

Multi-Scale Entropy and Neural Networks for Detection of Depth of Anaesthesia within EEG Signals

Mohamed H. Almeer

Computer Science & Engineering Dept, College of Engineering, Qatar University, POB 2713, Qatar

Abstract: *In recent years, multi-scale sample entropy (MSSE) is rapidly gaining popularity as an interesting tool for exploring neurophysiological mechanisms. In this paper, we propose a new method based on MSSE for on-line monitoring of the depth of anaesthesia (DoA) to quantify the anaesthetic effect with real-time electroencephalography by using MSSE. Empirical mode decomposition (EMD) was used successfully to filter EEG recordings of artefacts before the imitation of the recognition phases. Artificial neural networks were used to classify three scales of the MSSE into three stages of hypnosis. Results showed 95% accuracy of the tested clinical anaesthesia EEG recordings obtained for 20 patients.*

Keywords: *DoA; Anaesthesia; EMD; ANN; Sample Entropy.*

1. INTRODUCTION

Electroencephalogram (EEG) signals are a proven method for monitoring the depth of anaesthesia (DoA). In general, many anaesthesiologists monitor the depth of anaesthesia by means of clinical parameters such as blood pressure, heart rate, muscle movement, pupil size, lacrimation, and airway pressure [1-3]. However, these physiological parameters change in response to anaesthetic agents and are also affected by other factors such as body weight and age. Therefore, these parameters alone may not provide an accurate assessment of anaesthesia. Since anaesthetic agents affect the brain cortex, monitoring brain activity using EEG signals would be a suitable method to measure the DoA. EEG signals are generally non-stationary, although they may include some stationary epochs; these signals reflect the non-stationary activity of the brain.

Many studies have been conducted on the measurement of DoA using various methods and with multiple algorithms [21, 22, 25]. Neural networks [10, 11, 14, 15, 17, 18, 20], both feed-forward and recurrent, as well as DWT wavelets [8, 11] have also been used to classify the anaesthesia levels. Empirical mode decomposition (EMD) [11] and its variants have been employed to filter the samples of artefacts and obtain a clean EEG before attempting recognition. Some studies have investigated the value of different frequency bands in EEG [16] in the measurement of the DoA; for example, the alpha-frequency band showed a strong correlation to the level of consciousness during surgery under anaesthesia. Furthermore, some studies have utilised the chaotic feature as described by the chaos theory to follow the hypnotic levels [9].

Many scientific standards have been developed recently to measure DoA, but the most preferred important monitor is the bispectral analysis index (BIS) [4, 5]. BIS, one of the high-grade spectral analysis methods, is a complex measure encompassing the domains of time and frequency as well as high-order spectral parameters. It is determined by the analysis of the correlation of the different phases of the parts of the EEG signal. Generally, BIS readings are obtained from signals of four different EEG electrodes. BIS monitors are the most common tools used to measure the DoA, except in the case of some anaesthetic drugs. However, some investigations have indicated that the patient could awaken momentarily even if the BIS reading is between 40 and 60.

BIS entropy monitors have been recently developed to monitor the DoA using EEG signals. These monitors mainly display two readings simultaneously: response entropy (RE) and state entropy (SE)

[27]. Both parameters reflect the degree of regularity (or irregularity) of the EEG signal. The RE and SE measures are based on spectral entropy which relies heavily on fast Fourier transformation (FFT) for the calculations [28]; therefore, the spectral entropy may miss some nonlinear and non-stationary information within the EEG signal.

Approximate entropy (*ApEn*) [30] and sample entropy (*SampEn*) [31] are two recently developed and widely used entropy measures that have been shown to be efficient in the detection of complexity of any time series. In particular, *SampEn* has been shown to perform even better than *ApEn* in detecting the DoA of patients in surgeries [29-31]. *SampEn* measures complexity based on only one time scale, which makes it likely to miss some features. On the other hand, multi-scale entropy (MSE) [7, 13, 23, 33, 34, 35] used multiple scales to determine the complexity and can therefore capture the dynamic changes associated with multiple time scales. This property makes MSE much more suitable than other techniques in computing the degree of complexity for the EEG time series.

