IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 1, 2019, accepted June 19, 2019, date of publication July 16, 2019, date of current version August 12, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2929266

Wavelet-Based EEG Processing for Epilepsy Detection Using Fuzzy Entropy and Associative Petri Net

HSIU-SEN CHIANG, MU-YEN CHEN, AND YU-JHIH HUANG
Department of Information Management, National Taichung University of Science and Technology, Taichung 40401, Taiwan
Corresponding author: Mu-Yen Chen (mychen.academy@gmail.com)

This work was supported by the Ministry of Science and Technology of Taiwan under Grant MOST 107-2410-H-025-010-MY2.

ABSTRACT Epilepsy is a common neurological disease that can cause seizures and loss of consciousness and can have a severe negative impact on long-term cognitive function. Reducing the severity of impact requires early diagnosis and treatment. Epilepsy is traditionally diagnosed using electroencephalography (EEG) performed by trained physicians or technicians but this process is time-consuming and prone to interference, which can negatively impact accuracy. This paper develops a model for epilepsy diagnosis using discrete wavelet transform to analyze sub-bands within the EEG parameter and select EEG characteristics for epilepsy detection. The minimize entropy principle approach is used to build fuzzy membership functions of the characteristics of each brain wave and are then used as the basis for the construction of an associative Petri net model. Using our APN model, the associative Petri net approach provides diagnosis accuracy rates of 93.8%, outperforming similar approaches using decision tree, support vector machine, neural network, Bayes net, naïve Bayes, and tree augmented naïve Bayes. Thus, the proposed approach shows promise for fast, accurate, and objective diagnosis of epilepsy in clinical settings.

INDEX TERMS Epilepsy, electroencephalogram, wavelet transform, associative petri net.

I. INTRODUCTION

Epilepsy is one of the most common neurological diseases. It is a cranial nerve disease which is a paroxysmal disorder of brain functions and, among serious neurological disorders, the incidence of epilepsy is second only to cerebrovascular stroke. Epilepsy can cause recurrent, frequent abnormal brain discharge which can damage brain cells, impact cognitive function, and even cause shock. Treatment efficacy decreases as the disease progresses.

Epilepsy detection and diagnosis relies heavily on electroencephalography (EEG). It is estimated that most epilepsy patients show abnormal EEG patterns, while a small of patients still show normal EEG activity [1]. Epilepsy diagnosis using EEG requires close observation for extended periods to detect seizures in real time, but this detection is dependent on the experience and subjective interpretation of the administering technician, making it time-consuming and less than fully objective. EEG signals are also prone to external interference which can significantly reduce diagnosis efficacy. Therefore, a faster and more objective approach for the clinical diagnosis of epilepsy is needed.

This study analyzes the physiological parameters of EEG through discrete wavelet transform to generate multiple sub-bands and extract their features. Using the minimize entropy principle approach (MEPA), in conjunction with associative Petri net (APN) classification, we compare the accuracy of these features in classifying epilepsy brainwave with other common methods including decision tree, support vector machine, neural network, Bayes net, Naïve Bayes, and Tree Augmented Naïve Bayes, and develop a diagnosis model for epilepsy to reduce the time and cost of diagnosis.

The associate editor coordinating the review of this manuscript and approving it for publication was Tie Qiu.

II. LITERATURE REVIEW

A. ELECTROENCEPHALOGRAM (EEG)

EEG can effectively show abnormal discharge patterns in the brain [2], allowing it to reflect patterns potentially associated with epileptic seizure activity and other nonspecific abnormities [3] through spikes, sharp waves, and paroxysmal rapid activities which are often accompanied by slow waves [4]. Spikes rise and fall vertically with a higher

amplitude of 100∼200μV. Sharp waves are triangular waves which rise quickly and fall more slowly, with an amplitude of more than 200μV. These represent transient electrical discharge originating from synchronized neurons and are associated with cerebral blood-flow and cerebral metabolism that may affect cognition [5]. Seizures typically originate in the brain's temporal lobe, and spike amplitude during measurement will be limited by the anterior temporal electrode [6]. During seizures, EEG electrodes attached to the scalp provide reliable brain information to understand seizures [7], which can also be used as a basis for early detection of epilepsy [8].

B. RELATED RESEARCH ON EPILEPSY DETECTION

In the feature extraction of epilepsy frequency domain, there are literatures proposed that decomposing brainwaves into several sub-bands provides more information than using the original brainwave signal [9]. Brainwave signal is usually decomposed into four sub-bands: δ band, θ band, α band, and β band.

Most features extracted from frequency domain are invalid and the accuracy is generally not high. Nevertheless, there is exception; Übeyli and Güler (2007) used power spectral density (PSD) as a feature and classified by mixture of experts (ME) and modified mixture of experts (MME) to obtain an accuracy rate of 95.53% and 98.6% [10].

The use of discrete wavelet transform (DWT) in the time domain is the most practical method for EEG signal classification [10]–[16]. In order to reduce the dimension of the eigenvector and the computational complexity, statistical features such as maximum (MAX), minimum (Min), mean and standard deviation (SD) can be used to represent the signals [10], [11], [17]. Other studies [18] used different features to cut continuous energy signals into window functions. Guo et al. (2011) used the mean, SD, and curve length of asymmetric frequency bands as features [12]. The study of scholars Acharya et al. (2011) used the principal component of wavelet packet decomposition (WPD) as features [19]. The study of Chen (2014) used the coefficients of dual tree complex wavelet transform as features [15].

