A SVM-Based EEG Signal Analysis: An Auxiliary Therapy for Tinnitus

Pei-Zhen Li[1,2], Juan-Hui Li[1,2], and Chang-Dong Wang[1,2](✉)

[1] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China
changdongwang@hotmail.com

[2] Guangdong Key Laboratory of Information Security Technology, Guangzhou, China

Abstract. Tinnitus is a kind of auditory disease characterized by an ongoing conscious perception of a sound in the absence of any external sound source. It is a common symptom for which no effective treatment exists. Though many non-invasive functional imaging modalities have been rapidly developed and applied to this field, yet, whether the EEG signal can be utilized to distinguish tinnitus patients from normal populations has not been investigated. In the present study, we perform a binary classification based on EEG signal to distinguish tinnitus patients from normal populations. In this study, 22 subjects are involved in the experiment with 15 of them being tinnitus patients and the others being normal controls. The collected EEG signals are preprocessed in frequency domain and well represented as features that depict each subject. Then the linear support vector machine is applied to classify the subjects. Satisfactory results have been achieved, where the accuracy of the classification could reach 90.91% in spite of the undeniable fact that the collected EEG signals contain noises. Accordingly, the present study reveals that the EEG signals can be utilized to distinguish tinnitus patients from normal populations, which could be regarded as an auxiliary therapy in tinnitus.

Keywords: Tinnitus · EEG signal · Binary classification · Support Vector Machine · Auxiliary therapy

1 Introduction

Tinnitus is a kind of auditory disease characterized by an ongoing conscious perception of a sound in the absence of any physical sound source, which is perceived continuously by 5–15% of the adult population. But so far there is no effective treatment for it, though several studies have shown that focal stimulation of the temporal cortex by repetitive transcranial magnetic stimulation (rTMS) [1–3] can suppress tinnitus perception, but the maximal amount of tinnitus suppression by rTMS decreases with time [1–4]. With the rapid development and widespread application of non-invasive functional modalities, such as electroencephalogram (EEG), magnetoencephalogram (MEG), and functional

C.-L. Liu et al. (Eds.): BICS 2016, LNAI 10023, pp. 207–219, 2016.
DOI: 10.1007/978-3-319-49685-6_19

magnetic resonance imaging (fMRI), such non-invasive methods make it possible to observe the central nervous system (CNS) in vivo. In particular, EEG is used most frequently in tinnitus since EEG signal is closely related with neural electric activity and has extremely high temporal resolution as well as spatial definition at the level of the scalp compared with other modalities. On the one hand, the notorious tinnitus is in an urgent need for treatment, on the other hand, the non-invasive functional modalities are rapidly developed. It becomes necessary to apply those advanced modalities to the therapies for tinnitus, in particular, it requires some effective and applicable methods to analyze and deal with medical data generated from such non-invasive devices.

To this end, the present study proposes a simple but effective method to analyze the EEG signals. We utilize the well-known machine learning algorithm, namely support vector machine (SVM) to perform a linear classification based on the features extracted from EEG signals, which works well in distinguishing tinnitus patients from normal controls. Extensive experiments have been conducted to confirm the accuracy of our method in the binary classification. This new finding can be regarded as an auxiliary therapy in tinnitus and has great significance in accelerating the clinical therapies of tinnitus especially when the tinnitus is no longer uncommon in populations while the medical resources is limited.

2 The Method

2.1 Support Vector Machine

Support vector machine (SVM) is a classical supervised machine learning model with associated learning algorithm based on statistical learning theorem, which is widely used for classification. Given a set of training samples, each marked for belonging to one of two categories, the SVM training algorithm constructs a non-probabilistic binary linear classifier to assign new testing samples into one category or the other. In the SVM model, each sample is represented by a point in a high-dimensional space, where a hyperplane separates the samples into two categories. New samples will be mapped to the same space and tagged as one category according to which side of the hyperplane they fall on. A good classifier is obtained by a hyperplane that has the largest functional margin making a least generalization error. SVM aims to improve the generalization ability of the learning via seeking for a structural risk minimization, which makes it perform well in spite of the small sample size. This is why we choose SVM for classification due to a small number of subjects and large quantities of features [5]. Figure 1 demonstrates the main idea of SVM in the case of 2-dimensional space.

