JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-15, No.-2, February (2020) pp 270-280 ISSN (Print) 0973-8975

SUPPORT VECTOR MACHINE APPROACH FOR HUMAN IDENTIFICATION BASED ON EEG SIGNALS

Shaymaaadnan Abdulrahman[1], Mohamed Roushdy[2], Abdel-Badeeh M. Salem[3]

[1]Department of Computer Engineering, Imam Ja'afar Al-Sadiq University, Baghdad, Iraq
PhD Student at Ain Shams University, Egypt

[2] Faculty of Computers & Information Technology, Future University in Egypt, New Cairo, Egypt

[3]Department of Computer Science, Ain Shams University, Cairo, Egypt

https://doi.org/10.26782/jmcms.2020.02.00023

Abstract

The signals of the electroencephalogram (EEG) have been applied for detecting as well as registering the electrical efficiency in the human brain. In this paper, EEG signals have been utilized for human identification. The reliability regarding a lot of biometric systems aren't adequate due to the possibility of being copied or faked. Thus the brain signatures have been applied as potential biometric identifiers. The aim of this paper is to apply sample entropy and graph entropy as feature extraction. While in classification Support vector machine (SVM) and K-Nearest Neighbor (KNN) have achieved. Machine Learning Repository (UCI) used as dataset. Experimental consequences on this dataset demonstrate substantial enhancement in the classification accuracy as compared with other testified results in the literature. Results showed that the classification accuracy with SVM for biometric identification is 90.8% while with K-NN is 83.7% .Our study using13channels to feature extraction.

Keywords: Electroencephalogram (EEG), Support vector machine, K-Nearest Neighbor, Machine learning

I. Introduction

EEG signals are considered as neuro signals that are developed because of many brain's electrical activities. Many kinds regarding the brain's electrical activities are corresponding to various states of the brains of animals and human. Such signals could be captured and processed for providing important data. Placing electrodes on scalp will provide the EEG signals. To record the brain activities of humans, the 10-20 EEG placement system has been applied [IX]. This system is widely applied for placing electrodes on scalp. With regard to the discussed system, more electrodes could be added for recording extra details. Figure 1 shows the electrode's placement.

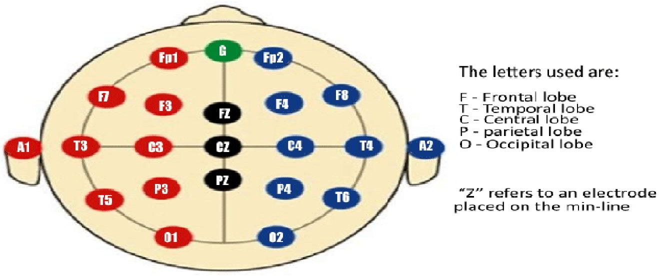

Fig. 1:EEG Electrode Placements

Typically, efficient EEG signals frequency is in the range of (0.5-40Hertz) [XXV]. EEG signals have several frequency bands. The delta signals are captured in frequency range (0.5-3.5Hertz). It has a tendency toward being slowest waves and highest in amplitude. Also, it could be detected in babies as well as in slow wave sleep in adults. Theta signal Frequency regarding these signals are in the range of (3.5 -7.5Hertz). Actually, extremely low theta waves are used for representing fine line between being in sleep or awake. In adults high theta levels are considered to be abnormal. The frequency of alpha signals is in range of (7.5-12Hertz). Beta signals can be defined as the brain signals where the frequency is in the range of (12-30Hertz). As shown in Figure 2.Usually alpha signal can be recognized on the two sides in symmetrical distribution. The Gamma signals are the signals with range of frequency of (31 Hertz and more). It does reflect consciousness.