For the classification methods, the more common and accurate methods include the neural network theory and the probabilistic neural network(PNN), with an accuracy rate over 93% [20]–[23]; the k-nearest neighbors algorithm (KNN), with an accuracy rate of 99.3% [12]; multilayer perceptron network (MLPNN), with an accuracy rate over 89% [17], [18], [20]–[24]; artificial neural network referred to as the fuzzy inference system (ANFIS), with an accuracy rate of 98.68% [11]; support vector machine (SVM), with an accuracy rate about 99.2% [21], [24] and optimum path forest (OPF) with an accuracy rate of 89.2% [16].

III. RESEARCH METHODOLOGY

This study uses discrete wavelet transform to extract features via high-frequency and low-frequency filters, and uses a variety of classification methods to compare its accuracy rate for epilepsy detection including discrete wavelet transform, minimize entropy principle approach, and associative Petri net.

A. DISCRETE WAVELET TRANSFORM

Wavelet transform (WT) is mainly used for signal preprocessing, noise reduction, and feature extraction. WT can analyze EEG signals of different scales and capture more details than short-time Fourier transfer (STFT). Discrete wavelet transform (DWT) is a wavelet transform that can be used in the frequency and time domains to divide original data into consistent data and highly variable data to respectively run high-frequency and low-frequency filtering on the original series. The series generated by the low-frequency filter retains the consistent data of the original series and the series generated by the high-frequency filter retains the highly variable data of the original series [25].

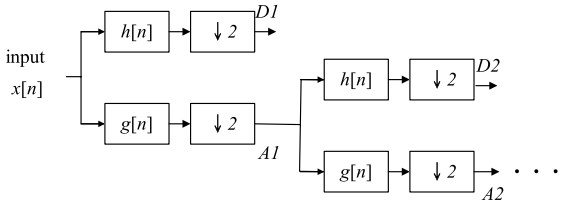


FIGURE 1. Process of discrete wavelet transform.

The architecture of DWT is shown in Fig. 1: $x[n]$ is the discrete input signal with length N; $g[n]$ is the low pass filter that filters out the high frequency of the input signal and outputs the low frequency; h[n] is the high pass filter that filters out low frequency and outputs high frequency; $\downarrow Q$ is the downsampling filter; and a indicates the a$^{\text{th}}$-layer in the architecture.

B. MINIMIZE ENTROPY PRINCIPLE APPROACH

MEPA uses the concept of entropy to minimize the level of confusion in the data and then establish a fuzzy membership function, membership degree and linguistic value. Entropy indicates the distribution uniformity of any kind of energy in space, with a higher distribution uniformity indicating greater entropy. The probability distribution of entropy is used to measure the distribution of uncertainty. MEPA divides data into different segments via interval segmentation and then evaluates the degree of clutter of messages in every specific data segment based on one key objective: to minimize entropy or data randomness. This assessment can identify the interval segment that produces the minimum degree of data randomness.

Suppose we want to find the threshold segmentation line x_i to divide the data set into two regions $[a, a + x]$ and $[a + x, b]$ in the data collection interval $[a, b]$. To find the most appropriate threshold segmentation line, MEPA will establish a threshold segmentation line for every different data type in the data collection interval $[a, b]$; therefore, a set of threshold segmentation lines X, x_i will be obtained. Next, use Eqs. (1), (2), (3), and (4) to calculate the entropy

value at the established threshold x_i for each different data type [26]. The threshold with the minimum entropy value will be the determined threshold used to divide the entire data set into two regions p and q. The induction is performed using the entropy minimization principle; we can find intervals in which the distribution of samples of any class is as relatively uniform as possible and obtain the optimal clusters within the interval [26].

$$S(x) = p(x)S_p(x) + q(x)S_q(x) \quad (1)$$

$$S_p(x) = -[p_1(x) + lnp_1(x) + p_2(x) + lnp_2(x)] \quad (2)$$

$$S_q(x) = -[q_1(x) + lnq_1(x) + q_2(x) + lnq_2(x)] \quad (3)$$

$$p(x) + q(x) = 1 \quad (4)$$

$p_k(x)$ and $q_k(x)$ respectively denote the conditional probability of k-type samples in $[a, a+x]$ and $[a+x, b]$. In the set of threshold segmentation lines, the segmentation line with the minimum entropy value is the optimum threshold. The value estimates of $p_k(x)$, $q_k(x)$, $p(x)$ and $q(x)$ are descripted in the literatures [26].

C. ASSOCIATIVE PETRI NET

Associative Petri Net (APN) integrates the Apriori algorithm to Petri net that becoming a new type of Petri network structure. It can be constructed in the original knowledge-based system to generate new knowledge rules to form a network structure that can describe the reasoning process. Then the causal relationship between the input/output places is deduced using the unique transition reasoning mechanism of the associative production rule (APR) generated by APN [27].

APN contains three types of nodes: (1) places, which are circles and describe a certain state; (2) squares, which represent the threshold of association between the antecedent and consequence, and are used to evaluate whether there is an association between two place nodes; (3) transitions, bars, which contain an associative function G, as referred to as certainty function (CF), that transform the association through which different antecedents impact consequences into appropriate representations. When the conditions are satisfied, the antecedent is transformed into consequence via a transition mechanism. The main concept of APN lies in the discussion of dynamic processes. Therefore, its implementation rule emphasizes its ability to enable and fire the transition; that is, firing transitions that are enabled. During the transition state, inference ability is triggered by different conditions to further explore the influence and correlation of antecedent and consequence. The APN can be defined as a 13-tuple, APN $= (P, T, S, C, D, \Lambda, \Gamma, I, O, \alpha, \beta, G, W)$ and a detailed description is in the literatures [27] and [28].

