SDPS-2017
Printed in the United States of America, November, 2017

SOUND ANALYSIS AND MACHINE LEARNING IN NONINVASIVE CLASSIFICATION OF NEUROLOGICAL CONDITIONS

Khaled Elzaafarany[1], Gyanendra Kumar[2], Moustafa H. Aly[3], Arie Nakhmani[1]
[1]Department of Electrical and Computer Engineering, University of Alabama at Birmingham, AL, USA
[2]Department of Neurology, Mayo Clinic, Phoenix, AZ, USA
[3]College of Engineering, Arab Academy for Science and Technology, Egypt

{khaledz@uab.edu, Kumar.Gyanendra@mayo.edu, mosaly@aast.edu, anry@uab.edu}

ABSTRACT

Machine learning and signal analysis are widely used to assist medical practice. Modern medicine is in constant search for effective noninvasive methods for diagnostics. Unfortunately, many of the developed automatic methods are prone to noise and disturbances and have high computational complexity. A transcranial Doppler (TCD) is a noninvasive and reliable device that can monitor the blood flow rate in the brain and can help neurologists to diagnose many brain problems like edema, trauma, hemorrhage, and aneurysm. The proposed algorithm is a blend of statistical and machine learning tools that are used in Big Data analysis. The algorithm's goal is monitoring the TCD signals in the real-time for detection of cerebral vasospasms, which produces enormous amounts of data. We handled the data by carefully selected time and frequency domain features which allowed designing classifiers with the desired sensitivity and specificity. In addition, the proposed convergence of digital sound analysis and medical fields could prove to be useful in the future modeling of various brain disorders.

INTRODUCTION

Nowadays, sound processing methods are well developed and widely used in our cell phones, smart watches, and computers. Though modern machine learning based algorithms could automatically transcribe music and recognize speech, those algorithms were not explored enough about their applicability in medical diagnostics. The physicians still rely on their experience and senses to diagnose various diseases instead of using automated techniques. However, sound analysis and machine learning start to take their place in medicine, and this article shows a successful attempt of using sound analysis framework in neurology.

One of the medical devices that is used for diagnostics is the Doppler apparatus, which is noninvasive, inexpensive and safe. It was first used by Satomura in 1959 and its popularity is growing since then (Wright, Gough, 1997).

The physical Doppler effect allows measuring object velocity based on the sound arriving from the object. In medical applications, blood flow velocity in human body could be measured based on the Doppler effect, but it requires sound emitter and receiver to penetrate the blood vessels and receive a variation of sound frequency due to scattering and reflection. Measurement in the brain arteries is especially challenging, because it requires advanced spectral resolution methods. That kind of measurement is provided by transcranial Doppler device (TCD). A TCD can monitor the increase and decrease of blood flow velocity, or change in blood flow resistivity. Based on TCD measurement, presently neurologists detect stenosis, edema, trauma, hemorrhage, and aneurysm, but additional disorders are researched (Li et al., 2014). Angiography and magnetic resonance could be used as alternative techniques, but they are expensive and more complicated.

TCD signal analysis is active field of research. Serhatlioglu et al. used Fast Fourier Transform (FFT) for feature extraction and used a neural network as a classifier for TCD signals (Serhatlioglu, Guler, 2003). Ozturk et al. used chaotic invariant features of TCD signal as the input of neuro-fuzzy classifier (Ozturk et al., 2008). Guler et al. used FFT and adaptive autoregressive moving average (A-ARMA) methods for a spectral analysis of TCD signals; the A-ARMA shows better spectral resolution than FFT (Guler et al., 2002). Other studies (Uguz et al., 2008), (Uguz et al., 2010) used fuzzy discrete hidden Markov model (FDHMM) and the Rocchio-based hidden Markov model (HMM) for enhancing TCD signals classifier.

As seen in many studies, integral transforms are used widely, but the techniques cannot display time information that would be sufficient for a good classification. This means that many features of interest will be lost during the processing due to poor spectrum resolution especially in non-stationary TCD signals. Therefore, we tried to use both time and frequency features for classification due to its optimal time-frequency resolution in all ranges of frequencies.

In the following section, the framework for signal processing and analysis is presented for machine learning applications in medical diagnostics. In the Results and Discussion section, we demonstrate the application of the proposed framework to detection of cerebral vasospasms in stroke patients. Finally, the Conclusion section summarizes the findings and proposes future research directions.