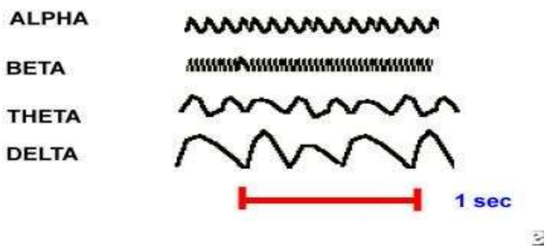

Fig. 2: Types of EEG Waveforms

The biometrics in the authentication systems are applied for identifying and allowing access control on the basis of human features. Therefore, it is applied widely in modern smart technology as well as clinical field. The EEG-based biometrics is the main focus of this study. An identity could be detected via the use of many biometric systems, such as iris recognition, the shape of the ear, signature dynamics, face recognition, finger-prints, retina, the geometry of hand, hand vein, speech recognition, and so on [XI]. Researchers are preferring to apply the EEG brain signals since they offer extra safety as well as strong security systems. It can't be faked or get forcefully, due to the fact that the signals of EEG are changing with various mental states [XVI]. Typically, EEG based biometric procedure need many methods. Initially, there is a need to acquire data of the subject. This data is obtained either from a website or by getting data from the person directly through the use of modern devices such as electrodes placed on the (head/ chest). Secondly it requires a preprocessing step. In this processing step, a filter is used for eliminating the noise effect and artifacts. Thirdly, a module of feature extraction is required for identifying the representative elements of the signal. Matching module that is going to develop scores for verifying personalities[II]. Classification regarding the EEG signals recordings viacognitive preparing was of high importance lately. In this research article, EEG data of thirteen channels from Machine Learning Repository (UCI)dataset have used.Nonlinear features such as sample entropy and graph entropy have been processed based on EEG data. These computed features fed to the various classifiers like (SVM and KNN). Performance hascomputed based on classification accuracy.

II. Rellated Work

Mustafa, et al, suggested EEG data which were recorded throughout certain mental arithmetic operation, silent reading has been examined through discrete wavelet transform (DWT), also feature vectors have been acquired. The classification applied SVMs. The results have been provided for twenty-six channels mental arithmetic and silent reading tasks. Based on twenty-six channels, the classification accuracy has been (90.71)%.

DuaaAIQattan and Francisco Sepulveda, acquired aEEG brain activity has been acquired at the same time as imagining execution regarding six one-handed sign from the American Sign Language (ASL) have been examined and utilized entropy features as feature extraction from the EEG of brain.

While Syed Anwar,was proposed DWT and PCA as feature extraction and BF and SVM were applied for classifier.The accuracy was(83%) when used 14 channels of electrodes .

The S-transform method has employed for extracting the fundamental signal features for the classification techniques as proposed by V. Sankara Narayanan ,et al. SVMs, KNN, and random forest have been applied. Accuracy has been between (78)% and (82)%.

The EEG recognition procedure on the basis of the channel frequency selection has been suggested by Deming Zhang, et al [10]. Original signals of EEG have been filtered through various frequency bands. One-versus-one CSP is used to extract the feature vectors related to filtered EEG signals. SVMs have been utilized as classifiers. Accuracy has been (86.8) % .

Nivedha R, et al, suggested classifying the emotions of human through the signals of EEG in 4 states (relaxed, sad, happy, and angry). The study applied statistical and spectral features which are extracted via DWT and SVM classifier and the accuracy was(80.6)% for a combination of 32 electrodes.

All aforementioned works utilize machine learning and statistical approaches in different path. In feature extraction is used like(sample entropy ,PCA, DWT,CSP,AAR… etc.,) while in classification is used like(SVM, KNN, LDA,CNN, …etc). It is a possibility that they have utilized such methods to all the channel data or selected number of channels i.e. C3,C4,P3,P4,O1,O2,A1,A2 ,….. . etc.

In our study, data acquisition was from 13 channels like(AF8,C1,C2,C3,C4,CP1,CP5,CP6,FC5,FT7,P8,PO8,PZ) as feature extraction as shown inFigure 3.

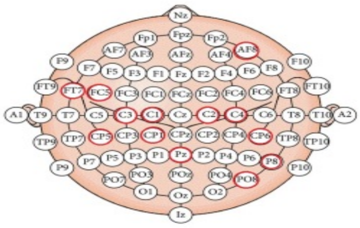

Fig.3: international 10- 20 electrode system

MATLAB of version (R2015b), as well as Intel Corei5 computer has been applied for classification as well as for feature extraction. The structure related to the rest of this study is described in the following way: Section IIRelated work, Section IIIBrain signals from medical aspects, Section IV proposed work , Finally Section classifier