2.2 Feature Selection

Feature selection is a vital step for classification. An appropriate feature selection leads to an accurate classifier since the feature vector plays an important role in

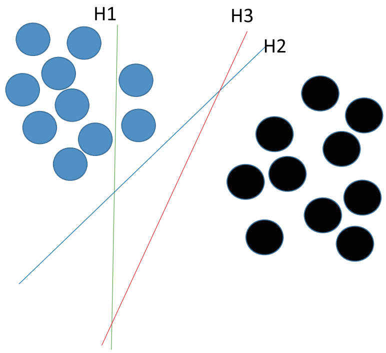


Fig. 1. SVM in 2-dimensional space. H1, H2, H3 are 3 hyperplanes, where H1 can not separate this two classes, while H2 and H3 can. H3 is the best choice with the maximum margin.

finding the "maximum-margin hyperplane" that divides the subjects represented by points in high-dimensional space. In the present study, each subject is mapped to a point represented by a feature vector with 129 features extracted from the corresponding 129 EEG signals collected from 129 electrodes.

According to the signal processing theorem, the EEG signals can be analyzed both in time domain and frequency domain. We propose to analyze in frequency domain due to the fact that EEG signals are often characterized and distinguished by their frequency bands: delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), and gamma (30–45 Hz) [6]. Therefore, the collected EEG signals are transformed into frequency domain after fast fourier transformation (FFT), where the number of FFT data points is equal to the sampled EEG data points. We compute the frequency spectrum for every EEG signals. As shown in Fig. 2, it is clear that each frequency spectrum consists of one primary tone at certain frequency point with the biggest absolute value in amplitude plus lots of other weak tones with small absolute value in amplitude. The primary tone is no wonder the most representative tone for this EEG signal, thus, the phase is calculated and the value is assigned to the feature vector after being mapped via a nonlinear cosine function, that is, the 129-element feature vector is composed of the strongest, cosine mapped component of the corresponding EEG signals. This is a technique to reduce the dimension of the feature vector and the disturbance in the sampled EEG data.

The cosine mapping can be viewed as a strategy aiming at recovering the essence of the signal while removing the strength or amplitude information. It makes sense in view of the large variability existing among individuals. Here, the phase and strength(amplitude) components are separated from one signal, where the signal strength(amplitude) varies much more from one subject to another due to many personal factors such as age, gender, emotion, etc. It may be more consistent when only signal phase is taken into consideration. To emphasize the function of this cosine mapping, we also perform the binary classification directly based on the extracted 129 phases as a contrast.

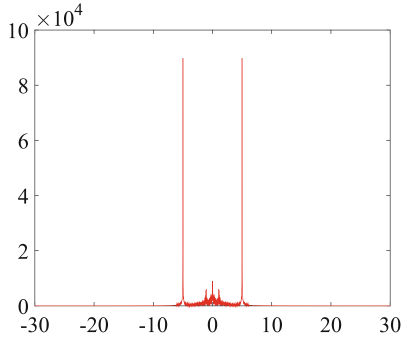


Fig. 2. The frequency-domain EEG signal of one electrode.

2.3 The Application of SVM

The linear support vector machine is applied to perform the classification due to a small number of subjects and large quantities of features [5]. During the process of classification, part of the subjects are used to train a SVM classifier, then all the subjects are used as testing data. The linear SVM is supplied by MATLAB toolbox. The performance of the classification is evaluated by accuracy, false positive rate (FPR) or Sensitivity (measuring the hit rate in tinnitus group) and false negative rate (FNR) or Specificity (measuring the hit rate in normal group), where the normal controls are labeled as 1 and tinnitus patients are labeled as -1. The formulas are displayed as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$FPR = \frac{FP}{FP + TN} \tag{2}$$

$$Sensitivity = \frac{TN}{FP + TN} \tag{3}$$

$$FNR = \frac{FN}{FN + TP} \tag{4}$$

$$Specificity = \frac{TP}{FN + TP} \tag{5}$$

where TP stands for the number of normal controls who were correctly classified; TN stands for the number of tinnitus patients who were correctly classified; FP stands for the number of tinnitus patients who were misjudged to be normal controls; and FN stands for the number of normal controls who were misjudged to be tinnitus patients.