1) ASSOCIATIVE PRODUCTION RULE (APR)

The reasoning mechanism of APN uses different types of APR and, based on different input antecedents, puts the association function between places and transitions into consideration to reason out the possible value of consequences via the unique reasoning algorithm. Given the antecedents or consequences and the uncertainty factor of reasoning process, in an APN a logical approach for processing uncertainty is incorporated to the reasoning mechanism to produce five different types of associative production rules (APR). According to different conditions, APR comprises "and" or "or" computations and is also called a complex APR. This study mainly uses three types of APRs and their calculation are described as follows [27]:

Type 1: IF d_j THEN d_k (CF $= c_{jk}$). This form of APR is calculated as formula (5):

$$\alpha(p_k) = \alpha\left(p_j\right)^* c_{jk} \quad when\ s_{jk} \geq \tau_{jk},\ c_{jk} \geq \gamma_{jk} \quad (5)$$

Type 2: IF d_{j1} or d_{j2} or . . ..or d_{jn} THEN d_k (CF $= c_{ji}$). This form of APR is calculated as formula (6):

$$\alpha(p_k) = Max\left\{\alpha\left(p_{j1}\right)^* c_{j1},\quad \alpha\left(p_{j2}\right)^* c_{j2}, \ldots \alpha\left(p_{jn}\right)^* c_{jn}\right\} \quad when\ s_{ji} \geq \tau_{ji},\quad c_{ji} \geq \gamma_{ji}, i = 1, 2, \ldots n \quad (6)$$

Type 3: IF d_j THEN d_{k1} or d_{k2} or. . ..or d_{kn} (CF $= c_{ji}$). This form of APR is calculated as formula (7):

$$\alpha(p_{k1}) = \alpha\left(p_{j1}\right)^* c_{j1},\quad \alpha(p_{k2}) = \alpha\left(p_{j2}\right)^* c_{j2}, \ldots \alpha(p_{kn}) = \alpha\left(p_{jn}\right)^* c_{jn} \quad when\ s_{ji} \geq \tau_{ji},\ c_{ji} \geq \gamma_{ji},\quad i = 1, 2, \ldots n \quad (7)$$

All types of APRs can be presented mathematically and graphically. By connecting the associated places and giving transitions an appropriate CF, an APN reasoning model that expresses the specialized domain knowledge can be derived.

2) REASONING ALGORITHM

Each place in an APN model is represented by a triple ($p_{\backslash x}$, $\alpha(p_x)$, IRS(p_x)), where $p_x \in P$, $P = \{p_1, p_2, \ldots\ldots .p_n\}$ is a finite set of places. The p_x represents a place (node) in an APN and d_x denote the proposition of p_x. The degree of truth of proposition d_x is defined as $\alpha(p_x)$; the threshold of degree of truth of each proposition is given as λ. If $\alpha(p_x) \geq \lambda_x$, then the proposition d_x exists. IRS(p_x) and RS(p_x) denote the immediate reachability set and reachability set of p_x, respectively. Taking into account the degree of association of every antecedent proposition and consequence proposition, this study adopted the Apriori algorithm in association rule [29] to find the associative degree between places p_x and p_y. Let s_{xy} is the support degree and c_{xy} denote conference degree between places p_x and p_y. The thresholds of support and confidence degree are given with τ_{xy} and γ_{xy}. When the associative degree satisfies their minimum threshold τ_{xy} and γ_{xy}, they are considered interesting; these threshold values can be defined by user-experience or experts in this domain. If the support and confidence of values are higher than threshold values ($s_{xy} \geq \tau_{xy}$ and $c_{xy} \geq \gamma_{xy}$), the transition t_{xy} is enabled to fire. Moreover, an appropriate certainty function and corresponding confidence value (CF $= c_{xy}$) is given, otherwise the association does not exist and the corresponding confidence value is zero (CF $= 0$). When a transition t_i fires, the tokens in the input places pass all squares and

move to output places. The degree of truth of proposition d_y is calculated by $\alpha(p_x)^* c_{xy}$.

If there is an APN network structure, a reasoning algorithm is adopted to generate all reasoning paths from starting place p_s to goal place p_g. Through our proposed a reasoning algorithm [27], the degree of truth of the goal proposition can be predicted under different antecedents. An example of epilepsy diagnosis will be used to illustrate the reasoning algorithm in this study.

IV. EXPERIMENT DESIGN AND PERFORMANCE EVALUATION

A. DATA SET DESCRIPTION

The dataset used in this study is a public dataset published by the Epilepsy Center of Department of Epileptology, University of Bonn (http://epileptologie-bonn.de/cms/front_content.php?idcat = 193) [30], and has been widely used in studies related to EEG signals [10], [11], [18], [24]. Class A is the brainwave measured from normal subjects awake with eyes open. Class B is the brainwave measured from normal subjects awake with eyes closed. Class C is the brainwave measured in the regions around the hippocampal formation of epileptic patients during a partial seizure. Class D is the brainwave measured in the hippocampal formation of epileptic patients during a partial seizure. Class E is the collection of data measured centralizing in the hippocampal formation during a clinician-confirmed generalized seizure. Each class contains 100 samples, for a total of 500 samples each lasting 23.6 seconds with a sampling frequency range of 0.53∼40Hz. The sampling rate per second is 173.61Hz, and the resolution is 12 bits. This study classified Class A and Class B as normal, with a total of 200 samples; and Class C, Class D, and Class E were classified as epilepsy, with a total of 300 samples.