III. Brain Signals FromThe Medical

Electroencephalogram signals are created from manyactivity in the neurons system. When the neuron system is activate local current flows are generated [XV]. Usually, measuring the brain signal is through synaptic excitation of the dendrites of neurons system in cerebral cortex in the head of human. The brain of human will have information through various sensors, including hearing, touch, sight, smell, and taste. Thus, the received information will be constructed via the brain to create meaning messages. The ability of moving the arms, legs, also they human thoughts, speech, and memories are all achieved via the brain. Peripheral system and central system are the 2 main types of nervous system [XII]. Spinal cord and brain are the two major portions of the central nervous system, whereas peripheral nervous system consists of spinal nerves which are branching from spinal cord as well as cranial nerves which are branching from the brain of human. Peripheral nervous system is responsible for autonomic nervous system that control important tasks in the body, including hormone's secretion, digestion, breathing, and heart rate. The brain is shielded from injuries via the skull. It consists of eight bones (frontal two parietal, dual temporal sphenoids, occipital and ethmoid) as shown in Figure 4. As the presented study compared the brain with respect to the other mammals, the brain of human has the most developed and the largest cortex. Neural process associated to capabilities such as language as well as complex reasoning speech, and so on etc. That differentiate the humans from the other mammals happen in that part of brain [VII] [X].

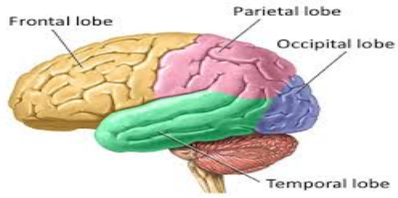

Fig. 4: Brain anatom **IV.Proposed work**

A. Data Acquisition

The EEG data used in the present study is available from the UCIrepository dataset[I] [V] . There are twelve input feature vectors and only one target vector in the data-set. Input feature vectors have been acquired though utilizing wavelet packet analysis on original signals in frequency band of (7-13 Hertz). The planning or relaxed states are represented via the target vector. For the purpose of experimental details as well as knowing more regarding the EEG data-set, one might refer to [XIX]. There are ninety-one samples in the testing data, also there are ninety-one samples in training data. As the data-set applied in [XIX], is fifty percent testing data and fifty percent training data. Comparable number regarding testing and training data have been applied in the presented study. The suggested research methodology of the presented study is displayed in Figure5.

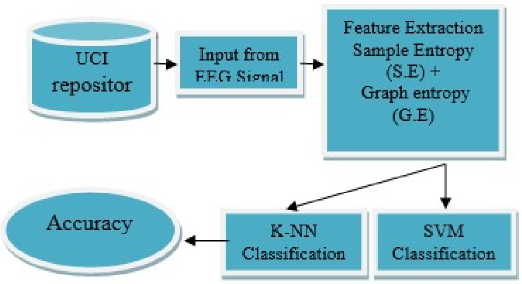

Fig. 5: diagram of proposed search

B. Feature Extraction

Sample entropy (S.E) and Shannon entropy are applied asfeature extraction. In 2000, Richman and Moorman have developed the sample entropy as an enhanced approximate entropy approach, that is defined to be non-linear dynamic parameter for measuring sequence complexity [XVII]. It is applied for evaluating the intricacy in the time arrangement information for a short length. This method has 3 input parameters (1) m: embedded dimension, (2) r: similarity criterion, (3) n: time series' length [XX]. In our case EEG signals, for calculating sample entropy 2 features Se1: m=2, r=0.15 Se2:m=2, r=0.2of every epoch of the EEG signals is extracted. Equation 1 is used to define the sample entropy. Based on edges or vertex, there are many graph entropy calculation approaches.