3 Experiment

3.1 Subjects

22 subjects participate in this study, and they involve two groups: 15 patients (tinnitus group) and 7 healthy control subjects (control group). Table 1 shows the demographic information of the subjects. All the participants are recruited from the second affiliated hospital of Sun Yat-sen university. The patients are diagnosed by expert clinicians via a series of test routines. After filling in the tinnitus handicap inventory, the score (THI value) will be recorded as a partition criterion to the tinnitus severity level in the tinnitus group. None of the employed patients use medications that is expected to influence the EEG signals. All subjects involved in this study have normal or corrected-to-normal visual acuity, and no color blindness. Informed written consent is obtained from all the subjects before conducting the experiment.

Table 1. Demographic information of subjects participated in the study (THI, i.e. tinnitus handicap inventory, a measure to the tinnitus degree).

Index	Type	Gender	Age	THI
No.1	tinnitus	male	29	60
No.2	tinnitus	female	26	34
No.3	tinnitus	male	42	82
No.4	tinnitus	male	61	62
No.5	tinnitus	female	37	58
No.6	tinnitus	male	26	24
No.7	tinnitus	male	39	66
No.8	tinnitus	male	35	54
No.9	tinnitus	N/A	N/A	N/A
No.10	tinnitus	male	22	40
No.11	tinnitus	male	28	54
No.12	tinnitus	female	59	52
No.13	tinnitus	female	47	78
No.14	tinnitus	female	47	88
No.15	tinnitus	male	N/A	N/A
No.16	normal	female	42	–
No.17	normal	female	27	–
No.18	normal	N/A	N/A	–
No.29	normal	N/A	N/A	–
No.20	normal	N/A	N/A	–
No.21	normal	N/A	N/A	–
No.22	normal	N/A	N/A	–

3.2 EEG Data Collection and Preprocessing

The resting state EEG signals are used in this study. All the subjects are requested to calm down and sit upright on a comfortable chair with eye opened in a fully lighted room, wearing an electrode cap with 129 electrodes positioned according to the 10–10 international electrode placement system. The 2-dimensional position is shown in Fig. 3. Continuous EEG data are recorded from 129 scalp sites using the EEG acquisition system. During the whole 4 min data acquisition process, the impedance of each electrode is kept below 40 KΩ. The raw EEG signals are divided into 2-second epochs, by which we can get 120 epochs in total(the sampling rate of the EEG acquisition system is 1000 Hz). All episodic artifacts including eye blinks, eye movements, teeth clenching, body movement and ECG artifacts are removed from the original EEG signals. The signals are bandpass filtered with the accepted frequency band ranging from 0.1 to 60 Hz. Here, sampled EEG signals in time domain are shown in Fig. 4(a) and (b) for tinnitus patient and normal control respectively.

The overall process of the experiment is summarized in Fig. 5.

