FUZZY-ROUGH CLASSIFICATION FOR BRAINPRINT AUTHENTICATION

Siaw-Hong Liew[1], Yun-Huoy Choo[1] and Yin Fen Low[2]

ABSTRACT

The electroencephalogram (EEG) signal is used as biometric modality, because it is proven to be unique, universal and collectable. This work aims to assess the performance of fuzzy-based techniques for brainprint authentication modelling. We benchmark the performance of Fuzzy-Rough Nearest Neighbour (FRNN) technique to the Discernibility Nearest Neighbour (D-kNN) and the Fuzzy Lattice Reasoning (FLR) techniques using the selected samples of brainwaves' data from the original UCI EEG dataset. All the three classifiers are available in the fuzzy-rough version of WEKA implementation tool. Selected 9 EEG channels located at the midline and lateral regions were used in the experimentation. The coherence, mean of amplitudes and cross-correlation feature extraction methods were used to extract the EEG signals. The area under ROC curve (AUC) measurement of FRNN was promising against the D-kNN and FLR techniques. The FRNN model has achieved the best performance of AUC measure at 0.904 in opposition to the D-kNN and FLR models, where both recorded 0.770 and 0.563, respectively. However, the classification accuracy shows significantly no difference among the three classifiers. The results confirmed that the classification accuracy of D-kNN and FLR techniques is not reliable, because they are highly contributed by the true negative cases. Hence, we conclude that the FRNN model is less biased to imbalance data problem as compared to the D-kNN and FLR models. Future work of this research should focus on optimizing the EEG channel and feature selection in order to obtain a better data representation of biometric brainprint for more efficient authentication in imbalance data problem.

KEYWORDS

1. INTRODUCTION

The aim of brainprint authentication is to accept or reject the identity claimed by an individual. There are numerous types of person authentication methods, such as knowledge-based, token-based and biometric methods. The commonly used Personal Identification Number (PIN) and password are examples of knowledge-based authentication and signature is an example of token-based authentication. However, password and signature are considered the weakest authentication models, because the password can be stolen, while the signature can be forged easily. Biometric systems such as fingerprint, iris, face, voice and hand geometry authentication systems were introduced to overcome the security incompetency of the traditional authentication methods. Among all, fingerprint and face recognition are common modalities in today's biometric authentication systems. Fingerprint scheme [1] is widely used, but is still prone to forgery. This technology recognizes only the ridge arrangement on the finger surface, where intruders can easily replicate the fingerprint using silicon or gelatine to infringe the security systems. Facial recognition is also less promising, because the human face structure will change as a person ages. The above mentioned limitations can be overcome using a more secure biometric modality; the human brainprint. The brainprint extracted through electroencephalogram (EEG) signals is a highly secure biometric modality for person authentication. Over the recent years, EEG-based person authentication is catching much researchers' attention [2]-[3].

Various types of soft computing techniques have been applied in EEG signal classification. Artificial neural networks (ANNs), fuzzy logic, K-Nearest Neighbour (kNN), linear discriminant analysis (LDA) and support vector machine (SVM) are examples of soft computing techniques for EEG signal classification. Gui et al. [4] investigated visual evoked potential (VEP) data collection using a low-cost sensor system. ANNs were used for EEG-based biometric authentication and the classification accuracy achieved was around 90%. Back-propagation NN, SVM and LDA were used to classify the EEG signals

1. S. H. Liew & Y. H. Choo are with Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia. Emails: siawhong.liew@gmail.com and huoy@utem.edu.my
2. Y. F. Low is with Faculty of Electronics and Computer Engineering, Universiti Teknikal Malaysia Melaka (UTeM), Melaka, Malaysia. Email: yinfen@utem.edu.my

for biometric authentication and the accuracy ranged from 80.8% to 89.5% [5]. Fuzzy logic is an outstanding model, because it can solve real-world classification problems, since there is no precise boundary existing between the categories of problems [6]. Furthermore, the fuzzy set theory is very flexible and allows simple computation of evidences in complex relations.

Due to that EEG signals are non-stationary, a combination of fuzzy set theory, rough set theory and kNN called Fuzzy-Rough Nearest Neighbour (FRNN) is chosen to classify EEG signals. It is because the uncertainty of fuzzy rough set is able to deal with this problem. The classification results will be compared with those of other fuzzy-rough approaches, such as Discernibility Nearest Neighbour (D-kNN) and Fuzzy Lattice Reasoning (FLR).

The rest of the paper is organized as follows: Section 2 provides a literature review on EEG signals for person authentication, feature extraction and soft computing techniques. Section 3 describes the classification techniques. Section 4 outlines the experimentation, which includes data pre-processing and preparation, feature extraction, experimental setting, performance measures and statistical tests. Section 5 depicts experimental results and discussion. Last but not least, Section 6 draws the conclusions and indicates the direction of future work.

2. RELATED WORK

Berger was the first who recorded the EEG signals in 1929 [7]. EEG is defined as the electrical activity recorded from the scalp surface [8]. EEG signals are the electromagnetic waves that are emitted from the human brain's neurons. EEG is the most practical capturing method that can be used in biometrics due to the advances in its hardware devices. The EEG recording is a completely non-invasive procedure that can be repeatedly applied to normal adults, patients and children with virtually no risk or limitation [8]. The main advantages of using brain electromagnetic waves are: the uniqueness and liveness of the EEG signals, in addition to that the recorded brain responses cannot be replicated and the individual's identity cannot be stolen. A research work [9] showed that the individual's EEG signals vary from every individual to another, even though they performed similar task or thought. Conditions of stress, anxiety, fatigue, medication, drowsiness, environment, …etc. can increase the difficulty of reproducing similar pattern of EEG signals [10]. For example, a person that has been under the influence of stress will generate different EEG signals when compared to his/her normal state.

EEG recording electrodes and their function are critical for obtaining high-quality data for interpretation [8]. One important problem of EEG signal recording is the artifacts. Examples of artifacts occurring in EEG signal recording are: eye blinking, head movements, muscle activities and electrocardiogram (ECG). Due to the very low amplitude of EEG signals, artifacts often contaminate the recordings, restricting or making difficult analysis or interpretation. Therefore, the position of the subject during EEG recording should be very comfortable to avoid unnecessary activities; a lying position diminishes the existence of some artifacts caused by feeble motion. One of the ideas that combined EEG signals with authentication systems was proposed by Thorpe et al. [11]. The studied authentication system was designed by using pass-thought, which is reliable due to the uniqueness of EEG signals. Apart from that, a consumer grade of EEG headset was used in Ashby et al. [12] for authentication purpose.

