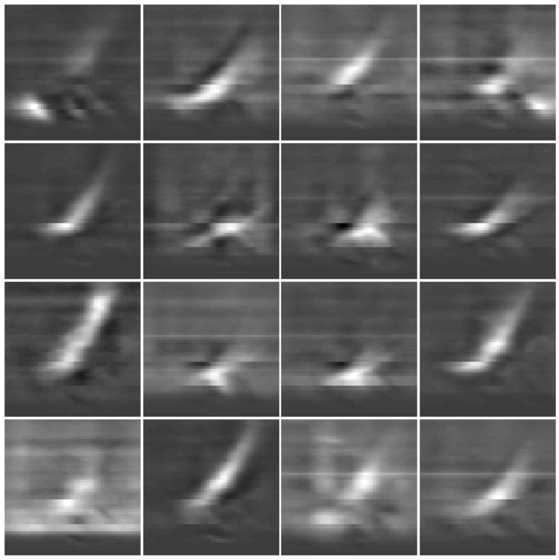

Fig. 7. Examples with right whale call reconstructed from top 50 principal components

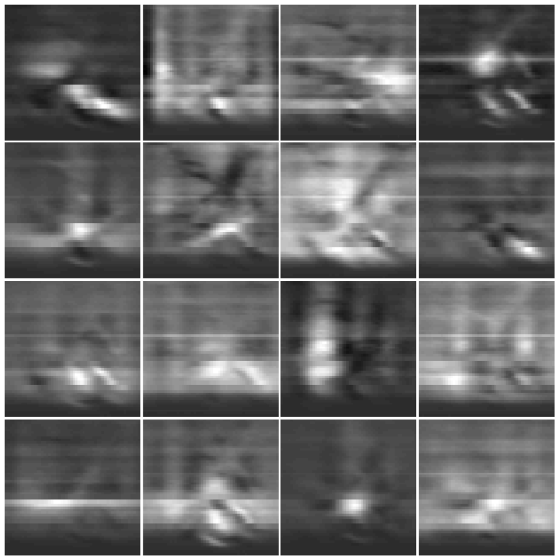

Fig. 8. Examples without right whale call reconstructed from top 50 principal components

C. *Artificial Neural Network (ANN)*

The number of types of ANNs and their uses is very high. The first neural model was created by McCulloch and Pitts [13] and since then there have been developed hundreds of different models considered as ANNs.

The ANN is very useful and effective model for signal classification in general as well as bioacoustics signals [14]. ANNs can use a wide range of features as input, such as those from images, spectrograms, or waveforms. The number of features can be huge, especially in the text classification problems. This flexibility allows the application of ANNs to different detection conditions and various types of signals.

For purpose of presented research a feed-forward ANN was used. It had 3 layers - an input layer, a hidden layer and an output layer. The number of units in the input layer was equal to a number of features, the hidden layer had 1000 units and the output layer had two units.

The ANN was trained on the train dataset using back-propagation algorithm [15].

D. Classification

Described above ANN was applied for three different cases: with 400 features as a feature set, with 100 features as the feature set, and with only 50 features as a features set. For each case in order to reduce number of features from 3136 to 400, 100 and 50 PCA algorithm was used.

Fig. 9 demonstrate the Receiver Operating Curve (ROC) of the proposed method evaluated for different numbers of features.

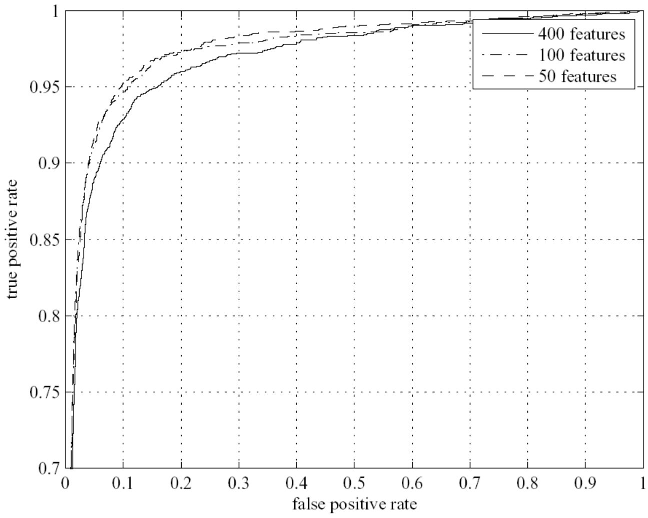

Fig. 9. ROC for classification

The performance of the proposed classifier on the test dataset is shown in Table I.