B. CONSTRUCTION OF EPILEPSY DIAGNOSIS MODEL

In this study, the epilepsy diagnosis model was constructed in two stages. First, the dataset was discretized by discrete wavelet transform (DWT) and important features were extracted. The second stage was the diagnosis of epilepsy, using the minimize entropy principle approach (MEPA) to produce fuzzy membership functions and rules, and integrate the associative Petri net (APN) for classification.

1) DATA PROCESSING

This study ran discrete wavelet transformation on the 500 samples using MATLAB software. The waveform used was the db2, which was relatively smooth and the downsampling parameter was set to 2 [11]. Discrete wavelet transformation discretized the data. The process is shown in Fig. 1. In the first discrete wavelet transform, the D1 sub-band with higher frequency and A1 the sub-band with lower frequency were divided via high-pass and low-pass filter. Next, the A1 sub-band was used for the second discrete wavelet transform. The D2 sub-band with higher frequency and the A2 sub-band with lower frequency were divided via high-pass and low-pass filter. The A2 sub-band was then used for the third discrete wavelet transformation. Similarly, the A3 sub-band was used for the fourth discrete wavelet transformation to divide the high-frequency D4 sub-band and low-frequency A4 sub-band. The dataset was divided into D1, D2, D3, D4, and A4 sub-bands.

2) FEATURE EXTRACTION

According to Übeyli et al., to reduce the dimensionality of the feature eigenvector and the computational complexity, statistical features such as maximum (Max), minimum (Min), mean, and standard deviation (SD) can be used to express the signals [20]–[24]. In this study, the maximum, minimum, mean, and standard deviation of each sub-band generated after discrete wavelet transformations were used as the features to classify epileptic states. A total of 20 features are summarized in Table 1.

TABLE 1. Features of sub-bands for epilepsy detection.

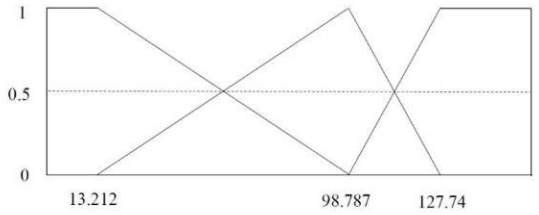
Class	Features	(Sub-bands)				
		D1	D2	D3	D4	A4
Normal	Max	125.54	233.66	652.19	903.58	787.93
	Min	-125.85	-289.51	-550.85	-852.35	-786.9
	Mean	-0.003	-0.004	0.031	-0.808	-37.572
	SD	2.994	10.379	36.96	48.755	29.684
Epilepsy	Max	779.05	1548.7	3339.7	3789.2	6094.8
	Min	-863.56	-1847.7	-3710	-4376.5	-5339.8
	Mean	-0.005	-0.02	-0.06	-0.0362	-26.470
	SD	21.54	89.05	229.356	342.812	350.364

3) FUZZY MEMBERSHIP FUNCTION

In this study, the MEPA is used to establish the fuzzy membership function of every feature in each sub-band. First, the entropy of dataset was calculated; the minimum was the optimal segmentation position and used to divide the dataset into two datasets. The entropy values of these two datasets were respectively calculated, and then the minimum entropy was used as the segmentation position. The entropy value at the segmentation position was used to establish the fuzzy membership function of the feature. Taking the Max of the D1 sub-band as an example, the fuzzy membership function of the feature from low to high was divided into $D1\_Max_L$, $D1\_Max_M$ and $D1\_Max_H$. Figure 2 shows the fuzzy membership function of Max in the D1 sub-band. The equations for the low, middle, and high semantic membership functions are shown below.

4) ASSOCIATIVE PETRI NET MODEL

Using the fuzzy membership functions constructed in section 4.2.3, we incorporate the linguistic value and membership degree into the propositions of input place and degree of truth in the associative Petri net. We establish the associating Petri



$$\text{D1\_Max} = \begin{cases} D1\_Max_L = \begin{cases} 1 & x < 13.212 \\ \dfrac{98.787 - x}{85.575} & 13.212 < x \le 55.9995 \end{cases} \\ D1\_Max_M = \begin{cases} \dfrac{x - 13.212}{85.575} & 55.9995 < x \le 98.7870 \\ \dfrac{127.74 - x}{28.953} & 98.7870 < x \le 113.2635 \end{cases} \\ D1\_Max_H = \begin{cases} \dfrac{x - 98.787}{28.953} & 113.2635 < x \le 127.74 \\ 1 & x > 127.74 \end{cases} \end{cases}$$

FIGURE 2. Fuzzy membership function of max_D1.

net model by calculating the association between each feature and epilepsy and setting the threshold value of support to 0.05 and threshold value of confidence to 0.05 by trial and error.

To improve the accuracy of epilepsy diagnosis, this study obtained the ranking of information gain of each feature to sort the feature with priority. Excluding the lowest ranked D2_Mean and D4_Mean, the accuracy of the associative Petri net model classification of epilepsy was tested by trial and error to finally establish the associative Petri net model as shown in Fig. 3.

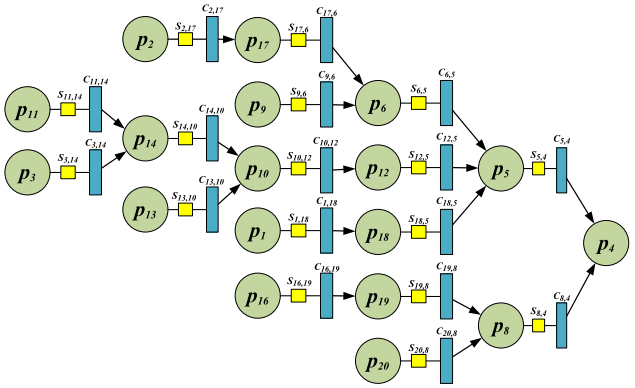


FIGURE 3. APN model for epilepsy detection.

