Epileptic Automated Detection from EEG Signal Using Statistical Features and Machine Learning Technique

[1]Rand Ameen Azeez, [2]Firas Sabar Miften, [3]Mustafa J. Hayawi

ABSTRACT

Background and objective: Epileptic seizures detection based on EEG signals is crucial in the diagnosis of patients with epilepsy. Experts mainly employed the visual inspection to identify epileptic seizures; however, it is a tedious job for them. This study, proposed designing an automated model to detect if the patient has epilepsy or not this could support the clinical research, facilitates the task of experts and speeds up the detection process. Thus, the aim of this research is to develop an automatic model to detect epileptic seizures from EEG signals.

Material and methods: In this study, we present a new method for seizure classification for EEG signals using a statistical features coupled with a least square support vector machine (LS_SVM) classifier. To achieve this task, each single EEG channel has been divided into four clusters. In addition, each cluster is segmented into sub-intervals. A vector of statistical features is pulled out from each sub-cluster form the final features set. The obtained features set are sent to the LS-SVM classifier. In this paper, epileptic database from Bonn University is used to evaluate the performance of the proposed model.

Results: In this paper, our developed model has been tested with epileptic dataset and the proposed model produced an average accuracy, sensitivity and specificity of 100%, 100% and 100%, respectively.

Keywords: *Epileptic seizures, LS-SVM, Statistical features, EEG signals.*

I. INTRODUCTION

EEG is a clinical tool captures to monitor patient's status in order to determine his/he neurological issues [1]. Epilepsy is a neurological syndrome happened when an abnormal discharge occurred in the brain. It is associated with rapid changes in the electrical activity of the brain. Experts are defined epileptic seizures as signals associated with as a slow-spike wave. The unknown cause of most seizures cases could affect subject's daily life with loss memory, consciousness by which they could endure a risk of sudden death [2, 3]. The recent research has shown that about 4% of population have a seizure [4, 5].

[1] University of Thi-Qar, College of Education for Pure Science, Department of Computer Science, Iraq
[2] University of Thi-Qar, College of Education for Pure Science, Department of Computer Science, Iraq
[3] University of Thi-Qar, College of Education for Pure Science, Department of Computer Science, Iraq

Received: 03 Apr 2019 | Revised: 21 May 2020 | Accepted: 12 Jun 2020

Classic approaches of analysing EEG are considered not suitable because they required huge time and burden for neurologists. However, Visual inspection of patient's recordings could cause errors and misdetection [6]. Moreover, EEG signals are combined with background noise and artefacts during recording. To overcome these issues, designing automatic approaches for epileptic seizures identification is needed to help the neurologists.

The brain network is described as a nonlinear and complex system. That means detecting seizures based on a single-channel EEG is not enough. Processing multi-channel EEG to identify seizures has been proved to be vital role to analyse human brain [7, 8]. Our purpose in this study is to support neurological experts by developing an accurate model based on multi-channel features extraction.

Investigation epileptic seizures based automatic detection models back to 1970s. Much studies have been carried out for example Costa et al. [9] classified EEG recordings into four cases inter-ictal, pre-ictal, ictal and postictal. In that study, a total of 14 features were adopted using signal energy, wavelet transform and nonlinear models. Those extracted features were fed to artificial neural network. They obtained 99% accuracy. In [10], authors suggested a key-point based local binary model for of EEG epilepsy diagnosis. In that method pyramid of Gaussian filters were used coupled with local binary patterns the key-points. The histogram of patterns was analysed to extract feature set. A SVM was employed for classification purpose.

Jaiswal and Banka [11] presented a model based on two feature extraction techniques called one-dimensional local gradient pattern and local neighbor descriptive pattern. The performances of that model were assessed using a 10-fold cross validation. An SVM, decision tree and ANN were applied for the classification purpose. 99.82% and 99.80% accuracy, and sensitivity were obtained respectively. Mengni Zhou et al. [12] applied a convolutional neural network to distinguish epileptic seizures. Time and frequency domains were employed to analyse EEG. Three experiments were simulated to solve binary classification problems as well as three-class problem. Average accuracies of 96.7, 95.4, and 92.3% were gained.

