Judgement on Shunt Sounds from Vascular Access using YOLO Deep Learning Model

Kyosuke Fujiwara
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Takayuki Yamamoto
Graduate School of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Lindsey Tate
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

Kazuya Kibune
Department of clinical engineering section, Tokatsu dialysis hospital & clinic
865-2 Hinokuchi, Matsudo-shi, Chiba, 271-0067, Japan

Hiroki Tamura
Faculty of Engineering, University of Miyazaki,
1-1, Gakuen Kibanadai-Nishi, Miyazaki, 889-2192, Japan

E-mail: hi18041@student.miyazaki-u.ac.jp, hi16046@student.miyazaki-u.ac.jp, teitorinzeirini.c4@cc.miyazaki-u.ac.jp, kmtssk@yahoo.co.jp, htamura@cc.miyazaki-u.ac.jp

Abstract

In order to explore more cost effective and accessible options for screening AV fistulas used as accesses for hemodialysis, we sought to link the results of echocardiograms to audio recorded from their respective shunts. We used machine learning techniques aimed at correctly classifying the audio data into the three classification categories derived from echocardiogram results. In this paper, we compare the discrimination rate of YOLOv2tiny alone to that of the combination of YOLOv2tiny and multivariate analysis. We conclude that bump wavelet analysis and linear multivariate analysis are the most suitable for audio data from hemodialysis shunts.

Keywords: Shunt sounds, Wavelet transform, YOLOv2tiny, Deep learning, Support vector machine

1. Introduction

An arteriovenous (AV) fistula, hereafter referred to as a shunt, is a blood vessel that connects a vein directly to an artery in order to allow a large volume of blood to pass for hemodialysis. The sound of a shunt is the sound made at its junction. Shunts carry a high risk of stenosis or blood clots, which are normally screened for using an echocardiogram (aka, "echo"). However, performing an echo requires oversight by a physician and the use of expensive equipment, which may not always be accessible to patients. In order to explore more cost effective and accessible options for screening, we sought to link the results of echocardiograms to audio recorded from their respective shunts. We used machine learning techniques aimed at correctly classifying the audio data into the classification categories derived from the echo results. Audio was recorded from shunts both before and after stress. We used wavelet transforms to image shunt sounds before and after stress, and we trained YOLOv2tiny to image the sounds. In this study, we found that YOLOv2tiny alone did not improve the discrimination rate. Therefore, we performed multivariate analysis using the multivariate calculated in the judgment of YOLOv2tiny. In this paper, we compare the discrimination rate of the result of YOLOv2tiny alone to the result of the combination of YOLOv2tiny and multivariate analysis, and we explain how to achieve a higher discrimination rate.

2. Proposed method

When examining a shunt, pre-stress measurements (such as peripheral vascular resistance, PI) and post-stress measurements are calculated by echography, and the values are comprehensively taken into account for diagnosis. I hypothesized that by linking the pre- and post-stress characteristics of the shunt sound with the diagnostic results, it would be possible to diagnose the shunt using only audio data. First, the audio data was converted into images using wavelet transform, and YOLOv2tiny was trained with the transform images. Next, we applied the trained YOLOv2tiny to all available data. The discrimination rates before and after stress were calculated for three echocardiogram classifications: normal, gray (uncertain), and abnormal. We used these six discrimination rates to compare analysis methods. In order to further improve discrimination rate, we used multiple regression analysis, decision tree, random forest, and SVM as learning methods.

2.1. *First analysis method*

Using only YOLOv2tiny, the discrimination rate is calculated using a model that is trained on pre-stress and post-stress together (Fig.1).

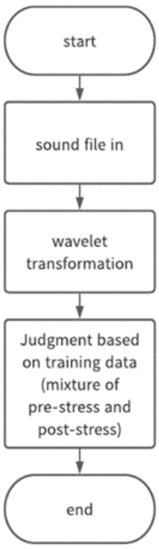


Fig.1 .Flowchart of the first analysis method

2.2. *Second analysis method*

Using multiple regression analysis, the discrimination rate is calculated using a model trained separately for pre-stress and post-stress (Fig.2).