5) REASONING OF EPILEPSY DIAGNOSIS

An example of epilepsy diagnosis reasoning is described with our proposed APN algorithm. As illustrated in Fig. 3, our proposed APN model has eight input places and nine middle places, and one final place. The support and confidence of each APR were calculated using equations shown

TABLE 2. Support and confidence of APRs.

APRs	Support (s)	Confidence(c)
$p_{11} \rightarrow p_{14}$	0.0960	0.6274
$p_3 \rightarrow p_{14}$	0.0660	0.6291
$p_2 \rightarrow p_{17}$	0.3460	0.9605
$p_{14} \rightarrow p_{10}$	0.5400	0.9783
$p_{13} \rightarrow p_{10}$	0.6600	0.9483
$p_{17} \rightarrow p_6$	0.3140	0.8345
$p_9 \rightarrow p_6$	0.5920	0.8916
$p_{10} \rightarrow p_{12}$	0.5200	0.9780
$p_1 \rightarrow p_{18}$	0.5900	0.9855
$p_{16} \rightarrow p_{19}$	0.3760	0.7801
$p_6 \rightarrow p_5$	0.5720	0.9556
$p_{12} \rightarrow p_5$	0.4480	1
$p_{18} \rightarrow p_5$	0.5160	0.8063
$p_{19} \rightarrow p_8$	0.5120	0.7065
$p_{20} \rightarrow p_8$	0.4820	0.8462
$p_5 \rightarrow p_4$	0.3220	0.9417
$p_8 \rightarrow p_4$	0.3800	0.9742

TABLE 3. IRS and RS of each places.

Place p_i	IRS(p_i)	RS(p_i)
p_1	$\{p_{18}\}$	$\{p_{18}, p_5, p_4\}$
p_2	$\{p_{17}\}$	$\{p_{17}, p_6, p_5, p_4\}$
p_3	$\{p_{14}\}$	$\{p_{14}, p_{10}, p_{12}, p_5, p_4\}$
p_4	ϕ	ϕ
p_5	$\{p_4\}$	$\{p_4\}$
p_6	$\{p_5\}$	$\{p_5, p_4\}$
p_8	$\{p_4\}$	$\{p_4\}$
p_9	$\{p_6\}$	$\{p_6, p_5, p_4\}$
p_{10}	$\{p_{12}\}$	$\{p_{12}, p_5, p_4\}$
p_{11}	$\{p_{14}\}$	$\{p_{14}, p_{10}, p_{12}, p_5, p_4\}$
p_{12}	$\{p_5\}$	$\{p_5, p_4\}$
p_{13}	$\{p_{10}\}$	$\{p_{10}, p_{12}, p_5, p_4\}$
p_{14}	$\{p_{10}\}$	$\{p_{10}, p_{12}, p_5, p_4\}$
p_{16}	$\{p_{19}\}$	$\{p_{19}, p_8, p_4\}$
p_{17}	$\{p_6\}$	$\{p_6, p_5, p_4\}$
p_{18}	$\{p_5\}$	$\{p_5, p_4\}$
p_{19}	$\{p_8\}$	$\{p_8, p_4\}$
p_{20}	$\{p_8\}$	$\{p_8, p_4\}$

in Table 2 below. The immediate reachability set (IRS) and reachability set (RS) of each place are summarized in Table 3.

We assume an EEG data record has twenty features after the wavelet transform and feature extraction. These feature values are represented by p_1, p_2, $p_3 \ldots$ and p_{20} in our proposed APN model, as shown in Fig. 3. Let d_1, d_2, $d_3 \ldots$ and d_{20} be the propositions of p_1, p_2, $p_3 \ldots$ and p_{20}. The places p_1, p_2, p_3, p_9, p_{11}, p_{13}, p_{16}, and p_{20} are called the starting place and the goal place is p_{21}. The truth degree of the propositions of the eight starting places $\{p_1, p_2, p_3, p_9, p_{11}, p_{13}, p_{16}, p_{20}\}$ based on the fuzzy membership function is {0.8406, 0.9217, 0.5926, 0.8599, 0.5167, 0.9246, 0.5329, 1} for the input of APN. Assume that all the threshold values τ_{ij} and γ_{ij} of support and confidence are respectively set to 0.1 and 0.35. Our reasoning process uses these changes to

the EEG characteristics to assess the likelihood of epilepsy. The APN reasoning process is as follows:

Set p_3 and p_{11} as the starting points and make the inference:

Step 1: moving from p_3 and p_{11} to p_{14}
$\alpha(p_{14}) = Max\{\alpha(p_3)*c_{3,14}, \alpha(p_{11})*c_{11,14}\} =$
$Max\{0.5926*0.6291, 0.5167*0.6274\} =$
$Max\{0.3728,0.3242\} = 0.3728$

Step 2: moving from p_2 to p_{17}
$\alpha(p_{17}) = \alpha(p_2)*c_{2,17} = \{0.5926*0.9605\}$
$= 0.5692$

Step 3: moving from p_{14} and p_{13} to p_{10}
$\alpha(p_{10}) = Max\{\alpha(p_{14})*c_{14,10}, \alpha(p_{13})*c_{13,10}\} =$
$Max\{0.3728*0.9783, 0.9246*0.9483\} =$
$Max\{0.3647,0.8768\} = 0.8768$