SIGNAL ANALYSIS FRAMEWORK

There is a direct analogy between speech and music analysis and other kind of sound analysis (Doppler). Doppler sound could be explored using developed for speech and music tools/features. There is an infinite number of ways to define features; we have selected those which work the best through a rigorous selection process. We propose to apply speech, sound, and music analysis approaches to the Doppler sound to detect and classify various brain disorders or events. The audio signal is acquired with TCD and analyzed for further classification. The framework consists of three stages:

1) Preprocessing
2) Feature Extraction
3) Classification

In the first stage, the signals are recorded, or processed directly from the audio output of the TCD. They are processed with standard filtering techniques to minimize noise. Fig. 1 shows one example of the cerebral vasospasm (bottom) signal and normal (top) TCD signal at the same time and voltage scale. It is visually obvious that those two signals are different, but it is not obvious how this difference could be quantified. In addition, in many cases normal and abnormal signals are very similar in their appearance. We have carefully selected from the features used in audio research those that might be useful in TCD analysis and beyond. In the third stage, the trained classifier arrives to the decision about the problem. The framework is providing supervised training of the classifier offline. Concerning the classifier, we used the decision tree classifier. The architecture of the proposed signal processing and classification algorithm is shown in Fig. 2.

A.TCD Signal Collection

In this study, 160 wave files of 3-15 seconds duration were recorded from TCD. Medical experts classified them as normal or cerebral vasospasm.

B. TCD-Signal Classification

We apply two essential steps: feature extraction and classifier for a condition classification.

B.1 Feature Extraction

This step is used in machine learning to compress the big dataset size into a small size and at the same time to increase the classifier efficiency. We used features from time and frequency domains. The time domain features are zero-cross rate (ZCR), energy, and energy entropy. The frequency features that were used are Mel-frequency cepstral coefficients (MFCC), Chroma coefficients, spectral centroid, spread, entropy, spectral roll-off, spectral flux and harmonic ratio.

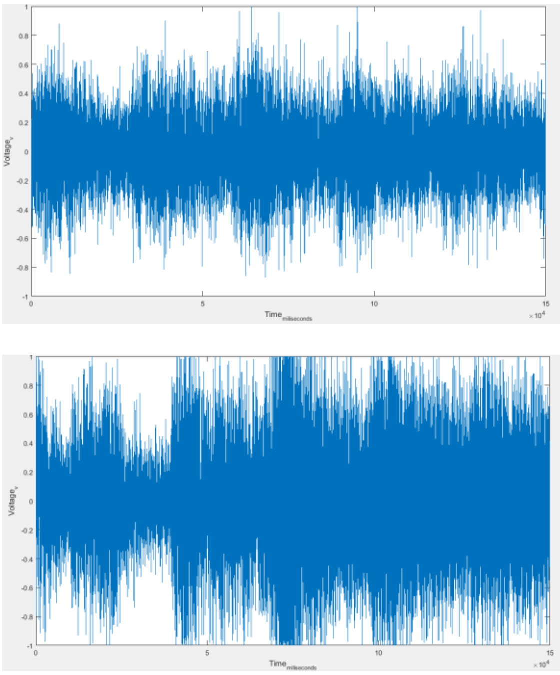

Fig. 1 TCD recording from control (top) and cerebral vasospasm (bottom).

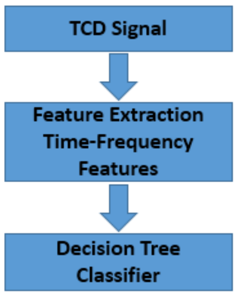

Fig. 2 The architecture of the proposed classification algorithm.

B.1.1 Zero-Crossing Rate and Harmonic Ratio

ZCR is used as a feature in the audio analysis (Tzanetakis & Cook, 2002). It counts the number of times the signal changes its sign per second. Concerning the harmonic ratio, we can assume the TCD signals as quasiperiodic to calculate the fundamental period by applying autocorrelation function (by shifting the signal and calculating the correlation of the original signal with the shifted one).

B.1.2 Energy and Energy Entropy

Energy and Energy Entropy are used in the audio content analysis (Zhang & Kuo, 2001), where the energy is calculated by dividing the signals into frames and each frame energy to samples. Energy-entropy can reflect the sudden variation of energy (Lartillot & Toiviainen, 2007). It can be computed by dividing energy of each frame to energies of sub-frames.

B.1.3 Mel-Frequency Cepstral Coefficients and Chroma

MFCC is used in speech processing field (Theodoridis & Koutroumbas, 2001). It measures the signal cepstral, where we apply the first 13 coefficients for CV detection. On the other hand, moreover, it can be used in music (Dan-Ning Jiang et al., 2002). It calculated by transforming the DFT coefficients to bins.