Marcel and Millán [13] achieved a high authentication performance of 93.4% in terms of accuracy. A total of 9 normal subjects were asked to perform 3 tasks (i.e., left-hand movement, right-hand movement and generation of words that begin with the same random letter) during 12 non-feedback sessions in 3 days, which means 4 sessions per day. The classification accuracy reached around 80% in the research work [14]. They analyzed the 8-channel EEG signals from a group of 40 volunteers who performed a simple experiment (i.e., relaxing with opened and closed eyes). In addition, the research work by Jian-Feng [15] used the BCI competition 2003 EEG dataset that was recorded from a total of 64 channels and sampled at 250 Hz. The authentication classification result ranged from 75% to 85%. Biometric authentication based on EEG signals conducted in [16] covered three tasks of classification accuracy of: reading task (97.3%), relax task (94.4%) and multiplication task (97.5%). The research work in [17] combined EEG headsets with the smartphone for EEG-based person authentication purpose. Besides, an EEG-based biometric authentication system was developed in [18]. The EMOTIV Epoch+ EEG headset was used to collect the EEG signals and the classification accuracy achieved was 96.97%.

Mean corresponds to the centre of a set of values. It is a time domain feature, which is calculated for the

reconstructed EEG signal amplitude and time duration. Mean has been used in [19] as one of the features for the filtered signals. The extracted signals are then distinguished to be normal or epileptic by using artificial neural network technique. Other than that, time domain features, such as mean, median, mode, standard deviation, minimum and maximum, were used in [20] in the analysis of EEG signals to detect brain abnormalities. Correlation is very similar to convolution, which is a mathematical operation. Then, cross-correlation is the measurement of the extent of similarity relationship between two signals. It is able to detect non-stationarity and is widely used for the analysis of time series of EEG signals. A set of five features is extracted and then utilized for training an SVM classifier to generalize the results. In this research, a healthy subject sample signals are acquired to represent a reference signal for data comparison. Cross-correlation has been aided to Support Vector Machine (SVM) classifier in EEG signal classification [21]. With the aid of cross-correlation, SVM is able to perform better in pattern recognition. The accuracy achieved was 94.5%. Hence, cross-correlation is a very useful technique to gain insight in EEG signals for feature extraction. Coherence is one of the feature extraction methods that is widely used for EEG signal analysis. Coherence is a linear correlation measure between two signals at different frequencies. It was first used as a feature in [22] for measuring the mean coupling between signals recorded from an electrode and its neighbours. In addition, mutual information, coherence and cross-correlation have been used in [23] for an EEG biometric system. The features extracted from the EEG signals have been proven unique enough among subjects for biometric applications. Research work in [23] used an unobtrusive authentication method that uses 2 frontal electrodes and 1 reference electrode placed at the left ear lobe only.

Due to low signal-to-noise ratio and non-stationarity of EEG signals, uncertainty modelling tools, such as fuzzy set and rough set, are needed to handle the related problems. Fuzzy set theory [24] and rough set theory [25] are good solutions in handling uncertainty and manipulating incomplete data. Fuzzy-rough set provides a higher degree of flexibility in dealing with imprecision and vagueness existing in real-world data [26]–[30]. Fuzzy-Rough Nearest Neighbour (FRNN) model introduced by Jensen and Cornelis [31] is hybridized with the strength of fuzzy-rough set and Fuzzy Nearest Neighbour (FNN) approach to complement each other. The constructed fuzzy lower and upper approximations are used to avoid the use of fuzzy logical connectives altogether. However, fuzzy-rough set allows that the element belongs to more than one class. In addition, FNN model is an extension version of kNN algorithm to fuzzy set theory and it is proved that FNN model outperforms the standard nearest neighbour model [32]. FNN model allows partial membership of an object in different classes and takes into account the closeness of each neighbour with respect to the test instance. Unfortunately, FNN algorithm is found out to have a problem when dealing with imperfect data. Therefore, the hybridization of the strength between fuzzy-rough set and FNN algorithm, which is fuzzy-rough nearest neighbour (FRNN) algorithm, can allow both to complement each other in order to gain good performance.

FRNN algorithm uses nearest neighbours to compute fuzzy lower and upper estimations in order to predict the test objects [31]. With the existing of the fuzzy approximations, the FRNN algorithm outperforms other nearest neighbour approaches and Naïve Bayes prediction models in classification problems. This is proven from the experiment done in Sarkar [33]. Three nearest neighbour approaches; namely, conventional kNN algorithm, the FNN algorithm and the FRNN algorithm, were used to classify Wisconsin Breast Cancer problem [34]. The dataset consisted of 699 samples and each sample provided ten numerical attributes. A total of 16 samples with missing attributes were removed from the dataset. From this experiment, FRNN algorithm gained the highest classification performance among the three algorithms. Moreover, the time complexity of FRNN algorithm is the same as those of the conventional kNN algorithm and the FNN algorithm. Furthermore, the FRNN algorithm was applied in China Stock Market Distressed Company for prediction problems [34]. The FRNN algorithm is able to use unbalanced and unmatched training and testing datasets in prediction. The prediction accuracy achieved was 78.37% which is better than that of the FNN classification approach. This study concluded that the FRNN approach performs better than the conventional kNN approach and the FNN approach. FRNN approach not only can deal with unbalanced data, but also performs well when dealing with incomplete data.

D-kNN approach is an extension of the kNN algorithm which uses the concept of discernibility. D-kNN computes the discernibility of the neighbours and the distances from the test objects. The main benefit in D-kNN approach is that it does not allow the classes of dataset to overlap. D-kNN algorithm considers the structural properties of the neighbours. A comparison of performance among three classifiers with

nearest neighbour approach was carried out by using Bupa Liver dataset [35]. The classifiers are conventional kNN algorithm, Weighted kNN (W-kNN) algorithm and D-kNN algorithm. By comparing the classifiers, D-kNN yields the best classification accuracy and net reliability. However, the processing time is slightly longer than in the conventional kNN and W-kNN algorithms. W-kNN yields the worse in terms of accuracy, because the dataset only contains six features and all of these are equally importance. W-kNN performed well when the dataset contains a larger number of features.