Step 4: moving from p_{17} and p_9 to p_6
$\alpha(p_6) = Max\{\alpha(p_{17})*c_{17,6}, \alpha(p_9)*c_{9,6}\} =$
$Max\{0.5692*0.8345, 0.8599*0.8916\} =$
$Max\{0.4750,0.7667\} = 0.7667$

Step 5: moving from p_{10} to p_{12}
$\alpha(p_{12}) = \alpha(p_{10})*c_{10,12} = \{0.8768*0.9780\}$
$= 0.8575$

Step 6: moving from p_1 to p_{18}
$\alpha(p_{18}) = \alpha(p_1)*c_{1,18}$
$= \{0.8406*0.9855\}$
$= 0.8284$

Step 7: moving from p_{16} to p_{19}
$\alpha(p_{19}) = \alpha(p_{16})*c_{16,19}$
$= \{0.5329*0.7801\} = 0.4157$

Step 8: moving from p_6, p_{12}, and p_{18} to p_5
$\alpha(p_5) = Max\{\alpha(p_6)*c_{6,5}, \alpha(p_{12})*c_{12,5},$
$\alpha(p_{18})*c_{18,5}\} = Max\{0.7667*0.9556,0.8575$
$*1,0.8284*0.8063\} =$
$Max\{0.7327,0.8575,0.6679\} = 0.8575$

Step 9: moving from p_{19} and p_{20} to p_8
$\alpha(p_8) = Max\{\alpha(p_{19})*c_{19,8}, \alpha(p_{20})*c_{20,8}\} =$
$Max\{0.4157*0.7065, 1*0.8462\} =$
$Max\{0.2937,0.8462\} = 0.8462$

Step 10: moving from p_5 and p_8 to p_4
$\alpha(p_4) = Max\{\alpha(p_5)*c_{5,4}, \alpha(p_8)*c_{8,4}\} =$
$Max\{0.8575*0.9417, 0.8462*0.9742\} =$
$Max\{0.8075,0.8244\} = 0.8244$

Therefore, after reasoning we can obtain the propositions of p_4 and the degree of truth.

C. EVALUATION OF EPILEPSY DIAGNOSIS MODEL

The associative Petri net (APN) used in this study was compared with other common classification methods, including decision tree (DT), neural network (NN), support vector machine (SVM), Bayes net (BN), tree augmented Naïve Bayes (TAN), and Naïve Bayes (NB). The decision tree adopted the C4.5 algorithm and set the minimum number of leaf nodes of each branch at 2; the pruning threshold value was 0.25 and seed was 1; the SVM kernel used polynomial functions. Each method ran cross-validation 10 times. The data were divided into 10 subsamples, 9 for training and 1 for testing. The cross-validation was repeated for 10 times and each subsample was validated. The results are shown in Table 4. Precise rate, F-measure, G-mean and accuracy of the APN diagnosis model used in this study outperformed other machine learning methods.

TABLE 4. Evaluation of various classification methods.

Methods	Precise	Recall	AUC	Fmeasure	Gmean	Accuracy
DT	93%	90.7%	0.947	0.919	0.930	93.4%
NN	89.5%	93.2%	0.974	0.913	0.932	93.2%
SVM	80.5%	67.4%	0.773	0.733	0.757	76.6%
BN	95.5%	79.6%	0.966	0.868	0.877	88.4%
NB	94.5%	51.1%	0.862	0.663	0.684	61.6%
TAN	94%	89.5%	0.982	0.917	0.926	93.2%
APN	99%	87.2%	0.98	0.927	0.931	93.8%

This section compares the common classification methods used for epilepsy diagnosis. During the comparison process, True Positives (TP: the number of normal samples classified as normal), True Negatives (TN: number of epilepsy samples classified as epilepsy), False Negatives (FN: number of epilepsy samples classified as normal), and False Positives (FP: number of normal sample classified as epilepsy) are calculated and evaluated by the following indicators:

(1) Precision Rate: TP/(TP + FP), the correct proportion of normal sample classification.

(2) Recall Rate: TP/(TP + FN), the proportion of normal samples in all samples classified as normal.

(3) F-measure: 2* TP/(2* TP + FP + FN), a measure of the quality for classification system.

(4) G-mean: $\sqrt{(TP/(TP + FN) \times TN/(TN + FP))}$, assessment for the average classification accuracy of all classes.

(5) AUC: the area under the ROC curve, where a larger value indicates that at least one class is correctly classified.

(6) Accuracy: the correct number of epilepsy and normal classification/total number of samples.

This study combined MEPA and APN to perform epilepsy diagnosis, and can effectively and accurately determine whether the patients suffer from epilepsy. The experimental results of the APN model proposed in this study showed a precision rate of 99% and negative predictive value of 90.33% where the negative predictive value was calculated by TN/(TN + FN). Compared with other methods, the precision rate of the APN model was the best, indicating that the probability of misclassification of the APN model in classifying normal participants was lower than in other methods. This improved accuracy will help reduce medical costs due to diagnostic errors, such as follow-up testing. The negative predictive value of the APN model was lower than DT, NN, and TAN, indicating a lower probability of misclassification of epilepsy.

1) COMPARISON WITH PAST WORK

This study used APN to classify Class A and Class B as normal participants and Class C, Class D, and Class E as epilepsy patients (two classes) with an accuracy rate of 93.8%. Nunes *et al.* (2014) [16] classified Class A and Class B as normal

subjects, Class C and Class D as patients with partial seizure, and Class E as patients with generalized seizure, a total of three classes. The accuracy rates using OPF, Bayesian, SVM-RBF and ANN-MLP were respectively 89.2%, 87.4%, 84.4%, and 70.6% [16]. The data classification of Nunes' study was similar to this study, but the present study produced more accurate results.