Although the use of MSE improves the correlation between the hypnosis level and the entropy value, but ultimately its is a value between 0 and 3. Because of these multiple entropy values, this method does not allow for a straightforward assessment of the DoA [35]. Artificial neural network (ANN) is one of the best artificial intelligence methods that has the ability to learn from a set of presented data through matching inputs with the outputs and then generalise the learned relationship to other foreign inputs [36]. By updating unknown weights and functions until a match is reached, a good relation is constructed on the basis of the weights and functions that correlate the input to the output sets.

The purpose of this research study is to estimate the level of anaesthesia based on EEG time series using multi-scale *SampEn* with the help of an ANN model that maps the three scales of *SampEn* to the anaesthetic condition of the patient under surgery, by using only generic filtering obtained by the generic EMD algorithm.

2. MATERIAL AND METHODS

2.1. Data Source

The single-channel EEG recording tested in this study was collected from 20 patients aged between 23 and 70 years through a forehead-mounted sensor placed on patients went under surgery of ENT cases under general anaesthesia at the National Taiwan University Hospital (NTUH) of Taiwan. The EEG data recording has been acquired using Philips Intellivue MP60. The anaesthetic drug sevoflurane or sesflurane was used for tracheal intubation of all 20 patients.

2.2. Data Processing

The standard operation procedure with general anaesthesia was divided into 4 stages: pre-operation, induction, maintenance, and recovery. The collected EEG recordings were divided to 2 parts. One set of recordings was collected from 6 patients to estimate the distinction between different anaesthesia levels since they show clear separation boundaries, can help determine the sensitivity level of MSE for the generated time indices, and for the neural network training. The other set of recordings was used for the testing and validation of the proposed neural prediction system.

The EEG signals were recorded with a sampling frequency of 128 Hz (samples per second) and its microvolt readings were digitised using 12-bit analogue-to-digital converter, thus giving a scalar range between 0 and $2^{12} = 4096$ values. Figure 1 shows some samples for an EEG recording during general anaesthesia covering three anaesthetic levels in addition to a recovery level, as shown below.

2.3. Expert Assessment of Consciousness Level

During the period of anaesthesia, the patients were observed by two research nurses who carefully observed the signs of unconsciousness and recorded the vital parameters such as heart rate and arterial blood pressure besides other clinical signs. Further, they recorded the time points and durations of various anaesthetic events such as induction, muscle signs; stimulus and clinical signs of the patients; and the duration of surgery.

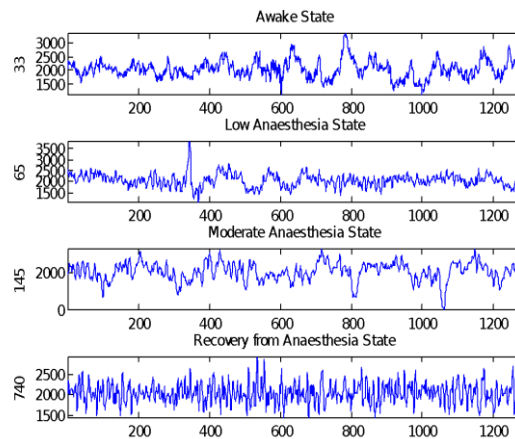


Figure1. Four single-channel EEG epochs, each of 10 seconds, representing brain function at different depths of anaesthesia. X-axis represents the epoch index (/10 seconds) and y-axis represents the captured epochs index (1-840).

Five expert anaesthesiologists were then asked to score the DoA of the patient on a scale of 10 to 100 based on the recordings of the patient’s signs only. A non-correlated and independent decision was then collected from each expert solely based on the data provided, without any contact with the patients. Finally, the average of the opinions of the 5 experts was determined by plotting on a graph, with time on the X-axis and their scores (from 10 to 100 for consistency the BIS-based anaesthesia level) on the Y-axis.