FLR is a rule-based classifier. The term "fuzzy lattice" was introduced by Naseem in 1994 [36] on the concept of fuzzy partial-order relation. The benefit of the lattice theory is capability of tackling with uncertain information and dealing with missing data [37]. Fuzzy lattices can be used in classification and clustering algorithms and have been successfully implemented in real-time problems, such as pattern recognition [38], air quality assessment and ambient ozone estimation [39]. FLR classifier was successfully applied in ambient ozone estimation [39] and the results with missing values and without missing values were compared. The FLR classifier gained similar values in terms of accuracy measure for the dataset with missing values and without missing values. The classification accuracy was 84.6% for dataset with missing values and 83.23% for dataset without missing values. Furthermore, the least time for training and testing was taken by FLR classifier. It used only around 1.5 seconds, while back-propagation neural networks took a training and testing time between 3 minutes and 25 minutes. In recent years, the FLR classifier has been used for image recognition, such as human facial expressions [40]-[41]. However, there is still lack of research on EEG signal classification using FLR.

Data pre-processing and feature extraction are the important steps in order to perform FLR classification. Seven different facial expressions; namely, neutral, angry, disgust, feared, happy, sad and surprised, were recorded. The dataset was divided into 75% of training data and 25% testing data. From this experiment, FLR classifier performed better than the conventional kNN algorithm [40].

3. CLASSIFICATION

In this study, FRNN, D-kNN and FLR techniques were used to accomplish brainprint authentication modelling. Brainprint authentication modeling consists only of 2 classes; client and impostor. The FRNN, D-kNN and FLR techniques can be found in fuzzy-rough version of Waikato Environment for Knowledge Analysis (WEKA). It is free downloaded from http://users.aber.ac.uk/rkj/book/wekafull.jar.

3.1 Fuzzy-Rough Nearest Neighbour (FRNN)

Fuzzy-Rough Nearest Neighbour (FRNN) was introduced by Jensen and Cornelis [31] in 2011. It is a hybrid model with the combination of fuzzy set, rough set and nearest neighbour classification approaches. In the FRNN algorithm, the lower and upper approximations are constructed by the nearest neighbours to allocate the decision class to the test object. The details of FRNN algorithm can be found in Algorithm 1 [31]. The FRNN algorithm calculates the similarity between the two objects and finally classifies the test objects into the most possible decision classes. FRNN classifies the test object based on single nearest neighbour with the highest similarity measure. Therefore, the value of k does not affect the classification performance. The FRNN technique captures uncertainty by using fuzzy-rough approximations. The construction of fuzzy upper and lower approximations is to avoid the use of fuzzy logical connectives completely. The connectives here are the keys in developing the fuzzy-rough set theory.

3.2 Discernibility Nearest Neighbour (D-kNN)

Discernibility Nearest Neighbour (D-kNN) classifier can handle overlapping classes of a dataset compared to the original kNN. The discernibility of the neighbours was first calculated, followed by their distances from the test objects. The algorithm of D-kNN is shown in Algorithm 2 [35]. The property of the neighbours is playing an important role in D-kNN prediction [35]. The ratio or distance of discernibility is computed for each neighbour data and the average of the ratios is taken for each class. D-kNN not only classifies the test elements based on the concept of nearest neighbours, but also based on the discernibility scores. The discernibility score of D-kNN classifier is produced for each object to be classified. After that, the average of the discernibility scores of the neighbouring objects and their distances from the objects are calculated for each one of the possible classes. Then, S_j is calculated for the classification score of each class. Eventually, the classification scores of the different classes are

compared in order to classify the test objects. The higher the classification score, the higher the chance to be the output of the classification.

Algorithm 1 Fuzzy-Rough Nearest Neighbour (FRNN) algorithm

1: **Input:** X, the training data; $\mathcal{C}$, the set of decision classes; y, the object to be classified

2: **Output:** Classification for y

3:   **begin**

4:     $N \leftarrow getNearestNeighbours(y, k)$

5:     $\tau \leftarrow 0, Class \leftarrow \emptyset$

6:     **foreach** $C \in \mathcal{C}$ **do**

7:       **if** $\left(\left(min_{a \in \mathbb{A}} \left(1 - \frac{|a(x) - a(y)|}{|a_{max} - a_{min}|} \right) \downarrow C \right)(y) + \left(min_{a \in \mathbb{A}} \left(1 - \frac{|a(x) - a(y)|}{|a_{max} - a_{min}|} \right) \uparrow C \right)(y) \right)/2 \geq \tau$

        then

8:        $Class \leftarrow C$

9:        $\tau \leftarrow \left(\left(min_{a \in \mathbb{A}} \left(1 - \frac{|a(x) - a(y)|}{|a_{max} - a_{min}|} \right) \downarrow C \right)(y) + \left(min_{a \in \mathbb{A}} \left(1 - \frac{|a(x) - a(y)|}{|a_{max} - a_{min}|} \right) \uparrow C \right)(y) \right)/2$

10:      **end**

11   **end**

12:     **output** $Class$

13:   **end**

Algorithm 2 Discernibility Nearest Neighbour (D-kNN) algorithm

---

1: **Input:** Input train objects (P), labels of train objects (T), number of neighbours (k), test objects (PT)

2: **Output:** Classification vector of test object (y)

3:   Initialization: set $i = 0$ (index of the objects), $j = 0$ (index over the classes)

4:   n $\leftarrow$ number of test objects (PT)

5:   N $\leftarrow$ number of train objects (P)

6:   z $\leftarrow$ discernibility vector for elements of P, using Algorithm-1 and P, T as inputs; $z = \{z_i\}, i = 1 \dots N$

7:   q $\leftarrow$ number of unique values of T (classes)

8:   do $i \leftarrow (i + 1)$

9:     **D** $\leftarrow$ vector of distances of PI (i) to P

10:    **sd** $\leftarrow$ sorted values of d

11:    **dk** $\leftarrow k$ first values of **sd**

12:    **v** $\leftarrow \{v_m : v_m = z_m/dk_m\}, m = 1 \dots k$

13:    do $j \leftarrow (j + 1)$ for current i

14:    $C_j \leftarrow$ subset of k nearest elements of P belonging to the j-th class

15:    Classification score $S_j \leftarrow mean\left(v_{c_j} \right), v_{c_j} = \{v_m : \forall m \in C_j\}$

16:   until $j = q$

17:    $b \leftarrow argmax S_j$

18:    $y_i \leftarrow$ class b

19:   until $i = n$

3.3 Fuzzy Lattice Reasoning (FLR)

Fuzzy Lattice Reasoning (FLR) is a classifier to extract rules from the input data based on fuzzy lattices. The sequence of the input data representation is important. FLR plays an important role in dealing with different types of data; for example, fuzzy sets, real vectors, images, symbols, graphs and waves. Other than that, FLR can deal with both points and intervals. Apart from that, FLR has the ability of knowledge representation and is capable of extracting implicit features beyond the data, which can represent the data as rules. Furthermore, FLR has the ability to combine different types of data, handle missing data and cope with both complete and incomplete lattices. FLR manages missing data by replacing them with least element O and great element I, respectively. For example, if the constituent lattice is $([0,1], \leq)$, then we can replace the missing data by intervals of $O = [1,0]$ and $I = [0,1]$, respectively [37].