$$\text{Sample Entropy } (m,r) = \ln\left[\frac{B^m(r)}{A^m(r)}\right] \quad (1)$$

$$A^m(r) = \frac{1}{N-m}\sum_{k=1}^{N-m} A_k^m(r) \quad (2)$$

$$B^m(r) = \frac{1}{N \quad -m} \sum_{k=1}^{N-m} B_k^m(r) \tag{3}$$

$B^m(r)$ the probability of the two sequences which

match for m points

$A^m(r)$ the probability for two sequences to match

for*m+1* points

$$B_k{}^m(r) = \frac{1}{N-m-1} B_K$$

Where Given N data points from a time series

$$\{\{x(n)\} = x(1), x(2), \ldots . x(N).\}.$$

Take m vectors

$$X_m(1), \ldots \ldots X_m(N-m+1) defined \ as \ X_m(i) = [x(k), x(k+1), \ldots x(k+m-1)]$$

$$\text{For } 1 \leq k \leq N-m+1. at\ the\ ith\ sample$$

Many types of graph entropy(G.E) calculation approachesusing either vertex or edges. This study describes the graph entropy with Shannon's entropy formula [XXVI]:

$$h = -\sum_{i=1}^{n} p(k) \log\big(p(k)\big) \tag{4}$$

where*p(k)* is the degree distribution regarding of *i*

The degree distribution stands for a probability degree kover a degree sequence. It can be obtainedthrough counting the number of nodeswith degree k divided by the degree sequence size.

IV. Classification

The signals obtained are classified by utilizing different machine learning classifier strategies. Usually Machine learning methods are applied for easy computation of the signal. The two different classification methods are discussed below.

A. Support vector machine (SVM)

In this study, SVM classifier to classify the EEG signals has used. There are numerous SVM techniques as in linear SVM, quadratic SVM, cubic SVM. SVM has been employed as binary classification method to classify two classes from numeric EEG data. Support vector machine is a another type of classification method . Each data item hasschemed as a point in number of features-dimensional space with the magnitude of every feature and particular coordinate [II]. It is a one of the finest machine learning classification that which yields results with high accuracy. And it refersto a supervised classification technique [XX] [XXI]. SVM is a group of relevant supervised machine learning strategies which test information and identify framework in data using for classification. SVM-based classification has been recognized to keep up the ideal equilibrium between accuracy precision accomplished on constrained supply of training and testing data. The SVM classifier can be defined as

$$g(x) = w.\Phi(x) + b \qquad (5)$$

$$w = \sum_{i=1}^{n} \alpha_i y_i \Phi(x_{pi}) \qquad (6)$$

$$b = \sum_{i=1}^{n} \alpha_i \; y_i \Phi(x_{pi}) \,.\, \Phi(x_{pj}) + y_i \qquad (7)$$

Where 'w' is the normal to separating hyper plane defined by $\Phi\ (x)$. For a set of data points $(xpi,)$ where (y= +1,−1) the margin between two classes is given by,$\frac{2}{||w2||}$. The optimal margin is found by minimizing constrained optimization problem which is further solved by reducing. It to quantization programming optimization problem that yields. In this paper the SVM algorithm with RBF kernel is implemented for both feature extraction and classification [XXII].

B.K-Nearest Neighbor (KNN)

K-Nearest Neighbors Algorithms (K-NN) which is one of the supervised learning algorithm, is applied to EEG data**.** Asindicated by k-NN calculation that was utilized in classification. Attributes which separated from classification andviewing the entire, distance among new individual which should be classified and earlier individuals and then the nearest k category is utilized (Natalia, et al., 2019). Because of that procedure test information has a place with the closest k neighbor classification that has more individuals in certain class. The most significant improvement issues in k-NN technique are identification of neighbor's number and strategy of distance calculation algorithm as shown in Figure 6. The comparison results between the existing works and proposed work are illustrated in Table 1.

Table 1: comparison of various classification algorithms

no	*Methods*	*# of channels*	*Accuracy*	*Dataset*
1	Discrete wavelet transform (DWT),+ SVM [VI]	**# 26** F7, F8, AF1, AF2, FZ, F4, F3, FC6, FC5, FC2, FC1, T8, T7, CZ, C3, C4, CP5,CP6, CP1, CP2, P3, P4, PZ,P8, P7, PO2	90.44%	DataSet of study
2	DWT+ PCA+RBF SVM [XXIII]	**# 14** C3,C4,P3,P4 ,O1,O2,A1, A2,F6,'FT7, FT8,FPZ ,plus two for	83%	Data Set of study

		eye blinking		
3	SVM and wavelet – ICA [XXII]	**#16** FP1,FP2, F3, Fz, F4, T7,C3, Cz, C4, T8, P3, Pz, P4, O1, Oz, O2	99.1%	Data Set of study
4	SVM , Wavelet transform, BAT, Binary BAT, [III]	**#16** Fz,FC1,FC2, C3,Cz,C4,CP5,CP1,CP2, CP6,P3,Pz,P4,O1 ,Oz,O2	94.11%	Data Set of study
5	***Sample entropy , graph entropy +SVM was proposed*** ***Sample entropy , graph entropy + KNN was proposed***	***#13 AF8,C1,C2, C3,C4,CP1, CP5,CP6,FC5,FT7,P8, PO8, PZ***	***90.8%*** ***83.7%***	***UCI Dataset***