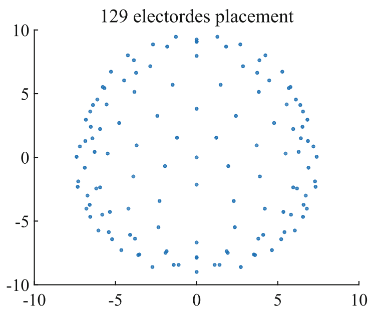
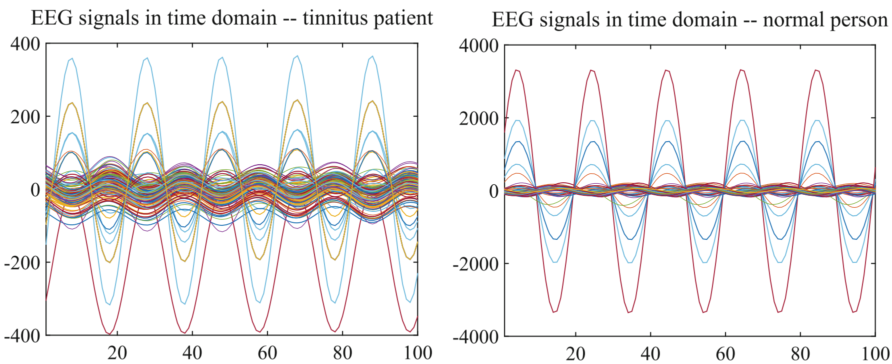


Fig. 3. The 2-dimensional placement for 129 EEG electrodes.



(a) EEG signals for tinnitus patient (b) EEG signals for normal person

Fig. 4. Time domain EEG signals.

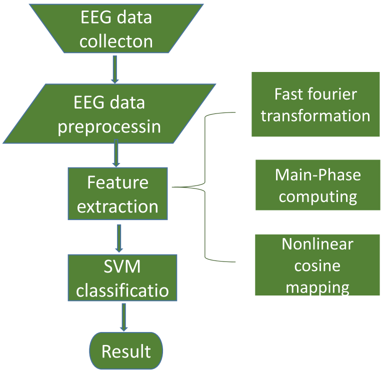


Fig. 5. The overall flow diagram for the experiment.