B.1.4 Spectral Centroid, Spread, and Entropy

Spectral Centroid, Spread, and Entropy are used for music, speech discrimination and watermarking (Scheirer & Slaney, 1997), (Kirovski & Malvar, 2003). Centroid means spectrum mass center and the spread is the second moment of the spectrum. With regards to spectral entropy, it is calculated like energy entropy, but the frame spectrum is divided into sub-bands.

B.1.5 Spectral Roll-off and Flux

Spectral roll-off can be used for differentiating between unvoiced and voiced signal (Kim et al., 2005). It is calculated by power spectral distribution. Spectral flux measures the spectral change for two sequential frames.

B.2 Classifiers

First, we apply feature scaling and mean normalization as a preprocessing step before the classifier to standardize the features values (Frohlich et al., 2003). We tried various classifying methods like a decision tree, K-nearest neighbors, support vector machine, and logistic regression. We found that decision tree gives the highest detection rate.

RESULTS AND DISCUSSION

In this study, we used the machine learning for classifying TCD signals and the obtained results are compared to a manual expert classification. A Supervised learning approach was used to train the classifier. We can enhance the efficiency of the algorithm by increasing the number of training information and the number of extracted features. The goal of supervised machine learning is to extract the model that makes an estimation based on evidence in the known input data and known output. Then, this model can predict the future output data. It is complicated to extract the correct model which is based on trial and error technique, for example, when the model has too many parameters to train, the training data will generate a sensitive model that will model minor variation which can be noise. On the other, hand too simplistic model will have low classification accuracy. Therefore, the right algorithm is a tradeoff between the training data, the number of features, accuracy, model speed, and complexity. The following schematic in Fig. 3 helps to overcome some of the machine learning challenges.

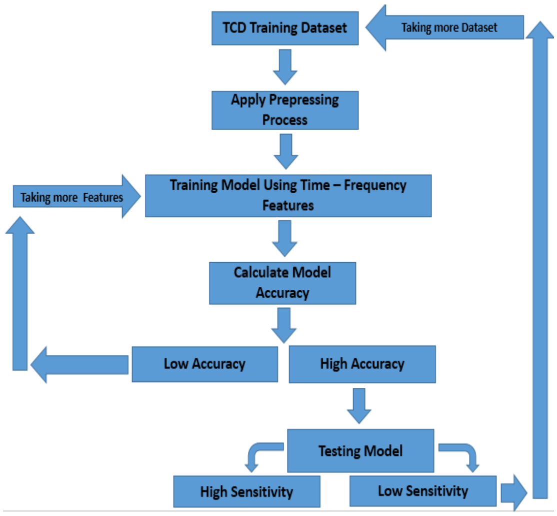

Fig. 3 Machine learning framework.

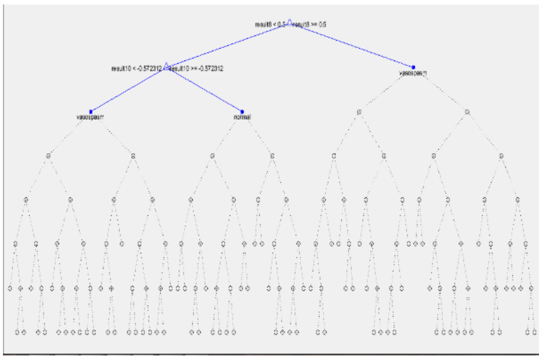

Fig. 4 Two out of 34 levels are shown in decision tree.

First, use preprocessing processes like data scaling and normalization to standardize the values, after that use a certain number of time-frequency features, and then apply the classifier. If the model accuracy after training the dataset is low, so we should increase the number of features. When we get a high detection accuracy, then, we should start testing the extracted model by using testing datasets. If the sensitivity of the model is low, so we should take more training dataset in the beginning. This loop should be repeated until we reach high accuracy and sensitivity model for the detection.

In the example of cerebral vasospasm detection, we used 10 time-frequency features to extract a high sensitivity model with using a decision tree classifier. The decision tree contains branches and leaves, where the decision rule can be made by an if conditional statement. Figure. 4 shows an example of two levels out of 34 levels of our decision tree classifier.

After extracting high accuracy and high sensitivity model, we tested the model to classify the cerebral vasospasm using real-time proceeding from the recorded TCD signals. On the tested signals the sensitivity and specificity were over 78%, which is an improvement over state-of-the-art results (Kumar et al. 2016), (Kumar, Elzaafrani & Nakhmani, 2017).