In this study, we suggest an automated model for epileptic seizures detection utilising multi-channel EEG signals. We believe our model could help neurological in analysis EEG recordings to support patients with epilepsy. The rest of the paper is organized as follows: Section 2 lists the previous methods. In Section 3, the EEG datasets are explained. Section 4 is defines the methodology. Section 5 reports the classification results. Section 6 draws the conclusions.

II. MATERIAL and METHODS

In this paper, we propose algorithm for classifying EEG signals using statistical features and least square support vector machine (LS-SVM) for classification. The model of the proposed method using the LS-SVM to analyse EEG signals is presented in Fig. 1. More details regarding the proposed method and dataset will be presented in the next sections

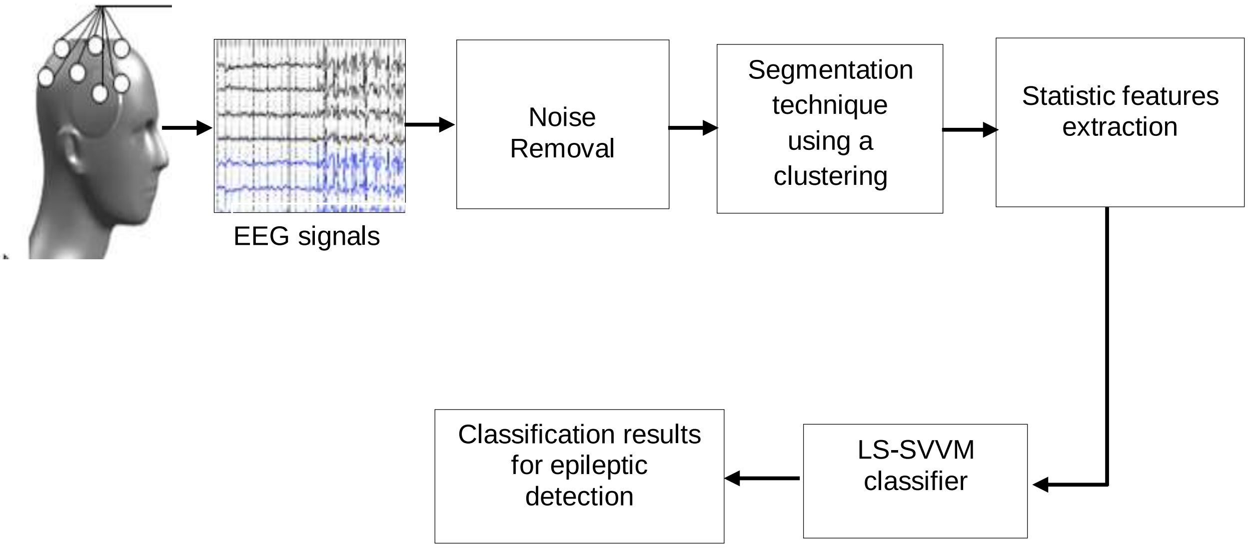

Fig. 1 Block diagram of the proposed methodology for EEG signal classification

EEG Recordings and Data Acquisition

The EEG dataset employed in this study are in [13]. The dataset contains five different sets represented as Z, O, N, F, and S with one included 100 single-channel EEG signal. EEG data was recorded at a sampling rate of 173.61 Hz. All EEG signals were filtered using Band-pass filter. Fig. 2 displays the 10–20 system used to record EEG signals. Fig.1 displays the 10-20 system used to electrode EEG signals. Sets Z and O were acquired from five healthy volunteers while Sets N, F, and S were recorded from five epileptic patients. Sets N and F were recorded activity recordings from seizure free. To make it clear, four sets Z, O, N and F were from normal EEG signals while set S was from epileptic EEG signals [14]. An example of each set is presented in Fig. 3.