To explore the accuracy of the classification of epilepsy patients with generalized seizures, Guo *et al.* (2010a, 2010b) [31], [32] classified Class A, Class B, Class C, and Class D into one class and Class E to another. Using the MLPNN and ANN for classification, they achieved respective accuracy rates of 97.77% and 98.27%. To compare against Guo et al. [31], [32], we classified Class A, Class B, Class C, and Class D into one class and Class E to another; using the APN for classification and information gain to filter variables, with an obtained accuracy of 98.6%, thus improving on the results of Guo *et al.* [31], [32]. Comparison results are summarized in Table 5.

TABLE 5. Comparison of prior studies for epilepsy detection.

Dimensionality	Class	Methods	Accuracy	Author
Line length feature	ACD-E	MLPNN	97.75%	Guo et al.(2010a) [31]
	ABCD-E	MLPNN	97.77%	
Approximate entropy	ABCD-E	ANN	98.27%	Guo et al.(2010b) [32]
Information Gain RELIEF Correlation feature selection	AB-CD-E	OPF	89.2%	Nunes et al.(2014) [16]
		Bayesian	87.4%	
		SVM-RBF	84.4%	
		ANN-MLP	70.6%	
Information Gain	ABCD-E	APN	98.6%	This study
	AB-CDE		93.8%	

V. CONCLUSION

This study constructed an epilepsy diagnosis model. First, high-pass and low-pass filters were used to generate D1∼D4 and A4 sub-bands and extracted the Max, Min, mean and SD of the features. Then, the information gain was used to filter out variables and MEPA was adopted to construct the fuzzy membership function of each feature to understand the relationship between each feature and seizures. The information gain was used to sort and rank the features. Finally, the APN was used for classification, achieving an accuracy rate of 93.8% thus outperforming other common machine learning classification methods. The diagnosis model established in this study (1) can quickly classify normal and epilepsy in the initial diagnosis and serve as an objective indicator for epilepsy diagnosis; (2) the APN model can graphically represent the classification rules, and can easily be transformed into a rule base for an expert system; (3) In addition to classifying into Class AB and Class CDE for further exploration, this study also classified into Class ABCD and Class E to compare with past studies. Using the information gain to sort and rank the features and finally using the APN for classification, we obtained an accuracy rate of 98.6%, which improved on previous results.

Since samples of patients with epilepsy were difficult to obtain, this study was limited to use of public datasets for testing. Although the accuracy rate was good, it may be biased in clinical diagnosis. Future studies should cooperate with doctors to collect patient samples of different severity to develop a diagnosis system based on this diagnosis model, combining mobile devices and a simple EEG App as the objective basis for clinical diagnosis to reduce costs and time needed to achieve an accurate diagnosis.

REFERENCES

[1] National Institute of Neurological Disorders and Stroke, "EEG monitoring," in *The Epilepsies and Seizures: Hope Through Research*. Rockville, MD, USA: Bethesda, 2004, p. 8.

[2] S. Noachtar and J. Rémi, "The role of EEG in epilepsy: A critical review," *Epilepsy Behav.*, vol. 15, pp. 22–33, May 2009.

[3] R. Kennett, "Modern electroencephalography," *J. Neurol.*, vol. 259, no. 4, pp. 783–789, Apr. 2012.

[4] M. de Curtis, J. G. R. Jefferys, and M. Avoli, "Interictal epileptiform discharges in partial epilepsy: Complex neurobiological mechanisms based on experimental and clinical evidence," in *Jasper's Basic Mechanisms of the Epilepsies*, 4th ed. Bethesda, MD, USA: NCBI, 2012.

[5] E. Rodin, T. Constantino, S. Rampp, and P. K. Wong, "Spikes and epilepsy," *Clin. EEG Neurosci.*, vol. 40, no. 4, pp. 288–299, 2009.

[6] B. A. Dworetzky and C. Reinsberger, "The role of the interictal EEG in selecting candidates for resective epilepsy surgery," *Epilepsy Behav.*, vol. 20, no. 2, pp. 167–171, Feb. 2011.

[7] M. Sammaritano, A. de Lotbiniére, F. Andermann, A. Olivier, P. Gloor, and L. F. Quesney, "False lateralization by surface EEG of seizure onset in patients with temporal lobe epilepsy and gross focal cerebral lesions," *Ann. Neurol.*, vol. 21, no. 4, pp. 361–369, Apr. 1987.

[8] M. Hildebrandt, R. Schulz, M. Hoppe, T. May, and A. Ebner, "Postoperative routine EEG correlates with long-term seizure outcome after epilepsy surgery," *Seizure*, vol. 14, no. 7, pp. 446–451, Oct. 2005.

[9] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 2, pp. 205–211, Feb. 2007.

[10] E. D. Übeyli and . Güler, "Features extracted by eigenvector methods for detecting variability of EEG signals," *Pattern Recognit. Let.*, vol. 28, no. 5, pp. 592–603, Apr. 2007.

[11] . Güler and E. D. Übeyli, "Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients," *J. Neurosci. Methods*, vol. 148, no. 2, pp. 113–121, 2005.

[12] L. Guo, D. Rivero, J. Dorado, C. R. Munteanu, and A. Pazos, "Automatic feature extraction using genetic programming: An application to epileptic EEG classification," *Expert Syst. Appl.*, vol. 38, no. 8, pp. 10425–10436, Aug. 2011.