CONCLUSIONS

It can be seen from the Table I that performance of the classifier increases with reduction of classification features. In case of 50 features the classifier produced an F_1 score of 90% and an area under the ROC curve (AUC) of 98%. In comparison, F_1 score achieved in work [12], which used a tree classifier as the train model, was close to 85%.

In this paper it is shown that by using PCA algorithm it is possible to achieve high performance with low number of classification features.

TABLE I. THE PERFORMANCE OF THE TRAINED CLASSIFIER ON THE TEST DATASET

Number of features	F_1	AUC
400	0,86	0,97
100	0,88	0,97
50	0,90	0,98

REFERENCES

[1] S. D. Kraus et al., "North atlantic right whales in crisis," Science, Vol. 309(5734), pp. 561-562, 2005.

[2] J. N. Matthews, S. Brown, D. Gillespie, M. Johnson, R. McLanaghan, A. Moscrop, D. Nowacek, R. Leaper, T. Lewis, and P. Tyack, "Vocalization rates of the North Atlantic right whale (Eubalaena glacialis)," J. Cet. Res. Man., Vol. 3, pp. 271-282.

[3] S. E. Parks and P. L. Tyack, "Sound production by North Atlantic right whales (Eubalaena glacialis) in surface active groups," J. Acoust. Soc. Amer., Vol. 117, pp. 3297-3306, 2005.

[4] D. Gillespie, "Detection and classification of right whale calls using an "edge" detector operating on a smoothed spectrogram," Can. Acoust., Vol. 32, No. 2, pp. 39-47, 2004.

[5] L. M. Munger, S. M. Wiggins, S. E. Moore, and J. A. Hildebrand, "North Pacific right whale (Eubalaena japonica) seasonal and diel calling patterns from long term acoustic recordings in the southeastern Bering Sea, 2000-2006," Marine Mammal Science, Vol. 24(4), pp. 795-814, 2008.

[6] S. Jarvis, N. DiMarzio, R. Morrissey, and D. Morretti, "Automated classification of beaked whales and other small odontocetes in the tongue of the ocean, bahamas," OCEANS 2006, pp. 1-6, 2006.

[7] T. Ura, R. Bahl, M. Sakata, J. Kojima, T. Fukuchi, J. Ura, and M. Yanagisawa, "Acoustic tracking of sperm whales using two sets of hydrophone array," Underwater Technology, 2004. UT '04. 2004 International Symposium on, pp. 103-107.

[8] A. Prodeus, O. Ladoshko, "On existance of optimal boundary value between early reflections and late reverberation," Proc. of IEEE 34th Int. Sc. Conf. Electronics and Nanotechnology (ELNANO), pp. 442-446, 15-18 April 2014, Kyiv, Ukraine.

[9] D. K. Mellinger and C. W. Clark, "Recognizing transient low-frequency whale sounds by spectrogram correlation," J. Acoust. Soc. Am., Vol. 107(6), pp. 3518-3529, 2000.

[10] I. R. Urazghildiiev and C. W. Clark, "Acoustic detection of North Atlantic right whale contact calls using the generalized liklihhod test," J. Acoust. Soc. Am., Vol. 120(4), pp. 1956-1963, 2006.

[11] P. J. Dugan, A. N. Rice, I. R. Urazghildiiev and C. W. Clark, "North atlantic right whale acoustic signal processing: Part i. comparison of machine learning algorithms," Applications and Technology Conference (LISAT), 2010 Long Island Systems, pp. 1-6, 2010.

[12] R. Abousleiman, G. Qu, O. Rawashdeh, "North Atlantic Right Whale Contact Call Detection," Proc. of the 1st Workshop on Machine Learning for Bioacoustic, Atlanta, 2013, pp. 33-38, 2013

[13] W. McCulloch and W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity," Bulletin of Mathematical Biophysics, Vol. 5, pp. 115-133, 1943.

[14] J. R. Potter, D. K. Mellinger and C. W. Clark, "Marine mammal call discrimination using artificial neural networks," J. Acoust. Soc. Am., Vol. 96, pp. 1255, 1994.

[15] L. Fausett, "Fundamentals of neural networks: architectures, algorithms, and applications," 1994.