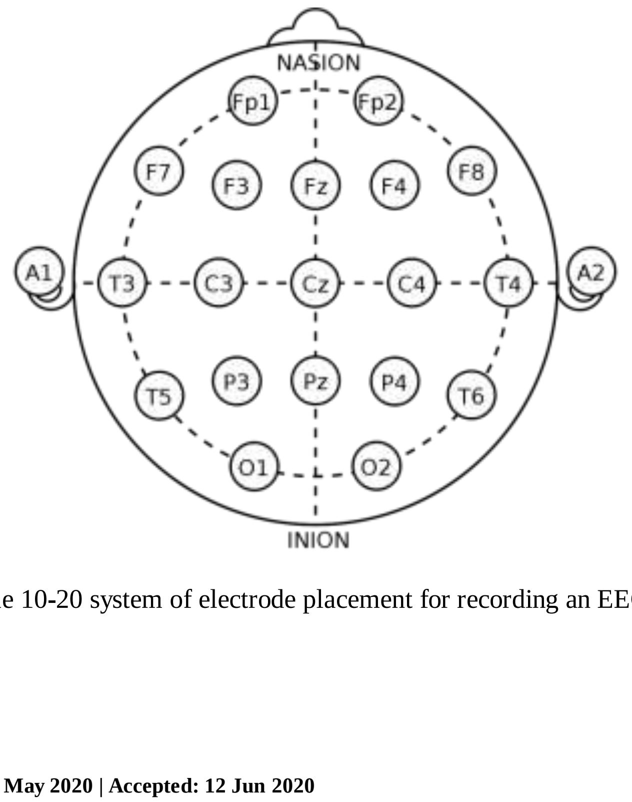

Fig. 2 The 10-20 system of electrode placement for recording an EEG pattern

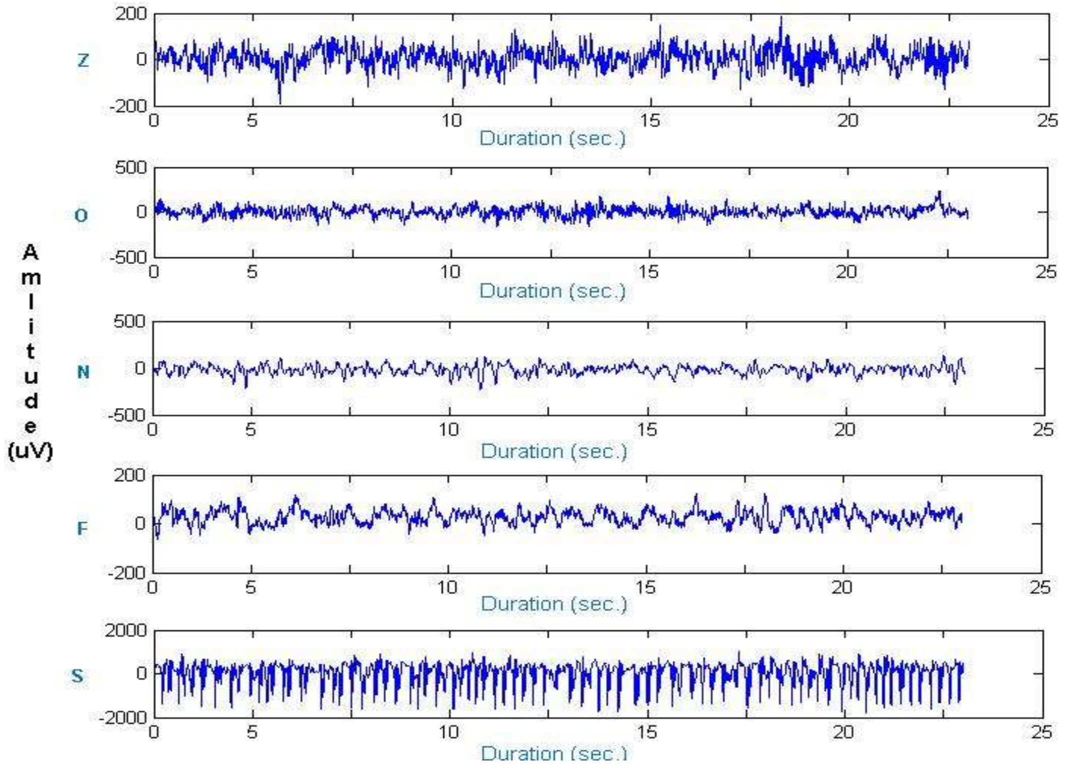

Fig. 3 EEG samples from sets, O, N, F and S

Segmentation Technique for Feature Extraction

Due to EEG data are non-stationary, it is important to divide EEG signal into intervals. Then, a vector of features is pulled out from each interval. We propose a feature extraction model to represent EEG signals Fig. 4 shows the steps of our model to extract the features and the segmentation method. These stages will be discussed in the following sub sections.