Algorithm 3 Rule Induction by the Fuzzy Lattice Reasoning (FLR) algorithm

---

1: A rule base $RB = \{A_1 \rightarrow C_1, \dots, A_L \rightarrow C_L\}$ is given – Note that RB could initially be empty, i.e. $L = 0$.

2: Present the next input rule $a_i \rightarrow c_i, i = 1, \dots, n$ to the initially 'set' family of rules in RB.

3: If no more rules in RB are 'set' then

4:     Store input rule $a_i \rightarrow c_i$ in RB;

5:     $L \leftarrow L + 1$;

6:     Goto 2,

7: Else, compute the fuzzy degree of inclusion $k(a_i \leq A_l), l \in \{1, \dots, L\}$ of antecedent a_i to the antecedents of all the 'set' rules in RB

8: Competition among the 'set' rules in RB.

9:     Winner is rule $A_J \rightarrow C_J$ such that $J = \arg \max_{l \in \{1,\dots,L\}} k(a_i \leq A_l)$.

10: The *Assimilation Condition*: Both $diag_1(a_i \vee A_J)$ is less than a maximum user-defined threshold size D_{crit} and $c_i = C_J$.

11: If the *Assimilation Condition* is satisfied then

12:     Replace the antecedent A_J of the winner rule $A_J \rightarrow C_J$ by the join-interval $a_i \vee A_J$.

13: Else, 'reset' the winner rule $A_J \rightarrow C_J$.

14: Goto 3,

Algorithm 4 Generalization by the Fuzzy Lattice Reasoning (FLR) algorithm

---

1: Consider a rule base $RB = \{A_1 \rightarrow C_1, \dots, A_L \rightarrow C_L\}$.

2: Present a rule antecedent 'a_0' for classification based on the rule base RB.

3: Calculate the fuzzy degree of inclusion $k(a_0 \leq A_l), l \in \{1, \dots, L\}$ of antecedent 'a_0' in the antecedents 'A_l', $l = 1, \dots, L$ of all rules in RB.

4: Competition among the rules in RB.

5:     Winner is rule $A_J \rightarrow C_J$ such that $J = \arg \max_{l \in \{1,\dots,L\}} k(a_i \leq A_l)$.

6: The antecedent a_0 is classified to the class with label C_J.

(a_i, C_L) is the representation for the input datum to the FLR model, where C_L represents the class label of datum a_i and can be interpreted as a rule "if a_i then C_L". An input datum (a_0, C_0) is presented to the network in the learning phase. The degree of inclusion between input and stored rules in RB will be calculated as $k(a_0, a_1), \dots, k(a_0, a_c)$. The FLR will choose the rule with $\arg \max_{l \in \{1,\dots,L\}} k(a_i \leq A_l)$ as the winner rule. If the winner rule A_J and input datum a_0 have the same class label and the size of $a_i \vee A_J$ is less than a user-defined threshold, then the winner rule will be updated. There is only one parameter that can be tuned in FLR; that is the threshold size, D_{crit}. D_{crit} is used to indicate the maximum size of a

hyperbox to be learned. Larger values of D_{crit} will result in more generalized rules while smaller values of D_{crit} will result in more specific rules.

4. EXPERIMENTATION

EEG signal classification is a difficult task as a result of that the characteristics of EEG signals are non-stationary, in addition to high dimensionality and low signal-to-noise ratio (SNR). Thus, data pre-processing and data preparation steps are important.

4.1 Data Pre-processing and Data Preparation

In this study, a free EEG dataset is used which can be taken from UCI Machine Learning Repository [42]. The online available EEG dataset consists of three versions, which are small dataset, large dataset and full dataset. Each of the datasets contains an individual, 10 individuals and 122 individuals, respectively. The UCI EEG dataset was recorded for both alcoholic and non-alcoholic persons. Since this study focused on person authentication modeling, only non-alcoholic dataset will be used. Alcoholic data is not suitable for this study, because data collected from alcoholic persons might be less accurate due to their brains having been affected by alcohol. Large dataset will be used in this study, but one of the individuals from the large dataset will be replaced by an individual from the full dataset. This is because there are many redundant trials by the individuals which will affect the result. Each individual accomplished 60 trials. This EEG dataset consists of the measurements of 64 electrodes (61 active electrodes + 3 reference electrodes) placed on the scalps and the sampling rate was at 256Hz.

The stimuli were composed of 90 images that were chosen from a total of 260 black-and-white Snodgrass and Vanderwart image set [43]. The subjects were requested to recognize the image as soon as the image is displayed on the computer screen. The distance of the computer screen from the subject's eyes was 1 meter. The image remained on the screen for 300ms and the Inter-Stimulus Interval (ISI) for each test was set to 3200ms. The visual stimulus presentation is illustrated in Figure 1.
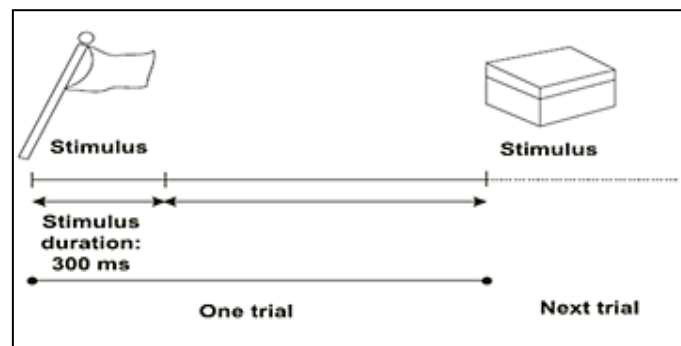

Figure 1. Visual stimulus presentation.