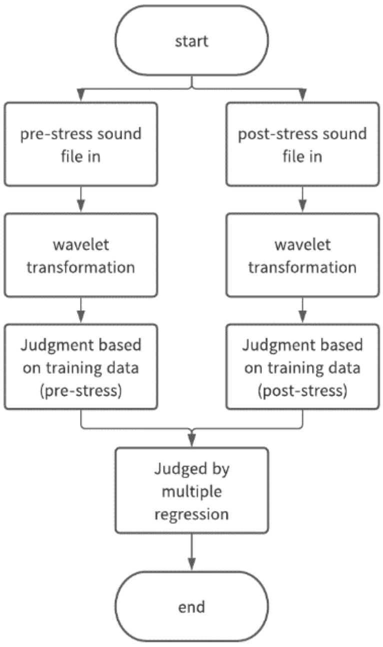


Fig.2 .Flowchart of the second analysis method

2.3. *Third analysis method*

From Fig.2, using separately pre-stress and post- stress, the judgment part has been changed. The discrimination rate was calculated by Decision tree, Random Forest, Linear SVM, and Nonlinear SVM.

3. Experiment

For the YOLOv2tiny model, the ratio of training data to test data was 7:3 for 100 subjects. The discrimination rate was calculated as the percentage of correct answers in the test data. The labels used were: good (i.e., normal; n = 45 subjects), need for follow-up (i.e., gray; n = 15 subjects), and need for treatment (i.e., abnormal; n = 40 subjects). For each pre- and post-stress, there were 100 trials (total 200 trials of audio data).

In the multiple regression analysis, all the data were trained and labeled as: normal = 1, gray = 0, and abnormal = (-1). After calculating the equation, the values were obtained by re-substitution. In the judgment, values lower than (-0.33) were considered abnormal, values between (-0.33) and 0.33 were considered gray, and values above 0.33 were considered normal.

Decision trees, random forests, and SVMs were all trained using MATLAB by MathWorks, and the discrimination rates were verified by single-tailed cross-validation. Morlet, morse, and bump wavelet transforms were tested.

There were 100 observations with 6 variables. Because of the small amount of audio data, the data was padded by clipping 5 seconds from the beginning of the audio and shifting it by 1 second. This resulted in 5,815 observations of image data. There were approximately 30 observations for each pre-stress audio and post-stress audio per person. Since the number of seconds of shunt sound varied, there was some variation in the padded image data, which made it difficult to label the same subject before and after stress. Therefore, we averaged the approximately 30 data mentioned earlier and adjusted the data so that there would be one observation of the three pre-stress variables and one observation of the three post-stress variables per person.

4. Result

The discrimination rate for YOLOv2tiny was calculated by assuming that the correct answer was the one where the judgment and label matched. For decision trees, random forests, and SVMs, the discrimination rate was calculated as the average of 100 cross-validation trials.

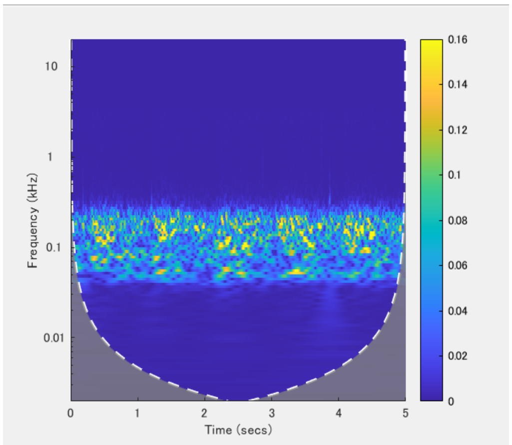


Fig.3 . Representative image of a normal subject after wavelet transform (bump)

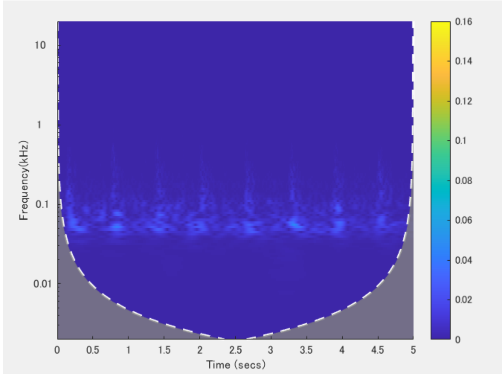


Fig.4 . Representative image of an abnormal subject after wavelet transform (bump)

From both wavelet transformations (Fig. 4 and 5), we observed periodic sounds continuously between 0.01 KHz and 1 KHz. Normal subjects characteristically had continuous waveforms with many high amplitude waveforms between 0.01 KHz and 1 KHz (Fig. 4). On the other hand, the abnormal subjects had only a few small amplitude waveforms and a series of high frequency sounds over a short time period (Fig. 5).