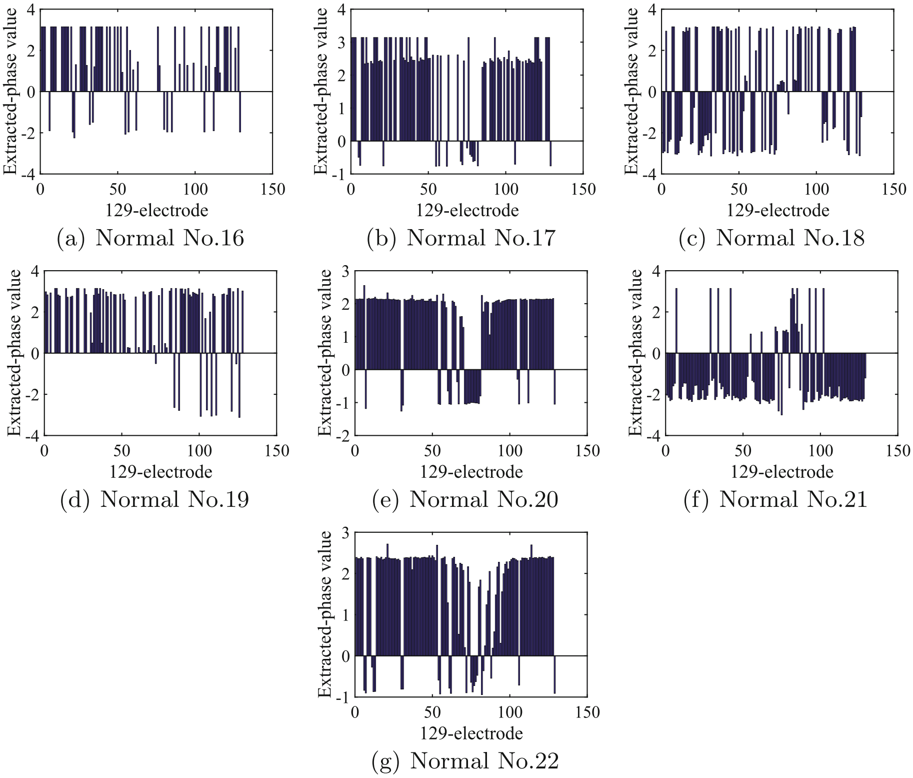
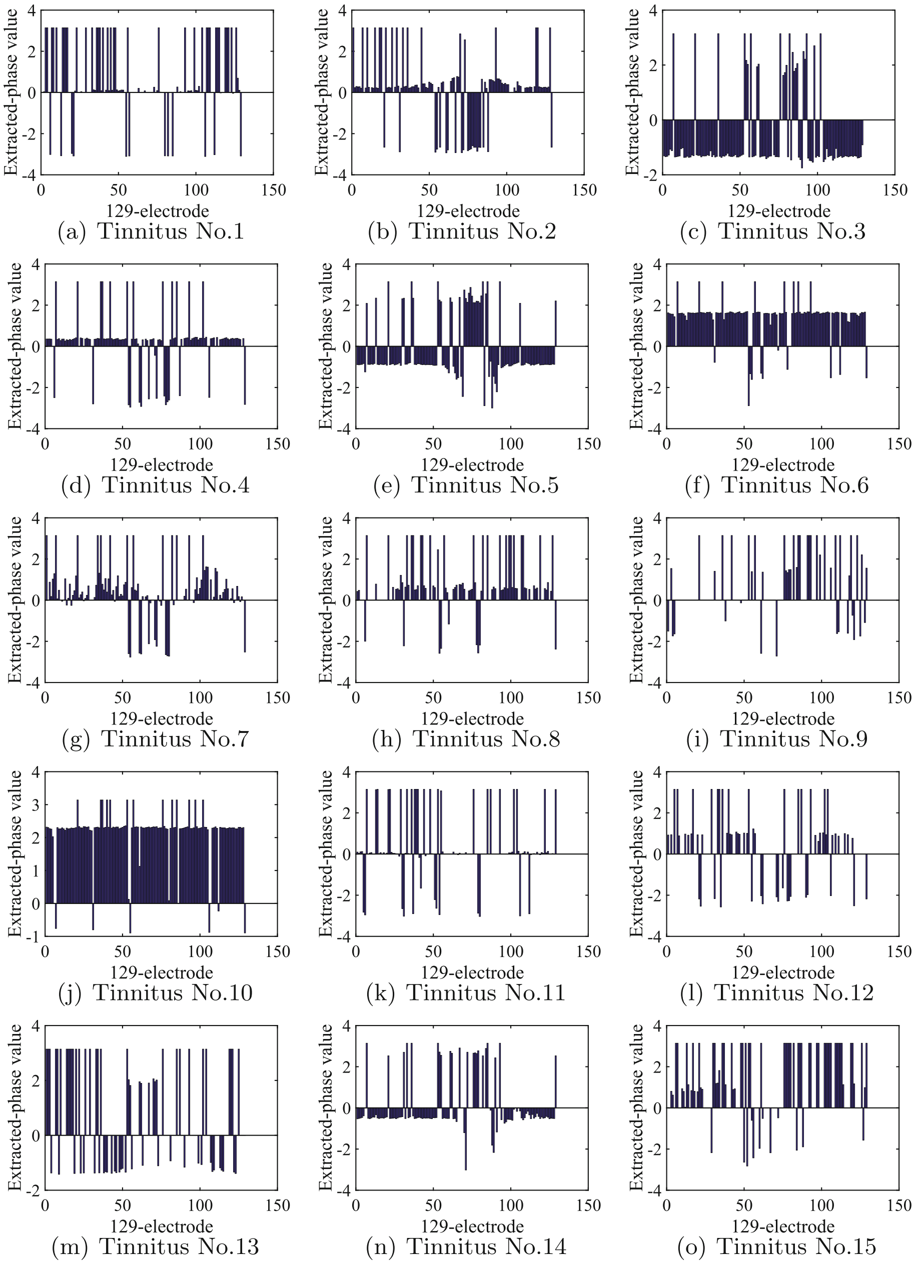


Fig. 6. Extracted phases (without cosine mapping) for 7 normal persons.