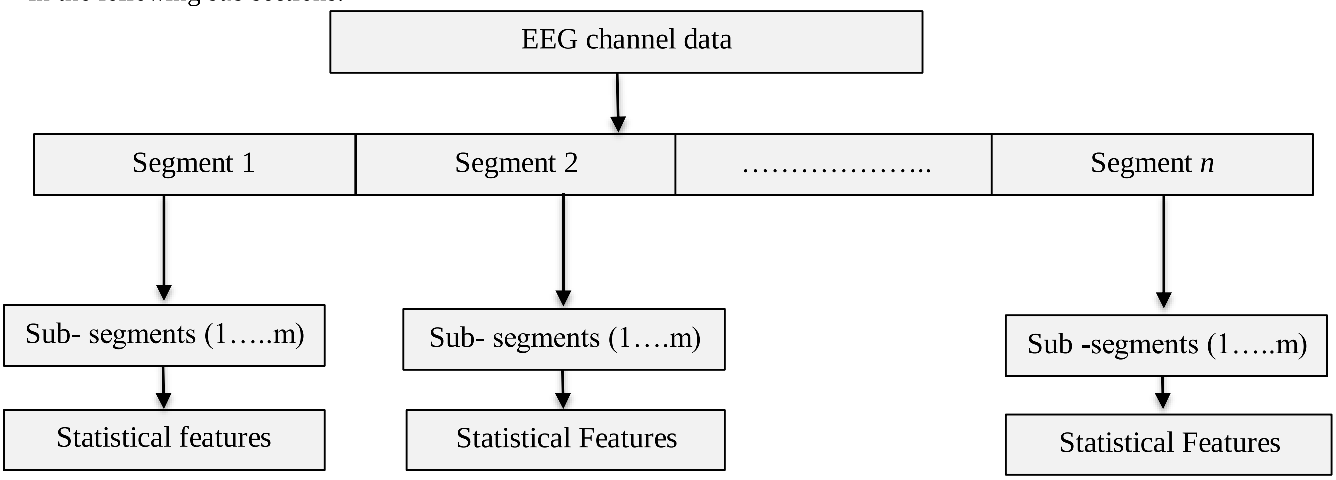

Fig. 4 Segmentation technique diagram for obtaining different segments, sub- segments and statistical features

Received: 03 Apr 2019 | Revised: 21 May 2020 | Accepted: 12 Jun 2020

- **Stage 1: Determination of segments**

Each EEG channel is represented a vector. This vector is segmented into number of intervals based on specific time duration where the number of intervals is n≥1.

- **Stage 2: Determination of sub-segments**

In this stage, each interval is divided into clusters using a specific period of time. We selected the number of clusters as m≥1. We defined the sizes of each interval and cluster empirically.

- **Stage 3: Statistical feature extraction**

Selecting powerful features improves the classification results due to the feature extraction remove the unwanted information from the original data. However, it is necessary to choose the key features from EEG dataset. In this paper, ten statistical features are adopted to represent EEG signals and they are extracted from each cluster. The features are denoted as *Maximum (X1), Range (X2), Standard deviation (X3), Minimum (X4), Mean (X5), Mode (X6), Median (X7), First quartile (X8), Third quartile (X9), inter-quartile range (X10).* The feature set is coded as *{X1, X2, X3, X4, X5, X6, X7, X8, X9, X10}*. EEG data is classified into symmetric and skewed [15]. We found that EEG data are symmetric distributions and other are skewed distributions. For these reasons, we used these features to investigate the symmetric and skewed behaviours of EEG data.

Implementation of the Proposed Method

Our proposed method is tested using the epileptic EEG data [13] which contains has five sets. In this method, each EEG channel is segmented into 4 intervals and each interval is divided into clusters. With each interval contains 1024 data points in 5.9 s while each cluster contains 32 data points of 0.1844 s. The ten statistical features namely *{X1, X2, X3, X4, X5, X6, X7, X8, X9, and X10}* are pulled out from each cluster to form a matrix of [2560×100]. That represents any two class signals. The matrix is fed to the LS-SVM algorithm. The LS-SVM is trained and tested using the extracted features.