4.1. *Results of the proposed method 1*

The overall results for the proposed method 1 are 64% (morlet),57%(morse)and57%(bump). The discrimination rate of gray is low in all mother wavelets, and the discrimination rate of bump seems to be the most balanced.

Table 1. Discrimination rate with YOLOv2tiny only (morlet)

		Correct		
		normal	gray	abnormal
Method1	normal	84.6	14.4	1
	gray	57	9.3	33.7
	abnormal	22.5	16.2	61.3

4.2. *Results of the proposed method 2*

The overall results for the proposed method 2 are 89% (morlet),93%(morse)and97%(bump). Identification rate of bump was the highest at 97% (Table2). However, do not use cross-validation or other methods, so this result is not sufficiently generic.

Table 2. Discrimination rate when YOLOv2tiny is combined with multiple regression analysis (bump)

		Correct		
		normal	gray	abnormal
Method2	normal	97.8	2.2	0
	gray	0	100	0
	abnormal	0	2.5	97.5

4.3. *Results of the proposed method 3*

The overall results of the proposed method 3 are 88%(morlet), 89%(morse), 89%(bump) for Decision Tree, 87%(morlet), 90%(morse), 91%(bump) for Random Forest, 91%(morlet), 91%(morse), 95%(bump) for Linear SVM, and 90%(morlet), 91%(morse), 90%(bump) for Nonlinear SVM. The discrimination rate of morse and bump was the same 89% in the decision tree (Table3). In Random Forest, bump had the highest discrimination rate at 91% (Table4). In linear SVM, bump had the highest identification rate at 95% (Table5). In nonlinear SVM, morse had the highest identification rate at 91% (Table6). As a result, the results of the linear SVM were high, which indicates that this data is linear data. It can be seen that gray(uncertain) has a low discrimination rate in all results.

Table 3. Discrimination rate when YOLOv2tiny is combined with Decision Trees(bump)

Decision Tree		Correct		
		normal	gray	abnormal
Method3	normal	93.3	6.7	0
	gray	13.3	66.7	20
	abnormal	0	7.5	92.5

Table 4. Discrimination rate when YOLOv2tiny is combined with Random Forest (bump)

Random Forest		Correct		
		normal	gray	abnormal
Method3	normal	100	0.0	0
	gray	6.7	80	13.3
	abnormal	10	5	85

Table 5. Discrimination rate when YOLOv2tiny is combined with Linear Support Vector Machines (bump)

Linear SVM		Correct		
		normal	gray	abnormal
Method3	normal	97.8	2.2	0
	gray	6.7	80.0	13.3
	abnormal	0	2.5	97.5

Table 6. Discrimination rate when YOLOv2tiny is combined with Non-Linear Support Vector Machines (morse)

Nonlinear SVM		Correct		
		normal	gray	abnormal
Method3	normal	91.1	0.0	8.9
	gray	13.3	73.3	13.3
	abnormal	2.5	0.0	97.5

5. Conclusion

Rather than classify wavelet transform images solely with YOLOv2tiny, we improved the discrimination rate of shunt sounds by performing multivariate analysis on the variables derived from the YOLOv2tiny decision.

In the first analysis method, the highest result for the normal classification was 84.6% (Table 1), 63.2% for abnormal (Table 2), and 17% for gray (Table 3). Since YOLOv2tiny alone was not sufficient to improve the discrimination rate, we used multivariate analysis to improve the discrimination rate. In the multiple regression analysis (i.e., the second analysis method), the discrimination rates for bump were quite high (Table 1). However, this result did not allow us to verify the method's versatility, so we used cross-validation in the third analysis method to verify its generalizability. The results were 89% for decision trees, 91% for random forests, 95% for linear SVM (Table 5), and 91% for nonlinear SVM (Table 6). Overall, the mother wavelet that was best suited for shunt sounds analysis was bump, and linear multivariate analysis was found to be suitable for the data.