In the general machine learning model building, there are some common suggestions for train/test splits, such as 60/40, 70/30, 80/20 or even 90/10, if the dataset is relatively large [34]. The higher percentage of train data tends to generate a better model, but sacrifices the objectivity of test results due to low number of test data. Therefore, the larger the dataset, the higher the train/test proportion which may be applied. However, machine learning experiment seldom implements the 90/10 proportion, unless the dataset used is extremely large. We used the 80/20 train/test proportion in this study, where 480 instances were used for model building *versus* 120 instances for model testing.

In this stage, we have checked the trials in order to avoid the redundant trials between train set and test set. The dataset has equivalent distribution of trials between S1 object, S2 match and not-match for both train set and test set. In this study, we only selected 100 data points, which corresponds to approximately 300 milliseconds (ms). This is because the VEP normally occurred within the first 300ms. Besides, the EEG signals of S2 were different from those of S1 due to that S2 involves brain information about the match or not-match analysis. Only the electrodes located at the midline and lateral sides were considered. This is due to that midline and lateral electrodes provide stronger strength from electrical signals when responding to visual stimuli [44]. The lateral electrodes are PO7, PO8, O1, O2 and OZ, while the midline electrodes are FPZ, FZ, CZ, PZ and OZ.

4.2 Feature Extraction

A set of feature vectors were retrieved from the raw EEG dataset. The extracted feature vectors act as a different observation for the purpose of classification. Besides, feature extraction can reduce the dimensions of the input attributes as compared to the raw EEG dataset. In this study, coherence, cross-correlation and mean of amplitudes are selected from a particular literature review. The three feature extraction methods are described as follows:

a) Coherence: Coherence is used in order to compute the degree of linear correlation between two signals. The correlation between two signals at different operating frequencies can be revealed by coherence [17]. EEG-based coherence analysis is proven to be suitable for use in biometrics [45]. Coherence is ranging from 0 to 1, where the value of 0 indicates that the two signals are independent, while the value of 1 indicates that the two signals are completely linearly dependent. The coherence is calculated as follows:

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f)P_{yy}(f)} \tag{1}$$

where,

$C_{xy}(f)$ is a function of the power spectral density (P_{xx} and P_{yy}) of x and y and the cross-power spectral density (P_{xy}) of x and y.

b) Cross-correlation: Cross-correlation, known as a sliding dot product, is used to compute the similarity between two signals. It is also frequently used to obtain the existence of a known signal sequence in an unknown one. It is a function of the relative delay between the signals and the application in pattern recognition. Two input signals will be calculated for cross-correlation:

Channel 1 with itself: ρ_X,

Channel 2 with itself: ρ_Y,

Channel 1 with channel 2: ρ_{XY}.

The correlation ρ_{XY} between two random variables x and y with expected values μ_X and μ_Y and standard deviation σ_X and σ_Y is given as:

$$\rho_{X,Y} = \frac{cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E\big((X - \mu_X)(Y - \mu_Y)\big)}{\sigma_X \sigma_Y} \tag{2}$$

where, $E(\ )$ is the expectation operator and $cov(\ )$ is the covariance operator.

c) Mean of Amplitudes: Mean, also known as average, is the sum the of all EEG potential values divided by the number of data points. The mean is calculated as follows:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i \tag{3}$$

where, n is the number of data points and x_i is the value of the data.

4.3 Experimental Setting

In FRNN algorithm, fuzzy logic connectives are crucial for developing the fuzzy-rough set theory. A triangular norm (t-norm), $\mathcal{T}$ is any increasing, associative and commutative $[0,1]^2 \rightarrow [0,1]$ mapping satisfying $\mathcal{T}(1, x) = x$, for all x in $[0,1]$. On the contrary, an implicator is any $[0,1]^2 \rightarrow [0,1]$ mapping ℓ satisfying $\ell(0,0) = 1, \ell(1, x) = x$, for all x in $[0,1]$. Based on the [31], the Kleene-Dienes implicator for x, y value in $[0,1]$ was implemented. In addition, the experimental setting for D-kNN was the same with FRNN classifier. Kleene-Dienes was chosen for both t-norm and implicator. Moreover, there is only one parameter that can be tuned in FLR algorithm; that is the threshold size ρ. We have set $\rho = 0.1$ [38] in our experiment.

4.4 Performance Measures and Statistical Test

The experimental result is analyzed based on the accuracy and the area under the receiver operating characteristics (ROC) curve (AUC). The AUC measure is used as one of the performance measures in

117

Jordanian Journal of Computers and Information Technology (JJCIT), Vol. 05, No. 02, August 2019.

this study, because it is more reliable and statistically consistent as compared to the accuracy measure in [46]. The accuracy and AUC of FRNN will be compared with the results obtained from D-kNN and FLR. The purpose of this comparison is to test whether FRNN can perform better than other classification algorithms, such as D-kNN and FLR.

Beforehand, the normality distribution of data is verified by using the Anderson-Darling test. The Anderson-Darling test [47] is modified from Kolmogorov-Smirnov (K-S) test. By comparing to the (K-S) test, Anderson-Darling test contributes more weights to the distribution tails. The critical value is calculated for the specific distribution. The Anderson-Darling test is calculated as:

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F^*(x)]^2 \psi \, (F^*(x) dF^*(x) \tag{4}$$

where, ψ = non-negative weight function which can be defined from:

$$\psi = F^*(x)(1 - F^*(x))^{-1} \tag{5}$$

The normality distribution of data must be determined before performing a statistical test. A statistical test is performed in order to determine the confidence level of the dataset which lead to reaching conclusions. Parametric test is chosen when data is normally distributed, while non-parametric test will be chosen when data is not normally distributed. Parametric tests, such as Z-test, paired-sample t-test or F-test, will yield higher accuracy when data is normally distributed. Simultaneously, if data is normally distributed and a non-parametric test is performed, then the results will not be as accurate as in the case of parametric test.

From the normality test using Anderson-Darling test, the accuracy of FRNN, accuracy of D-kNN and AUC of D-kNN are normally distributed, while the accuracy of FLR, AUC of FRNN and AUC of FLR are not normally distributed. Therefore, a paired-sample t-test is performed between accuracy of FRNN and accuracy of D-kNN. In contrast, Wilcoxon signed-rank test is performed when the results are not normally distributed.