Least Square Support Vector Machine (LS-SVM) for Classification

LS-SVM a popular classification model used for classification in many fields. It was designed by Suykens and Vandewalle [16]. Then, it was modified as a support vector machine (SVM) [17]. The LS-SVM resolves linear equations instead of the quadratic programming problem. The training points are employed to optimise the LS-SVM. In this study, we investigated a two-class problem and a multi-class classification problem by using the LSSVM. It was trained to distinguish the features of EEG data.

Performance Measures

The following metric were used to test the proposed model including accuracy, sensitivity, specificity, false positive rate (FPR)).

$$\text{Accuracy} = \frac{TP + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}};\ \ \text{Sensitivity} = \frac{TP}{\text{TP} + \text{FN}}\ \ ;\ \text{Specifcity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{False Positive Rat (FPR)} = \text{FP} + \text{TN}$$

III. RESULTS and DISCUSSION

The developed model to EEG signals was tested using epileptic dataset in six cases A–E. All 1024 vectors of 100 dimensions were used to train the model while the others 256 vectors for the testing purpose. Six experiments or cases were designed to detect epileptic EEG seizures. Five experiments were employed to classify a pair of EEG classes (compare just two class) and one experiment for all classes (compare all classes). The EEG cases are sorted out as follows:

Case1: Set A vs. Set E

Case2: Set B vs. Set E

Case3: Set C vs. Set E

Case4: Set D vs. Set E

Table 1. Experimental results of the proposed method

	Performance metrics			
EEG Case	**Accuracy**	**Sensitivity**	**Specifici**	**FPR**
Case 1	100%	100%	100%	0%
Case 2	100%	100%	100%	0%
Case 3	100%	100%	100%	0%
Case 4	100%	100%	100%	0%
Case 5	100%	100%	100%	0%
Case 6	100%	100%	100%	0%

Case5: Set A vs. Set D

Case6: All classes vs. all classes

Table 1 report the obtained results based on our model in terms of accuracy, sensitivity, specificity and the FPR for all the EEG cases.

From the results that has been obtain can see that the proposed method overcome on most of the methods of epileptic diagnosis that existing in this time. Table 2 show comparison of others researchers who use the same dataset with proposed model.

Table 2. Comparison researches that use epileptic dataset with proposed system

Research	Accuracy	sensitivity	Specificity
N. Sriraam et al. [18].	/	97.1%	97.8%
Raghu et al. [19].	99.45%	/	/
Annushree Bablani et al. [20].	96.7%	/	/
Hadi Ratham Al Ghayab et al. [21].	99.90%	99.80%	100%
Arindam Gajendra Mahapatra and Keiichi Horio [22].	99.91%	100%	/
Bhati D. et al. [23].	98.2%	/	/
Siuly et al. [24].	94.18%	94.92%	93.44%

Our findings demonstrated that the proposed method is efficient in EEG signals analysis and it best from most existing methods in all performance metrics in additional it take very short time execute this make the proposed method suitable for real time application.

IV. CONCLUSIONS

This study proposes a model for EEG data classification based on segmentation technique coupled with statistical features and least square support vector machine (LS-SVM) classifier. LS-SVM was implanted with RBF kernel function. We developed the proposed system to support experts by which the presence or absence of the epilepsy can be detected. Several experiments are designed using publicly EEG dataset. From the results of tests or experiments, and when compared to the results of existing methods, it can be concluded that the proposed method is very effective in determining whether a person has epilepsy.

REFERENES

1. Dabye AS, Issaka M, Gueye L (2015) Localization of epileptic seizure with an approach based on the PSD with an auto-regressive model. arXiv :1506.00947.
2. Kamath C (2013) A new approach to detect epileptic seizures in Electroencephalograms using Teager energy. ISRN Biomedical Engineering, Hindawi. https ://doi.org/10.1155/2013/358108.
3. Kamath C (2013) Teager energy based filter-Bank Cepstra in EEG classification for seizure detection using radial basis function neural network. ISRN Biomedical Engineering, Hindawi. https ://doi.org/10.1155/2013/498754.