(a) Tinnitus No.1 (b) Tinnitus No.2 (c) Tinnitus No.3
(d) Tinnitus No.4 (e) Tinnitus No.5 (f) Tinnitus No.6
(g) Tinnitus No.7 (h) Tinnitus No.8 (i) Tinnitus No.9
(j) Tinnitus No.10 (k) Tinnitus No.11 (l) Tinnitus No.12
(m) Tinnitus No.13 (n) Tinnitus No.14 (o) Tinnitus No.15

Fig. 7. Extracted phases (without cosine mapping) for 15 tinnitus patients.

Table 2. Part1: classification result without cosine mapping.

	Number of normal controls	Number of tinnitus patients
Ground truth	7	15
Results	9	13

Table 3. Part1: evaluation for the classification.

Performance	Value
Accuracy	81.82%
FPR	20.00%
Sensitivity	80.00%
FNR	14.29%
Specificity	85.71%

3.3 Result

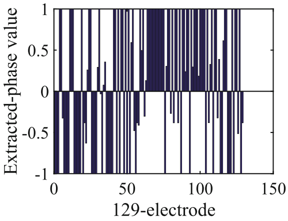
Part1: Classification without Cosine Mapping. Figures 6 and 7 report the extracted phases for the 22 subjects (7 normal controls, 15 tinnitus patients).

From these figures, we can see that the phase distribution in the normal group turns out to be denser and tidier than tinnitus group, though the collected data contains noise in view of the fact that some subjects may be distracted in the process of EEG data collection. When the classifier is trained to be good enough, we are sure to get a satisfied result.

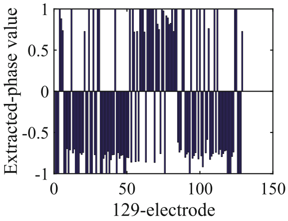
Tables 2 and 3 report the results in terms of accuracy, FPR, FNR, sensitivity and specificity. The results are quite satisfactory with accuracy higher than 80%, sensitivity being 80% and specificity reaching up to 85.71%, which indicates an anti-noise classifier in classifying EEG data with unavoidable personal disturbance. But still, we try another method in order to magnify the differences between two groups. The even better results are reported in the next part.

Part2: Classification with Cosine Mapping. Figures 8 and 9 report the cosine mapped features for the 22 subjects (7 normal controls, 15 tinnitus patients).

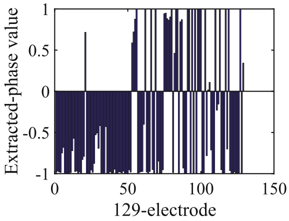
From these figures, we can see that there exist obvious differences in the distribution of cosine mapped values for tinnitus group and normal group, specifically,