The standard for anaesthetic level defined as per the BIS machines is explained in Table 1. Values between 40 and 60 are defined as ‘anaesthetic depth suitable for surgery’, those below 40 indicate that the anaesthesia is deep, and those between 60 and 85 indicate that the anaesthesia is light and may be suitable only for some types of surgeries.

Table1. Range of BIS value for different levels for depths of anaesthesia

Depth of Anaesthesia (Consciousness Level)	Level Description	Range
Awake		BIS (100 – 80)
Light Anaesthesia		BIS 80 – 60
Moderate Anaesthesia	Suitable for Surgery	BIS 40 – 60
Deep Anaesthesia		BIS 10 -40

Finally, it is noteworthy that in this study, 20 anaesthesia charts were digitised and were used as a gold standard for determining the DoA and were used to train the neural network as training targets.

2.4. Sample and Multi-Scale Entropy

SampEn [44] is a modification of ApEn; both are used to examine physiological time series signals such as EEG and other signals. SampEn is independent of the data length and can be implanted more easily in software but is computationally different. In ApEn, the comparison between the template vector and the rest of vectors includes the comparison with itself, and it is always possible to have the logarithm of the probability.

SampEn is a measure of complexity but does not include self-similar patterns as found in ApEn. As in ApEn, if an embedding dimension m is given as well as tolerance r and number of the data points N, then we can define SampEn as negative logarithm of the probability that two sets of similar data points of length m have for distance < r. Then, two sets of similar data points of length m+1 also have distance < rare and represented by

$$\text{SampEn}(m, r, N), \text{ or by } \text{SampEn}(m, r, t, N)$$

Where t is the sampling time.

If a time-series data set of length $N = \{ x_1, x_2, x_3, \dots, x_N \}$ with a constant time interval τ , we define a template vector of length m, such that $X_m(i) = \{x_{im}, x_{i+1}, \dots, x_{i+m-1}\}$ have Chebyshev or Euclidean

distance. Then, compute the total distances $\langle R = r \cdot SD$ (standard deviation of original signal). The distance function has the formula:

$$d[X_m(i), X_m(j)] < r, \text{ and } i \neq j \tag{1}$$

We compute the number of vector pairs in template vector of length m and $m+1$ having $d[X_m(i), X_m(j)] < r$. We can define the sample entropy:

$$\text{SampEn} = -\log(A/B) \tag{2}$$

where

A = number of template vector pairs having $d[X_m(i), X_m(j)] < r$ of length $m+1$ and

B = number of template vector pairs having $d[X_m(i), X_m(j)] < r$ of length m .

Therefore, $\text{SampEn}(m, r, \tau)$ will be always be either zero or any other positive value. A smaller value of SampEn also indicates more self-similarity in data set or less noise.

Previous studies on SampEn have been shown that $m = 1$ or 2 would be appropriate; further, having $r = 0.1$ to $0.2SD$ ($SD = \text{Standard Deviation}$) are more suitable values for clinical data analysis.

Traditional entropy measures quantify only the regularity (predictability) of time series based on a single scale. There is no direct correlation between regularity and complexity. Neither completely predictable or periodic signals, which have minimum entropy, nor completely unpredictable or uncorrelated random signals, which have maximum entropy, are actually complex. There is no consensus on the definition of complexity. Intuitively, complexity can be defined as the correlation over time and space scales.

The MSE method incorporates two aspects. First, a coarse graining process is applied to the time series data by averaging data within non-overlapping windows but of increasing length. The coarse-grained element $y_j^{(\tau)}$

$$y_j^{(\tau)} = \frac{1}{\tau} \sum_{i=(j-1)\tau+1}^{j\tau} x_i \tag{3}$$

Where τ represents the scale factor and $1 \leq j \leq N/\tau$. The length of each coarse-grained time series is N/τ .