4. Abbasi R, Esmaeilpour M (2017) Selecting statistical characteristics of brain signals to detect epileptic seizures using discrete wavelet transform and perceptron neural network. IJIMAI 4(5):33–38.
5. S. Pati , A.V. Alexopoulos , Pharmacoresistant epilepsy: from pathogenesis to current and emerging therapies, Cleve. Clin. J. Med. 77 (7) (2010) 457–467.
6. Zhou W, Liu Y, Yuan Q, Li X (2013) Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG. IEEE Trans Biomed Eng 60(12):3375–3381.
7. Ji Z, Sugi T, Goto S, Wang X, Ikeda A (2015) An automatic spike detection system based on elimination of false positives using the large-area context in the scalp EEG. IEEE Trans Biomed Eng 58:2478–2488.
8. Xia YS, Leung H (2006) Nonlinear spatial-temporal prediction based on optimal fusion. IEEE Trans Neural Netw 17:975–988.
9. Costa RP, Oliveira P, Rodrigues G, Leitão B, Dourado A. Epileptic seizure classification using neural networks with 14 features. Knowl-Based Intell Inform Eng Syst 2008;5178:281-8.
10. Tiwari AK, Pachori RB, Kanhangad V, Panigrahi BK. Automated diagnosis of epilepsy using key-point-based local binary pattern of EEG signals. IEEE J Biomed Health Inform 2017;21(4):888–96.
11. Jaiswal AK, Banka H. Local pattern transformation based feature extraction techniques for classification of epileptic EEG signals. Biomed Signal Process Control 2017;34:81–92.
12. Zhou M, Tian C, Cao R, Wang B,Niu Y, Hu T, Guo H and Xiang J (2018) Epileptic Seizure Detection Based on EEG Signals and CNN. Front. Neuroinform. 12:95. doi: 10.3389/fninf.2018.00095.
13. EEG time series data. Department of Epileptology University of Bonn; 2016, http://www.meb.uni bonn.de/epileptologie/science/physik/ eegdata.html .
14. Andrzejak RG, Lehnertz K, Mormann F, Rieke C, David P, Elger C. Indications of nonlinear deterministic and finitedimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys Rev E 2001;64(6).
15. R.D. De Veaux, P.F. Velleman, D.E. Bock, *Stats Data and Models*, 3rd ed., Pearson Addison Wesley, Boston, 2008.
16. J.A.K. Suykens, T.V. Gestel, J.D. Brabanter, B.D. Moor, J. Vandewalle, Least Square Support Vector Machine, World Scientific, Singapore, 2002.
17. V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.
18. N. Sriraam , S. Raghu, Kadeeja Tamanna, Leena Narayan , Mehraj Khanum, A. S. Hegde and Anjani Bhushan Kumar, (2018), "Automated epileptic seizures detection using multi-features and multilayer perceptron neural network", https://doi.org/10.1186/s40708-018-0088-8.
19. A S Raghu , Natarajan Sriraam , Alangar Sathyaranjan Hegde , Pieter L Kubben, (2019), "A novel approach for classification of epileptic seizures using matrix determinant", Expert Systems With Applications 127 (2019) 323–341.
20. Annushree Bablani, Damodar Reddy Edla, Shubham Dodia, (2018), "Classification of EEG Data using k-Nearest Neighbor approach for Concealed Information Test".
21. Hadi Ratham Al Ghayab . Yan Li . Shahab Abdulla .Mohammed Diykh . Xiangkui Wan, (2016), "Classification of epileptic EEG signals based on simple random sampling and sequential feature selection". https://DOI.10.1007/s40708-016-0039-.

22. Arindam Gajendra Mahapatra, Keiichi Horio, (2018), "Classification of ictal and interictal EEG using RMS frequency,dominant frequency, root mean instantaneous frequency square andtheir parameters ratio", https://doi.org/10.1016/j.bspc.2018.04.007.

23. Bhati D., Pachori R.B., Sharma M., Gadre V.M. (2020) Automated Detection of Seizure and Nonseizure EEG Signals Using Two Band Biorthogonal Wavelet Filter Banks. In: Naik G. (eds) Biomedical Signal Processing. Series in BioEngineering. Springer, Singapore.

24. Siuly, Yan Li , Peng (Paul) Wen, (2011), "Clustering technique-based least square support vector machine for EEG signal classification", https://doi.org/10.1016/j.cmpb.2010.11.014.