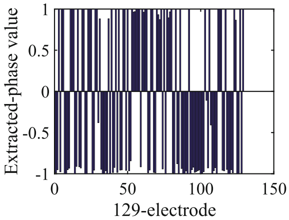
(a) Normal No.16



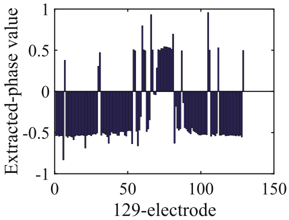
(b) Normal No.17



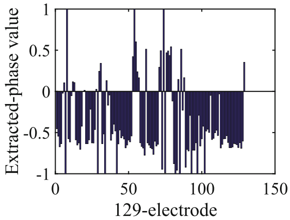
(c) Normal No.18



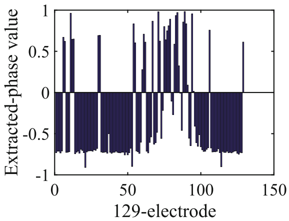
(d) Normal No.19



(e) Normal No.20



(f) Normal No.21

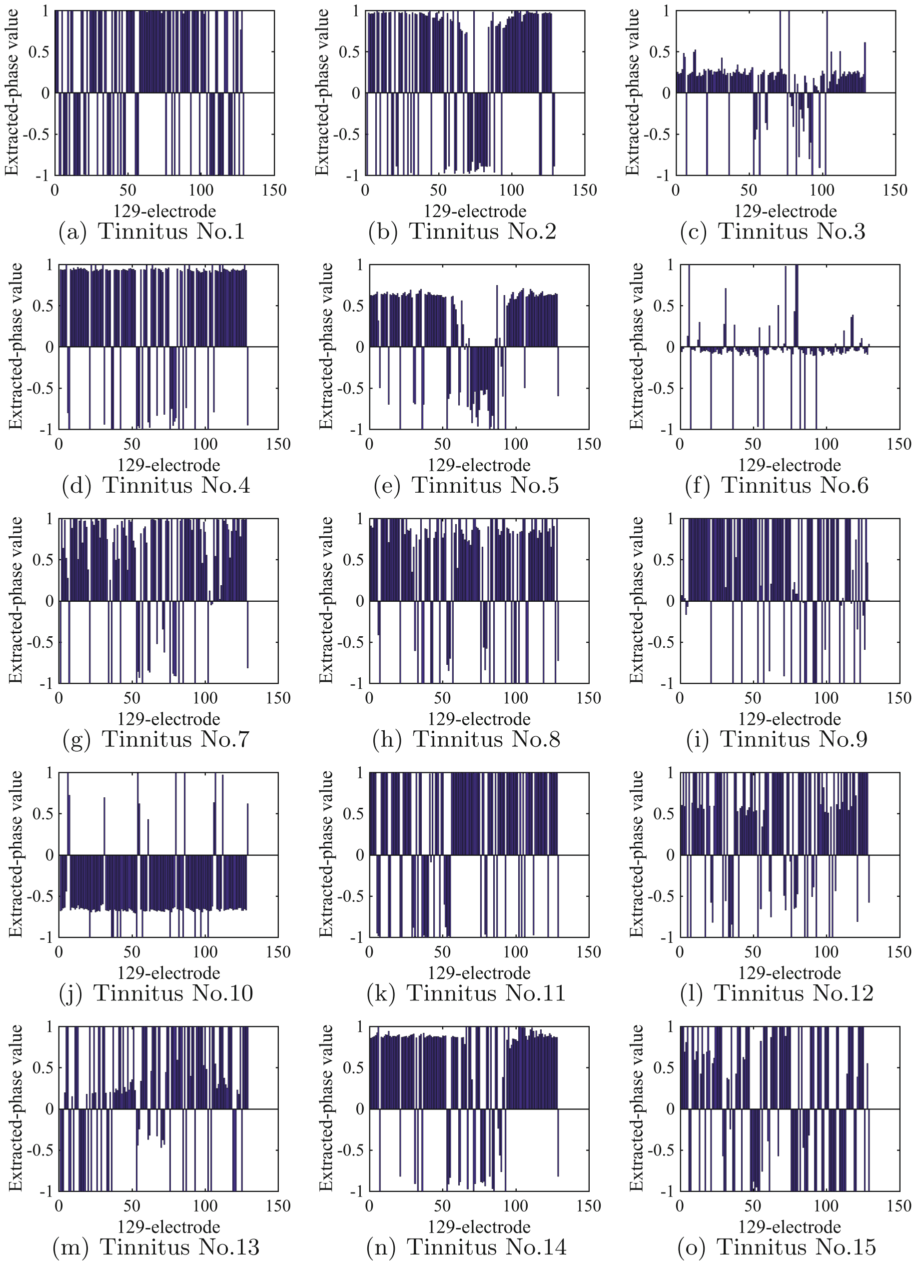


(g) Normal No.22

Fig. 8. Extracted phases (with cosine mapping) for 7 normal persons.

the negative part turns out to be much denser in tinnitus group than normal group, which is a difference-magnified version of classification in the previous part and may account for the better result in this part.

Tables 4 and 5 report the results in terms of accuracy, FPR, FNR, sensitivity and specificity. In this case where cosine mapping is applied to classification, only 2 tinnitus patients are misjudged to be normal controls and the specificity reaches up to 100%.