A paired-sample t-test is performed to compare the differences of means between paired observations by using the IBM SPSS Statistics 22. The paired-sample t-test is a statistical validation method which is used to compare the means from different sources in a dataset [48]. The reason behind this is to investigate the significance differences between two groups. The null hypothesis of paired-sample T-test states that the difference between two mean values is zero, which is represented as:

$$H_0: \mu_1 - \mu_2 = 0 \tag{6}$$

On the other hand, Wilcoxon signed-rank test is frequently used for non-parametric testing. It is an alternative method for paired-sample t-test. The Wilcoxon signed-rank test is used to evaluate the difference of medians between paired data. Wilcoxon signed-rank test is more powerful in distinguishing the differences between two samples [49]. The nominal data cannot be analyzed with Wilcoxon signed-rank test, because the difference of the nominal data points has no specific value.

In statistical test, the null hypothesis is rejected if and only if the p-value is less than 0.05, which means that there are statistically significant differences between the two samples. On the contrary, the null hypothesis is accepted if and only if the p-value is larger than 0.05, which means that there are no statistically significant differences between the two samples. A statistical test was performed to test the differences between the two classifiers for different cases of use at 95% confidence level.

5. EXPERIMENTAL RESULTS AND DISCUSSION

Table 1 shows the individual EEG signals' authentication performance for 10 subjects. D-kNN model ranked first with an accuracy of 92.09%, while the FLR model came second with an accuracy of 90.67%. Meanwhile, the FRNN model had the lowest classification accuracy, recorded to be 90.17%, as compared to the D-kNN model and the FLR model. The highest accuracy recorded for FRNN model was up to 99.17%, while the lowest accuracy was 80.83% only. On the other hand, the accuracy of D-kNN model was considerably high, where the highest value recorded was 98.33% and the lowest was 89.17%. Besides, the accuracy of FLR model was between 88.33% and 94.17%. In overall, the accuracy of FRNN model, D-kNN model and FLR model indicated good classification results.

Table 1. Comparison of accuracy and AUC between FRNN, D-kNN and FLR.

Subject	FRNN		D-kNN		FLR	
	Accuracy	AUC	Accuracy	AUC	Accuracy	AUC
Subject 1	87.50	0.924	91.67	0.735	92.50	0.773
Subject 2	86.67	0.788	89.17	0.661	90.83	0.579
Subject 3	88.33	0.922	91.67	0.758	90.00	0.500
Subject 4	80.83	0.704	89.17	0.488	88.33	0.491
Subject 5	93.33	0.954	94.17	0.880	90.00	0.500
Subject 6	88.33	0.924	90.00	0.814	90.00	0.500
Subject 7	99.17	1.000	98.33	1.000	94.17	0.708
Subject 8	90.83	0.895	91.67	0.733	90.00	0.500
Subject 9	90.00	0.936	90.83	0.758	90.00	0.500
Subject 10	96.67	0.990	94.17	0.875	90.83	0.579
Average	**90.17**	**0.904**	**92.09**	**0.770**	**90.67**	**0.563**

The FRNN model has achieved the highest average AUC with 0.904. However, the average AUC has been found to be 0.770 in D-kNN model followed by FLR model with 0.563. Thus, it was shown that the FRNN model outperformed the D-kNN model and the FLR model. In the FRNN model, the highest AUC achieved was a perfect 1.00 score, while the lowest was 0.704. Comparatively, the highest AUC for D-kNN model also achieved the perfect score 1, but 0.488 was the lowest AUC in this model. Lastly, 0.773 was the highest AUC in FLR model, while 0.491 was the lowest AUC in FLR model.

From the performance measure in terms of accuracy, D-kNN model is slightly higher than FRNN and FLR models. On the other hand, FLR model showed the worst result in AUC. From Table 1, we can visually observe that most of the results are around 0.5, which is rated as worst classification performance. The model is not capable of distinguishing between positive class and negative class. This result explained that the FLR model is not suitable to classify EEG signals for person authentication. The possible reason is the parameter setting of FLR model. As previously described, there is only one parameter which can be tuned; that is the size of threshold. As the parameter setting for FLR model is 0.1 [38], therefore it will affect the AUC obtained, since the perspective of AUC is different from that of accuracy. The larger values of threshold will result in more generalized rules [39]. The threshold used in this project is small, which is 0.1; therefore, the rules are more specific as smaller values of threshold will result in more specific rules. Thus, the AUC of FLR models is lower than in FRNN and D-kNN models.

By observing the overall classification results above, FRNN gained good performance in terms of accuracy and AUC compared to D-kNN and FLR models. As previously described, FRNN algorithm is a fusion model that combines the strength fuzzy-rough set and the FNN approach. The decision class is determined by using the fuzzy lower and upper approximations to compute the membership value of a test object [31]. The fuzzy lower and upper approximations play a crucial role in dealing with noisy data such as EEG signals. Hence, the FRNN is able to perform better.

Table 2. Statistical test for comparison of accuracy and AUC between FRNN, D-kNN and FLR.

Performance Measure	Mean	p - Value	Statistical Test
FRNN Accuracy	90.17	0.071	No significant differences
D-kNN Accuracy	92.09		
FRNN Accuracy	90.17	0.767	No significant differences
FLR Accuracy	90.67		
FRNN AUC	0.904	0.004	Significant differences
D-kNN AUC	0.770		
FRNN AUC	0.904	0.006	Significant differences
FLR AUC	0.563		

Table 2 shows the statistical test for the comparison of accuracy and AUC among FRNN, D-kNN and FLR models. The paired-sample t-test was only used for the comparison between the accuracy of FRNN model and the accuracy of D-kNN model. The p-value of this comparison is 0.071, which is greater than 0.05. Thus, we can conclude that the FRNN model and the D-kNN model are not significantly different. On the contrary, Wilcoxon signed-rank test was used for the rest of the comparisons. The p-value for the comparison between the accuracy of FRNN model and that of FLR model was recorded at 0.767, which is greater than 0.05. Hence, there are also no significant differences between the FRNN model and the FLR model. In summary, accuracy comparisons did not show significant differences among the models.

Comparatively, the statistical test for the comparison between the AUC of FRNN model and D-kNN model showed significant differences with a p-value of 0.004. From the mean values in Table 2, it is clearly proved that the AUC of FRNN model is higher than the AUC of D-kNN model. Thus, we can conclude that the FRNN model performed better than the D-kNN model. Furthermore, a statistical test was also carried out for the comparison between the AUC of FRNN model and that of FLR model. The p-value was recorded at 0.006 and indicating significant differences in the paired set. The AUC of FRNN model achieved a value of 0.904, while the AUC of FLR model achieved only 0.563, which is considered a poor result. In other words, the FRNN model significantly performed better than the D-kNN model and the FLR model.