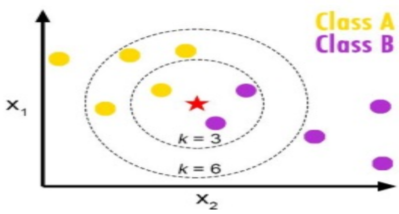

Fig.6: KNN Algorithm

V. Conclusion

The dataset utilized in this paper has been acquired from the UCI repository. Two ways applied ,the first way used 13 channels like (AF8 ,C1 ,C2 ,C3 ,C4,CP1,CP5 ,CP6,FC5,FT7,P8,PO8, and PZ) when used sample entropy and graph entropy with Shannon's entropy as feature extraction and SVM with RBF as classifier .Second way used the same channels with KNN classifier.The type of feature extraction techniques can play an effective role for getting the perfect quality features from EEG signal. . The outcomes are encouraging. The accuracy of personal identification with SVM and KNN are 90.8% and 83.7% respectively.Our experimental results showed that the SVM method is more reliable method than KNN in analyzing EEG signals as shown in Figure7 . Through this study the accuracy can be improved to give better results and more authentication bytrying to use other channels and use another types of signals in human body like ECG, EMG .

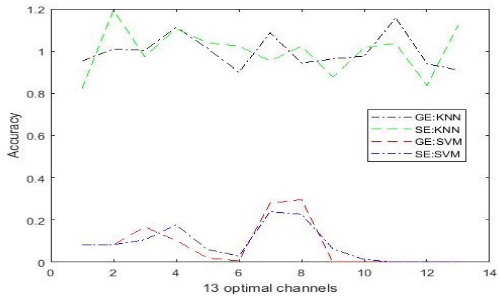

Fig. 7: Classification with 13 channels

References

I. Blake, C., and C. J. Merz. "UCI repository of machine learning databases, Department of Information and Computer Science, University of California, Irvine, CA, 1998." URL:< http://www. archive. ics. uci. edu/ml (2015).

II. Chong Yeh Sai, Norrima Mokhtar, Hamzah Arof, Paul Cumming, Masahiro Iwahashi," Automated Classification and Removal of EEG Artifacts With SVM and Wavelet-ICA", IEEE Journal of Biomedical and Health Informatics , Volume: 22 , Issue 3 , Page(s): 664 - 670 , MAY 2018.

III. Damastuti Natalia, Aisjah Aulia Siti , Masroeri Agoes A,Cassification of Ship-Based Automatic Identification Systems Using K- Nearest Neighbors"International Seminar on Application for Technology of Information and Communication, IEEE, pp 331-335, 2019.

IV. Deming Zhang, Guodong Yin, Weichao Zhuang, Xianjian Jin " Recognition Method for Multi-Class Motor Imagery EEG Based on Channel Frequency Selection "IEEE, 2018 .

V. Dharani, M., and M. Sivachitra. "Motor imagery signal classification using semi supervised and unsupervised extreme learning machines." 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS). IEEE, 2017.

VI. Duaa AIQattan, Francisco Sepulveda," Towards Sign Language Recognition UsingEEG-Based Motor Imagery Brain Computer Interface "5th International Winter Conference on Brain-Computer Interface (BCI), I EEE 2017, DOI: 10.1109/IWW-BCI.2017.7858143

VII. Gilbert, S., Dumontheil, I., Simons, J., Frith, C., Burgess, P.: Wandering minds: the default network and stimulus-independent thought. Sci. Mag. 315(5810), 393–395 ,2007

VIII. G. Paul and J. Irvine, "Fingerprint authentication is here but are we ready for what it brings?," IEEE Consumer Electronics Magazine, 2015.