(a) Tinnitus No.1 (b) Tinnitus No.2 (c) Tinnitus No.3
(d) Tinnitus No.4 (e) Tinnitus No.5 (f) Tinnitus No.6
(g) Tinnitus No.7 (h) Tinnitus No.8 (i) Tinnitus No.9
(j) Tinnitus No.10 (k) Tinnitus No.11 (l) Tinnitus No.12
(m) Tinnitus No.13 (n) Tinnitus No.14 (o) Tinnitus No.15

Fig. 9. Extracted phases (with cosine mapping) for 15 tinnitus patients.

Table 4. Part2: classification result with cosine mapping.

	Number of normal controls	Number of tinnitus patients
Ground truth	7	15
Results	9	13

Table 5. Part2: evaluation for the classification.

Performance	Value
Accuracy	90.91%
FPR	13.33%
Sensitivity	86.67%
FNR	0.00%
Specificity	100%

4 Conclusion

In this study, we utilize the machine learning algorithm, SVM, to distinguish tinnitus patients from normal controls, where the features corresponding to each subject are extracted from the 129 EEG signals. The EEG signals are carefully analyzed and preprocessed in frequency domain in order to fully express the variance between different groups (i.e., tinnitus group and normal group) and to be qualified as features in the vital process of classification. The main finding of this study is that the machine learning algorithm, SVM, can be utilized to deal with the EEG signals, and it works well in distinguishing tinnitus patients from normal controls in terms of the classification results. This new finding is of great significance from the following perspectives: first, it made much progress in the treatment of the mysterious tinnitus. The well trained classifier can help doctors to analyze the EEG data in an efficient way, particularly, the misjudged subjects may reflect more information, for example, if a person claims to suffer from tinnitus but only to be classified to the normal group, we would conclude that this is a patient with slight tinnitus. Therefore, the binary classifier could also play an important role in categorizing different tinnitus patients. In this sense, the new finding can be regarded as an auxiliary therapy in tinnitus. Secondly, this is a good attempt of applying ML algorithm to deal with medical data. The satisfied combination is bound to inspire more studies in this direction in the future.

Acknowledgment. This work was supported by NSFC (No. 61502543), Guangdong Natural Science Funds for Distinguished Young Scholar (2016A030306014), the PhD Start-up Fund of Natural Science Foundation of Guangdong Province, China (2014A030310180), the Fundamental Research Funds for the Central Universities (16lgzd15).

References

1. De Ridder, D., Verstraeten, E., Van der Kelen, K., De Mulder, G., Sunaert, S., Verlooy, J., Van de Heyning, P., Moller, A.: Transcranial magnetic stimulation for tinnitus: influence of tinnitus duration on stimulation parameter choice and maximal tinnitus suppression. Otol. Neurotol. **26**(4), 616–619 (2005)
2. Kleinjung, T., Steffens, T., Londero, A., Langguth, B.: Transcranial magnetic stimulation (TMS) for treatment of chronic tinnitus: clinical effects. Prog. Brain Res. **166**, 359–551 (2007)
3. Plewnia, C., Reimold, M., Najib, A., Brehm, B., Reischl, G., Plontke, S.K., Gerloff, C.: Dose-dependent attenuation of auditory phantom perception (tinnitus) by pet-guided repetitive transcranial magnetic stimulation. Hum. Brain Mapp. **28**(3), 238–246 (2007)
4. Khedr, E., Rothwell, J., Ahmed, M., El-Atar, A.: Effect of daily repetitive transcranial magnetic stimulation for treatment of tinnitus: comparison of different stimulus frequencies. J. Neurol. Neurosurg. Psychiatry **79**(2), 212–215 (2008)
5. He, Z., Xiao, B., Liu, F., Wu, H., Yang, Y., Xiao, S., Wang, C., Russell, T.P., Cao, Y.: Single-junction polymer solar cells with high efficiency and photovoltage. Nat. Photonics **9**(3), 174–179 (2015)
6. Knyazev, G.G., Volf, N.V., Belousova, L.V.: Age-related differences in electroencephalogram connectivity and network topology. Neurobiol. Aging **36**(5), 1849–1859 (2015)