CONCLUSION

We designed and applied machine learning framework based on audio and music features extraction to TCD signals for automatic diagnosis of cerebral vasospasms without manual assistance. Time and frequency features were extracted from the signal, and they were used as input to the classifier. Our diagnostic system detects cerebral vasospasms with high sensitivity and specificity, which can help in medical practice. In the future research, we could use the wavelets for expanding the feature set to enhance accuracy. Moreover, one could try to use the classifier decision output and design embedded real-time system for alarming or for treatment purposes.

ACKNOWLEDGEMENT

K. Elzaafarany thanks, Arab Academy of Science and Technology for supporting his research.

REFERENCES

Frohlich, H., Chapelle, O., & Scholkopf, B. (2003). *Feature selection for support vector machines by means of genetic algorithm*, Proceedings of 15th IEEE International Conference on Tools with Artificial Intelligence, PP:142-148.

Guler, I., Hardalac, F. & Kaymaz, M. (2002). Comparison of FFT and adaptive ARMA methods in transcranial Doppler signals recorded from the cerebral vessels. *Computers in Biology and Medicine*, 32, PP:445–453.

Jiang, D-N., Lu, L., Zhang, H.M., & Cai, L-H. (2002). *Music type classification by spectral contrast feature*, IEEE International Conference on Multimedia and Expo.

Kim, H-G., Moreau, N. & Sikora, T. (2005). MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval. (John Wiley & Sons).

Kirovski, D. & Malvar, H.S. (2003). Spread-spectrum watermarking of audio signals. *IEEE Transactions on Signal Processing*, 51(4), PP:1020-1033.

Kumar, G., Elzaafrani, K., & Nakhmani, A. (2017). *Machine Learning Approach to Automate Detection of Cerebral Vasospasm Using Transcranial Doppler Monitoring*. American Academy of Neurology.

Kumar, G., Shahripour, R.B. & Harrigan, M.R. (2016). Vasospasm on transcranial Doppler is predictive of delayed cerebral ischemia in aneurysmal subarachnoid hemorrhage: a systematic review and meta-analysis. *Journal of neurosurgery, 124*, PP: 1257-1264.

Lartillot, O. & Toiviainen, (2007). *A Matlab Toolbox for Musical Feature Extraction from Audio*, Proceedings of 10th International Conference on Digital Audio Effects (DAFx-07) .

Li, M., Huang, et al., (2014). An analysis of cerebral blood flow from middle cerebral arteries during cognitive tasks via functional transcranial Doppler recordings. *Neuroscience research, 84*, PP: 19-26.

Ozturk, A., Arslan, A. & Hardalac, F. (2008). Comparison of neuro-fuzzy systems for classification of transcranial Doppler signals with their chaotic invariant measures, *Expert Systems with Applications*, 34, PP:1044–1055.

Scheirer, E., & Slaney, M. (1997). *Construction and Evaluation of a Robust Multifeature Speech / Music Discriminator.* IEEE International Conference on Acoustics, Speech, and Signal Processing, 2, PP:1331-1334.

Serhatlioglu, S., Hardalac, F. & Guler, I. (2003). Classification of transcranial Doppler signals using artificial neural network. *Journal of Medical Systems,* 27 (2), PP:205–214.

Theodoridis, S. & Koutroumbas, K. (2001). Machine Learning and Its Applications. *Springer*, PP:169-195.

Tzanetakis, G. & Cook, P. (2002). Musical genre classification of audio signals. *IEEE Transactions on Speech and Audio Processing,* 10, PP:293-302.

Uguz, H. et al. (2008). A biomedical system based on fuzzy discrete hidden Markov model for the diagnosis of the brain diseases. *Expert Systems with Applications*, 35 (3), PP:1104–1114.

Uguz, H. & Arslan, A. (2010), A new approach based on discrete hidden Markov model using Rocchio algorithm for the diagnosis of the brain diseases. *Digital Signal Processing*, 20(3), PP:923–934.

Wright, I., Gough, N. & Rakebrandt, F. (1997). Neural network analysis of Doppler ultrasound blood flow signals: a pilot study. *Ultrasound in Medicine and Biology*, 23(5), PP:683–690.

Zhang, T. & Kuo, C. (2001). Audio content analysis for online audiovisual data segmentation and classification. *IEEE Transactions on Speech and Audio Processing*, 9, PP:441-457.