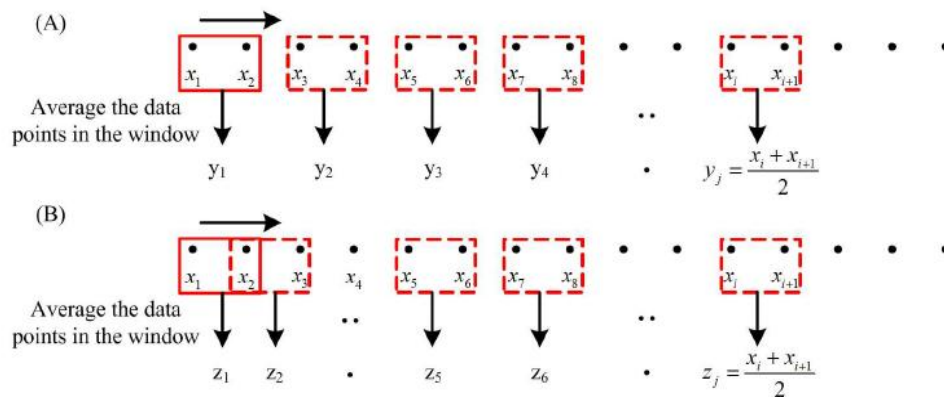


Figure 2. An illustrations of multiscale decomposition procedures. (a) Coarse-graining procedure and (b) Moving average procedure [45].

Second, SampEn is calculated for each coarse-grained time series and then plotted as a function of the scale factor. SampEn is a ‘regularity statistic’. It ‘looks for patterns’ in a time series and quantifies its degree of predictability or regularity (see Figure 2).

2.5. Empirical Mode Decomposition (EMD)

EMD is an adaptive method to decompose a non-linear or non-stationary time series signal into its intrinsic mode functions (IMF) [46]. This process to obtain the IMFs is called sifting, and it can be outlined into the next phases:

1. Create upper envelope $E_u(t)$ by local maxima and lower envelope $E_l(t)$ by local minima of data $x(t)$.

2. Calculate the mean of upper and lower envelope

$$m_1(t) = \frac{E_{u1}(t)+E_{l1}(t)}{2} \tag{4}$$

3. Subtract the mean from original data

$$h_1(t) = x(t) - m_1(t) \tag{5}$$

4. Verify that $h_1(t)$ satisfies conditions for IMFs. Repeat steps 1 to 4 with $h_1(t)$, until it is an IMF.

5. Get first IMF (after k iterations)

$$c_1(t) = h_{1(k-1)}(t) - m_{1k}(t) \tag{6}$$

6. Calculate first residue

$$r_1(t) = x(t) - c_1(t) \tag{7}$$

7. Repeat whole algorithm with $r_1(t), r_2(t), \dots$ until residue is monotonic function.

8. After n iterations, $x(t)$ is decomposed per equation (1).

The decomposition process can be stopped when r_n becomes a monotonic function, i.e., no more IMF can be extracted from it. Nevertheless, it is known that only a determined number of IMFs have physical meaning, and it is only important to apply a certain number of IMFs to extract relevant information from the main signal.

2.6. Artificial Neural Networks (ANNs)

ANNs are considered to be excellent classifiers due to their properties of adaptive learning, robustness to signal distortion, and good generalisation effect. For the classification of the anaesthetic level, a Feed-Forward Back Propagation Multi-Layer Perceptron was chosen. This particular type of neural networks is considered a ‘*de facto*’ and is widely known in neural classification and recognition types of applications. Since it has multiple layers and nonlinear activation functions (tan sigmoid or log sigmoid), the network can learn and recognise nonlinear beside linear correlations between the input and the output sets. The proposed neural network is shown in Figure 3.