[13] U. R. Acharya, S. V. Sree, A. P. C. Alvin, and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework," *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9072–9078, Aug. 2012.

[14] U. R. Acharya, S. V. Sree, A. P. C. Alvin, R. Yanti, and J. S. Suri, "Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals," *Int. J. Neural Syst.*, vol. 22, no. 2, Apr. 2012, Art. no. 1250002.

[15] G. Chen, "Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features," *Expert Syst. Appl.*, vol. 41, no. 5, pp. 2391–2394, 2014.

[16] T. M. Nunes, A. L. V. Coelho, C. A. M. Lima, J. P. Papa, and V. H. C. de Albuquerque, "EEG signal classification for epilepsy diagnosis via optimum path forest—A systematic assessment," *Neurocomputing*, vol. 136, pp. 103–123, Jul. 2014.

[17] N. F. Güler, E. D. Übeyli, and I. Güler, "Recurrent neural networks employing Lyapunov exponents for EEG signals classification," *Expert Syst. Appl.*, vol. 29, no. 3, pp. 506–514, Oct. 2005.

[18] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time–frequency analysis," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 5, pp. 703–710, Sep. 2009.
[19] U. R. Acharya, S. V. Sree, and J. S. Suri, "Automatic detection of epileptic EEG signals using higher order cumulant features," *Int. J. Neural Syst.*, vol. 21, no. 5, pp. 403–414, Oct. 2011.
[20] E. D. Übeyli, "Combined neural network model employing wavelet coefficients for EEG signals classification," *Digit. Signal Process.*, vol. 19, no. 2, pp. 297–308, Mar. 2009.
[21] E. D. Übeyli, "Decision support systems for time-varying biomedical signals: EEG signals classification," *Expert Syst. Appl.*, vol. 36, no. 2, Part 1, pp. 2275–2284, Mar. 2009.
[22] E. D. Übeyli, "Probabilistic neural networks combined with wavelet coefficients for analysis of electroencephalogram signals," *Expert Syst.*, vol. 26, no. 2, pp. 147–159, May 2009.
[23] E. D. Übeyli, "Lyapunov exponents/probabilistic neural networks for analysis of EEG signals," *Expert Syst. Appl.*, vol. 37, no. 2, pp. 985–992, Mar. 2010.
[24] E. D. Übeyli, "Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines," *Comput. Biol. Med.*, vol. 38, no. 1, pp. 14–22, Jan. 2008.
[25] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Trans. Inf. Theory*, vol. 36, no. 5, pp. 961–1005, Sep. 1990.
[26] T. J. Ross, "Develpoment of membership functions," in *Fuzzy Logic with Engineering Applications*, 3rd ed. Hoboken, NJ, USA: Wiley, 2010, pp. 200–205.
[27] D.-H. Shih, H.-S. Chiang, and B. Lin, "A generalized associative Petri net for reasoning," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 9, pp. 1241–1251, Sep. 2007.
[28] H.-S. Chiang and Z. W. Wu, "Online incremental learning for sleep quality assessment using associative Petri net," *Appl. Soft Comput.*, vol. 68, pp. 774–783, Jul. 2018.
[29] R. Agrawal, T. Imielinski, and A. Swami, "Database mining: A performance perspective," *IEEE Trans. Knowl. Data Eng.*, vol. 5, no. 6, pp. 914–925, Dec. 1993.
[30] R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 64, no. 6, p. 061907, 2001.
[31] L. Guo, D. Rivero, and A. Pazos, "Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks," *J. Neurosci. Methods*, vol. 193, no. 1, pp. 156–163, Oct. 2010.
[32] L. Guo, D. Rivero, J. Dorado, J. R. Rabuñal, and A. Pazos, "Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks," *J. Neurosci. Methods*, vol. 191, no. 1, pp. 101–109, 2010.

HSIU-SEN CHIANG received the Ph.D. degree. He is currently a Professor with the Department of Information Management, National Taichung University of Science and Technology, Taiwan. His current research interests include data mining, petri net, biomedical science, and internet marketing. Dr. Chiang's research is published or is forthcoming in *Applied Soft Computing*, *Information Fusion*, *Bioinformatics*, IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, *Journal of Medical Systems*, *Journal of Medical and Biological Engineering*, *Expert Systems with Applications*, and a number of national, and international conference proceedings.

MU-YEN CHEN received the Ph.D. degree. He is currently a Professor with the Department of Information Management, National Taichung University of Science and Technology, Taiwan. His current research interests include artificial intelligent, soft computing, bio-inspired computing, financial engineering, and data mining. Dr. Chen's research is published or is forthcoming in *Information Sciences*, *Applied Soft Computing*, *Neurocomputing*, *Neural Computing and Applications*, *Journal of Educational Technology and Society*, *Journal of Information Science*, *The Electronic Library*, *Computers and Mathematics with Applications*, *Quantitative Finance*, *Expert Systems with Applications*, *Soft Computing*, and a number of national, and international conference proceedings. He has served as an Editor-in-Chief and an Associate Editor for international journals (e.g. *International Journal of Big Data and Analytics in Healthcare*, IEEE ACCESS, *Journal of Information Processing Systems*, *International Journal of Social and Humanistic Computing*) while he is an editorial board member on several SCI journals.

YU-JHIH HUANG received the M.S. degree from the Department of Information Management, National Taichung University of Science and Technology, Taiwan. His current research interests include data mining, petri net, biomedical science, and internet marketing.