6. CONCLUSIONS

Among the fuzzy set and rough set approaches, the FRNN model is proven to be significantly better than D-kNN model and FLR model in EEG signal classification for brainprint authentication modeling. The AUC of FRNN model is 0.904, which is considered an excellent classification result. However, further work should be done on the FRNN model to improve the accuracy and AUC, since a good authentication system should have a perfect classification. The classification results gained from the FRNN model are more stable and consistent as compared to the classification results of D-kNN model and FLR model. This study showed the importance and capability of fuzzy-rough approximations of handling uncertain and non-stationary signals.

ACKNOWLEDGEMENTS

REFERENCES

[1]   P. Rajeswari, S. V. Raju, A. S. Ashour and N. Dey, "Multi-fingerprint Unimodel-based Biometric Authentication Supporting Cloud Computing," Intelligent Techniques in Signal Processing for Multimedia Security, Springer, Cham, pp. 469–485, 2017.

[2]   S. Yang and F. Deravi, "On the Usability of Electroencephalographic Signals for Biometric Recognition: A Survey," IEEE Trans. Human-Machine Syst., vol. 47, no. 6, pp. 958–969, 2017.

[3]   K. P. Thomas and A. P. Vinod, "Toward EEG-based Biometric Systems: The Great Potential of Brain-Wave-based Biometrics," IEEE Systems, Man, & Cybernetics Magazine, pp. 6–15, 2017.

[4]   Q. Gui, Z. Jin and W. Xu, "Exploring EEG-based Biometrics for User Identification and Authentication," Proc. of the IEEE Signal Processing in Medicine and Biology Symposium (SPMB), pp. 1–6, 2014.

[5]   H. U. Jian-Feng, "Comparison of Different Classifiers for Biometric System Based on EEG Signals," Proc. of the 2nd International Conference on Information Technology and Computer Science, pp. 288–291, 2010.

[6]   O. F. Bay and A. B. Usakli, "Survey of Fuzzy Logic Applications in Brain-related Researches," Journal of Medical Systems, vol. 27, no. 2, pp. 215–223, 2003.

[7]   H. Berger, "Ueber das Elektroencephalogramm des Menschen," J. für Psychol. und Neurol., vol. 87, pp. 527–570, 1929.

[8]   M. Teplan, "Fundamentals of EEG Measurement," Measurem. Sci. Review, vol. 2, no. 2, pp. 1–11, 2002.

[9]   A. Zuquete, B. Quintela and J. P. Silva Cunha, "Biometric Authentication using Brain Responses to Visual

Stimuli," Proc. of the International Conference on Bio-inspired Sys. and Sig. Proces., pp. 103–112, 2010.

[10]   I. Svogor and T. Kisasondi, "Two-factor Authentication Using EEG Augmented Passwords," Proc. of the 34th International Conference on Information Technology Interfaces, pp. 373–378, 2012.

[11]   J. Thorpe and P. C. Van Oorschot, "Pass-thoughts : Authenticating with Our Minds," Proceedings of the 2005 Workshop on New Security Paradigms, pp. 45–56, 2005.

[12]   C. Ashby, A. Bhatia, F. Tenore and J. Vogelstein, "Low-cost Electroencephalogram (EEG) Based Authentication," Proc. of the 5th Int. IEEE/EMBS Conf. on Neural Engineering, pp. 442–445, 2011.

[13]   S. Marcel and J. D. R. Millán, "Person Authentication Using Brainwaves (EEG) and Maximum *A Posteriori* Model Adaptation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 29, no. 4, pp. 743–752, Apr. 2007.

[14]   R. B. Paranjape, J. Mahovsky, L. Benedicenti and Z. Koles', "The Electroencephalogram As a Biometric," Proc. of the Canadian Conference on Electrical and Computer Engineering, vol. 2, pp. 1363–1366, 2001.

[15]   J.-F. Hu, "Multi-feature Biometric System Based on EEG Signals," Proc. of the 2nd Int. Conf. on Interaction Sciences Information Technology, Culture and Human (ICIS '09), pp. 1341–1345, 2009.

[16]   C. R. Hema, M. P. Paulraj and H. Kaur, "Brain Signatures: A Modality for Biometric Authentication," Proc. of the International Conference on Electronic Design, pp. 1–4, 2008.

[17]   H. Olesen, J. Klonovs and C. K. Petersen, Development of a Mobile EEG-based Feature Extraction and Classification System for Biometric Authentication, Master Thesis, Aalborg Univ., Copenhagen, 2012.

[18]   I. Jayarathne, M. Cohen and S. Amarakeerthi, "BrainID : Development of an EEG-based Biometric Authentication System," Proc. of the 7th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2016.

[19]   S. Garg and R. Narvey, "Denoising and Feature Extraction of EEG Signals Using Wavelet Transform," International Journal of Engineering, Science and Technology (IJEST), vol. 5, no. 6, pp. 1249–1253, 2013.

[20]   M. Kalaivani and V. A. Devi, "Analysis of EEG Signal for the Detection of Brain Abnormalities," Proceedings of the International Conference on Simulations in Computing Nexus (IJCA), pp. 1–6, 2014.

[21]   S. Chandaka, A. Chatterjee and S. Munshi, "Cross-correlation Aided Support Vector Machine Classifier for Classification of EEG Signals," Expert Systems with Applications Journal, vol. 36, no. 2, pp. 1329–1336, Mar. 2009.

[22]   E. Gysels and P. Celka, "Phase Synchronization for the Recognition of Mental Tasks in a Brain-Computer Interface," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 12, no. 4, pp. 406–415, 2004.

[23]   A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau and G. Ruffini, "Unobtrusive Biometric System Based on Electroencephalogram Analysis," EURASIP J. Adv. Signal Process., vol. 2008, no. 1, pp. 1–8, 2008.

[24]   L. Zadeh, "Fuzzy Sets," Information and Control Journal, vol. 8, no. 3, pp. 338–353, 1965.

[25]   P. Zdzislaw, "Rough Sets," International Journal of International Conference on Computer and Applications (IJCIS), vol. 11, no. 5, pp. 341–356, 1982.

[26]   R. Jensen and Q. Shen, "New Approaches to Fuzzy-rough Feature Selection," IEEE Trans. Fuzzy Syst., vol. 17, no. 4, pp. 824–838, 2009.