IX. Helma, Christoph, Eva Gottmann, and Stefan Kramer. "Knowledge discovery and data mining in toxicology." *Statistical methods in medical research* 9.4 (2000): 329-358.

X. Kucyi, Aaron, et al. "Spontaneous default network activity reflects behavioral variability independent of mind-wandering." Proceedings of the National Academy of Sciences 113.48 (2016): 13899-13904..

XI. Middendorf, M., Mcmillan, G., Calhoun, G., Jones, K.S.: Brain computer interfaces based on steady state visual evoked response. IEEE Trans. Rehabil. Eng. 8(2), 211–214 ,2000.

XII. Minshew, N., Keller, T.: The nature of brain dysfunction in autism: functional brain imaging studies. Curr. Opin. Neurol. 23, 124–130 , 2010

XIII. Mustafa Turan Arslan, Server Göksel Eraldemir, Esen" Channel Selection from EEG Signals and Application of Support Vector Machine on EEG Data" IEEE, 2017.

XIV. Nivedha R, Brinda M, Devika Vasanth, Anvitha M and Suma K.EG based Emotion Recognition using SVM and PSO " IEEE, 2017.

XV. Rabie A. Ramadan and Athanasios V. Vasilakos," Brain Computer Interface: Control Signals Review "Neurocomputing, Volume 223, Pages 26–44. , 5 February 2017.

XVI. Raihan Khalil, Abdollah Arasteh, Ajay Krishno Sarkar "EEG BasedBiometrics Using Emotional Stimulation Data " IEEE Region 10 Humanitarian Technology Conference (R10-HTC), Dhaka, Bangladesh , 21 - 23 Dec 2017.

XVII. Rihman, J S, and J R Moorman. "Physiological Time-Series Analysis UsingApproximate Entropy and Sample Entropy." American journal of physiology. Heart and circulatory physiology 278(6): 2000,. http://www.ncbi.nlm.nih.gov/pubmed/10843903.,

XVIII. Syed Anwar, Tahira Batool, Muhammad Majid ," Event Related Potential (ERP)based Lie Detection using a Wearable EEG headset " IEEE 16th International Bhurban Conference on Applied Sciences & Technology (IBCAST), 2019.

XIX. Sivachitra, M., and S. Vijayachitra. "Planning and Relaxed State EEG Signal Classification Using Complex Valued Neural Classifier for Brain Computer Interface." PP 8-11, 2015.

XX. Shannon CE (2001) A mathematical theory of communication. ACM SIGMOBILE Mob Computer Common Rev 5(1):3–55.

XXI. S.Dhivya, A.Nithya, " A REVIEW ON MACHINE LEARNING ALGORITHM FOR EEG SIGNAL ANALYSIS " Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology, 2018.

XXII. Scholkopf, B., Sung, K., Burges, C., Girosi, F., Niyogi, P., Poggio, T.,Vapnik, V.,"Comparing Support Vector Machines with Gaussian Kernels to Radial Basis Function Classifers", N142,1996.

XXIII. V. Sankara Narayanan, R. Elavarasan, C.N. Gnanaprakasam, N. Sri Madhava Raja*, R. Kiran Kumar " Heuristic Algorithm based Approach to Classify EEG Signals into Normal and Focal" IEEE International Conference on System, Computation, Automation and Networking (ICSCA) page 1-5 , 2018 DOI: 10.1109/ICSCAN.2018.8541180.

XXIV. WAEL H. KHALIFA, MOHAMED I. ROUSHDY, ABDEL-BADEEH M. SALEM, " User Identification System Based on EEG Signals" The Sixth International Conference on Intelligent Computing and Information Systems (ICICIS).pp 14-16, 2013, Cairo, Egypt ,2013.

XXV. Yang, Z., Wang, Y., Ouyang, G.: Adaptive neuro-fuzzy inference system for classification of background EEG signals from ESES patients and controls. Sci. World J , 2014.

XXVI. Zhu, Guohun, Yan Li, Peng (Paul) Wen, and Shuaifang Wang"Analysis of Alcoholic EEG Signals Based on Horizontal Visibility Graph Entropy." (1–4): 19–25,2014.