Future research will investigate why the discrimination rates of the gray classification were consistently lower than those of abnormal and normal classifications for all methods. We suggest that one of the reasons for this is the small number of gray data; therefore, we will include more gray trials. Another possible explanation is the inherent ambiguity in the gray category, which is an echo undiagnosable as either normal or abnormal; this ambiguity may necessitate subdividing gray into two or more sub-categories. Additionally, the analysis method should be developed further specifically to improve the pre-stress trials with the same accuracy as post-stress trials. Ultimately, the goal of this line of research is to create an automated system for screening arteriovenous (AV) fistulas for stenosis and blood clots without the use of an echocardiogram.

Acknowledgments:

In this experiment, audio data was collected using a device under development provided by Togo Medikit Co., Ltd. We appreciate Togo Medikit Co., Ltd for their cooperation.

References

[1] Ikuo Toufukuzi "The Future of Telemedicine Journal of the Japan Society of Home Economics" 62(2):141-144 ,2011(Japanese)

[2] Mizuho Tagashira and Takahumi Nakagawa "Discrimination Method for Vascular Stenosis by Acoustic Analysis of Dialysis Shunt Sounds" Biomedical Engineering 59(1):31-39,2021(Japanese)

[3] Joseph Redmon, Ali Farhadi; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7263-7271

[4] G. Li, Z. Song and Q. Fu, "A New Method of Image Detection for Small Datasets under the Framework of YOLO Network," 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2018, pp. 1031-1035,

[5] Sifuzzaman, M,slam, M R, Ali, M Z, 2009, Application of Wavelet Transform and its Advantages Compared to Fourier Transform. Journal of Physical Science;Vol 13 [2009] 121-134

[6] Ngui, W.K., Leong, M.S., Hee, L.M., Abdelrhman, A.M., 2013. Wavelet Analysis: Mother Wavelet Selection Methods. AMM 393, 953–958.

[7] S. R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 21, no. 3, pp. 660-674, May-June 1991

[8] Breiman, L. Random Forests. Machine Learning 45, 5–32 (2001).

[9] Haifeng Wang and Dejin Hu, "Comparison of SVM and LS-SVM for Regression," 2005 International Conference on Neural Networks and Brain, 2005, pp. 279-283,

[10] A. Mathur and G. M. Foody, "Multiclass and Binary SVM Classification: Implications for Training and Classification Users," in IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 2, pp. 241-245, April 2008,

Authors Introduction

Mr. Kyousuke Fujiwara

Kyosuke Fujiwara born in 2000. He is currently studying in department of environmental robotics, and will receive the B.Eng from University of Miyazaki in 2022.
His current research is Analysis of Shunt Sound from Vascular Access in Machine Learning.

Mr. Takayuki Yamamoto

Takayuki Yamamoto born in 1998. He is currenttly studying in Graduate School of Engineering, and will receive the M.Eng from University of Miyazaki in 2022.
His current research is A Study on the Algorithm of Auscultation SoundAnalysis for Heart Sounds.

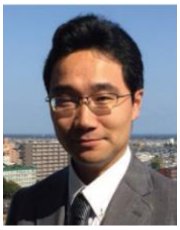
Dr. Lindsey Tate

She received a B.A. in Psychology (Neuroscience) from St. Edward's University in Austin, TX, USA in 2013. In 2016 and 2018, respectively, she received her M.S. and PhD in Cognitive Psychology from the University of Oklahoma, where she conducted perceptual and behavioral neuroscience research using EEG and MEG. Now she works as a researcher in the Faculty of Engineering at the University of Miyazaki. Her main research interests are applying machine learning and other algorithms to real-time non-invasive physiological data for use in clinical and non-clinical biofeedback systems.

Mr.Kazuya Kibune

He is currently a department of clinical engineering section manager in tokatsu dialysis hospital & clinics. He works as the clinical engineer and the clinical vascular technologist.

Prof. Hiroki Tamura



He received the B.E. and M.E. degree from Miyazaki University in 1998 and 2000, respectively. From 2000 to 2001, he was an Engineer in Asahi Kasei Corporation, Japan. In 2001, he joined Toyama University, Toyama, Japan, where he was a Technical Official in the Department of Intellectual Information Systems. In 2006, he joined Miyazaki University, Miyazaki, Japan, where he was an Assistant Professor in the Department of Electrical and Electronic Engineering. Since 2015, he is currently a Professor in the Department of Environmental Robotics. His main research interests are Neural Networks and Optimization Problems. In recent years, he has had interest in Biomedical Signal Processing using Soft Computing.