[27]   R. Sadeghi and J. Hamidzadeh, "Automatic Support Vector Data Description," Soft Computing Jour., pp. 147–158, 2016.

[28]   M. Zabihimayvan, R. Sadeghi, H. N. Rude and D. Doran, "A Soft Computing Approach for Benign and Malicious Web Robot Detection," Expert Systems with Applications, vol. 87, pp. 129–140, 2017.

[29]   J. Hamidzadeh, M. Zabihimayvan and R. Sadeghi, "Detection of Website Visitors Based on Fuzzy Rough Sets," Soft Computing, vol. 22, no. 7, pp. 2175–2188, 2018.

[30]   M. Zabihimayvan and D. Doran, "Fuzzy Rough Set Feature Selection to Enhance Phishing Attack Detection," Proc. of the IEEE International Conference on Fuzzy Systems, 2019.

[31]   R. Jensen and C. Cornelis, "Fuzzy-rough Nearest Neighbour Classification and Prediction," Theor. Comput. Sci., vol. 412, no. 42, pp. 5871–5884, Sep. 2011.

[32]   J. M. Keller and M. R. Gray, "A Fuzzy K-nearest Neighbor Algorithm," IEEE Trans. Syst. Man. Cybern., vol. 15, no. 4, pp. 580–585, 1985.

[33]   N. Mac Parthaláin and R. Jensen, "Fuzzy-rough Approaches for Mammographic Risk Analysis," Intell. Data Anal., vol. 14, no. 2, pp. 225–244, 2010.

[34] S. Raschka, "Model Evaluation, Model Selection and Algorithm Selection in Machine Learning," University of Wisconsin-Madison, arXiv.org > cs > arXiv:1811.12808, 2018.

[35] Z. N. Voulgaris, Discernibility Concept in Classification Problems, PhD Thesis, Uni. of London, 2009.

[36] N. Ajmal and K. Thomas, "Fuzzy Lattices," Inf. Sci. (NY), vol. 79, no. 3–4, pp. 271–291, 1994.

[37] V. Petridis and V. G. Kaburlasos, "Fuzzy Lattice Neural Network (FLNN): A Hybrid Model for Learning," IEEE Trans. Neural Networks, vol. 9, no. 5, pp. 877–890, 1998.

[38] Y. Jamshidi Khezeli and H. Nezamabadi-Pour, "Fuzzy Lattice Reasoning for Pattern Classification Using a New Positive Valuation Function," Adv. Fuzzy Syst., vol. 2012, pp. 1–8, 2012.

[39] V. G. Kaburlasos, I. N. Athanasiadis and P. A. Mitkas, "Fuzzy Lattice Reasoning (FLR) Classifier and Its Application for Ambient Ozone Estimation," Int. J. Approx. Reason., vol. 45, pp. 152–188, 2007.

[40] S. E. Papadakis, V. G. Kaburlasos and G. A. Papakostas, "Fuzzy Lattice Reasoning (FLR) Classifier for Human Facial Expression Recognition," World Scientific Proceedings Series on Computer Engineering and Information Science, pp. 633–638, 2012.

[41] S. E. Papadakis, V. G. Kaburlasos and G. A. Papakostas, "Two Fuzzy Lattice Reasoning (FLR) Classifiers and Their Application for Human Facial Expression Recognition," Jou. Mult. Log. Soft Comput., vol. 22, no. 4–6, pp. 561–579, 2014.

[42] H. Begleiter, "EEG Database," UCI Machine Learning Repository, [Online], Available at: https://archive.ics.uci.edu/ml/datasets/eeg+database.

[43] J. G. Snodgrass and M. Vanderwart, "A Standardized Set of 260 Pictures: Norms for Name Agreement, Image Agreement, Familiarity and Visual Complexity," Jou. Exp. Psychol. Hum. Learn. Mem., vol. 6, no. 2, pp. 174–215, Mar. 1980.

[44] J. V. Odom et al., "ISCEV Standard for Clinical Visual Evoked Potentials (2009 Update)," Doc. Ophthalmol., vol. 120, no. 1, pp. 111–119, Feb. 2010.

[45] V. Marsalek, P. Matousek, M. Mautner, Merta and R. Moucek, "Coherence of EEG Signals and Biometric Signals of Handwriting under Influence of Nicotine, Alcohol and Light Drugs," Neural Netw. World, vol. 16, pp. 41–60, 2006.

[46] C. X. Ling, J. Huang and H. Zhang, "AUC: A Better Measure than Accuracy in Comparing Learning Algorithms," Springer Verlag, Berlin, Heidelberg, pp. 329–341, 2003.

[47] T. W. Anderson, "Anderson-Darling Tests of Goodness-of-Fit," International Encyclopedia of Statistical Science, no. 2. Springer Verlag, Berlin, Heidelberg, pp. 52–54, 2011.

[48] A. Field, Discovering Statistics Using SPSS, 3rd Edition, Sage Publications, 2009.

[49] C. A. Bellera, M. Julien and J. A. Hanley, "Normal Approximations to the Distributions of the Wilcoxon Statistics: Accurate to What N? Graphical Insights," Jou. Stat. Educ., vol. 18, no. 2, pp. 1–17, 2010.

ملخص البحث:

تُستخدم إشارات التخطيط الكهربائي للدّماغ كوسيلة قياس بيولوجية، لأنها تتمتع بالفرادة والعمومية وقابلية الجمع. يهدف هذا العمل الى تقييم أداء تقنيات قائمة على المنطق المشوَّش من أجل نمذجة إثبات أصالة الصور باستخدام بصمات أدمغة الأشخاص. استخدمت مجموعة بيانات (EEG UCI) الأصلية للمقارنة بين ثلاثٍ من تقنيات التصنيف القائمة على المنطق المشوَّش، تم اختيار تسع قنوات تحتوي على إشارات التخطيط الكهربائي للدّماغ وتتموضع عند خط الوسط والمناطق الجانبية لاستخدامها في إجراء التجارب. ويقترح الباحثون أن يركز البحث المستقبلي على اختيار قناة التخطيط الكهربائي للدّماغ والسِّمات المتعلقة بها من أجل الحصول على تمثيلٍ أفضل للبيانات الخاصة ببصمات أدمغة الأشخاص بهدف التوصل الى قدرٍ أكبر من الدقة في إثبات أصالة صور وجوه الأشخاص للتغلب على مشكلة عدم التوازن.