The feed-forward network has one or more hidden layers beside the input and output layers. With those multiple layers of neurons, each utilising non-linear activation function enables the network to learn non-linear relation(s) between the input and output sets. Each neuron on different layers has a standard architecture which depends on storing multiple weights, a non-linear activation function, and a single bias. The following equation simplifies the equation core of the neurons:

$$Y = f(\sum_{j=1}^N w_j \cdot I_j) \tag{8}$$

where f is the required activation function that could be one of the next two functions: *LogSig* activation function or *TanSig* log activation function.

$$LogSig(n) = 1 / (1 + \exp(-n)) \tag{9}$$

$$TanSig(n) = 2/(1+\exp(-2*n))-1 \tag{10}$$

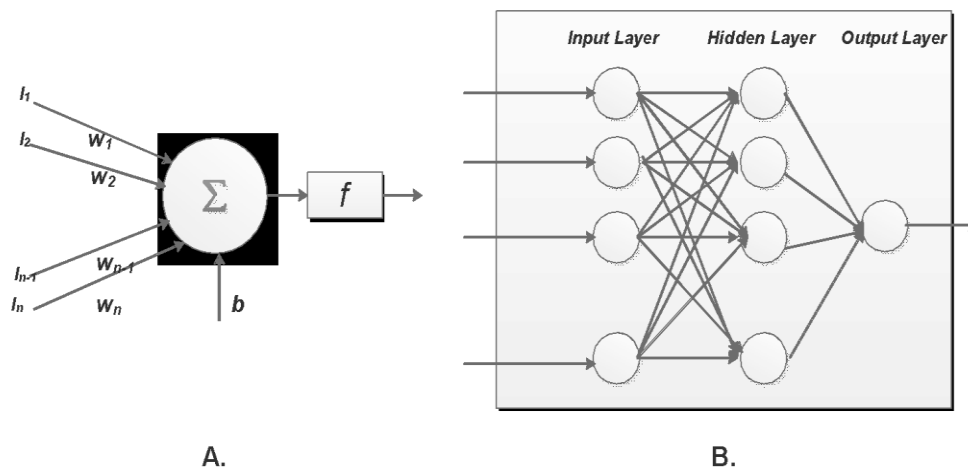


Figure3. Neural network for recognition of depth of anaesthesia. (A) shows one perceptron; while (B) shows a 3-layers multi-perceptron neural network.

3. RESULTS

In this study, some aspects of the raw EEG recording have prompted us to perform an early sorting of both trained and tested EEG recordings. The sorting was necessary because of the nature of the EEG signals itself. The signals showed 2 unpleasant properties: the artefact contamination from the patient signals themselves, as observed in electrocardiograms and electromyograms, as well as some signal discontinuity experienced during the EEG recording due to the patient’s movements as well as re-fixing of the electrodes and other unknown sources.

The artefacts and non-continuity in the recorded signal distort the original signal to a level that the *SampEn* gives misleading outputs, while bearing in mind the high sensitivity of such *SampEn* algorithm toward noise. However, considering the cleaning of the EEG recording, the application of EMD (Empirical Mode Decomposition) or any of its counters (such as EEMD, CEEMD, or MEMD) for filtration would shorten the amplitude of the signal such that it makes it difficult to discriminate between anaesthetic levels copied from *SampEn* by visual inspection although these EMD algorithms can clean and filter the signal to some extent. Because of this property, it was decided to compare two cases simultaneously: one with generic EMD filtering and one with no filtering.

3.1. Phase_1: Choosing the Appropriate Sections of the EEG Signals

All the patients’ EEG recording were carefully screened in two stages – first by visual inspection with the help of *Matlab* plotting method to identify then discard all the distorted, the non-continuous, clipped, and weak (less than threshold) epochs and then a second phase when *SampEn* of the whole patient is experimentally calculated and the Doctor Plot for consciousness level (1 to 100) is brought. From some careful entropy screening, it was noted that the entropy of the patient will change between 0.1 and 2.2. Next, all the patterns within specific regions where *SampEn* did not match with the expert’s plot were removed, and the remaining blocks which abide with the expert’s plot were taken to complete the training set with. Table-6 lists all the 6 patients’ appropriate blocks that were taken in the training process. From the table, it can be seen that the epoch is a 10-second recording of EEG signal yielding 1280 samples with 128 Hz sampling frequency.

Table2. Range of BIS value taken for different levels of depth of anaesthesia

Depth of Anaesthesia (Consciousness Level)	Level Description	Range
A wake (C1)		BIS (85 - 100)
Light Anaesthesia (C2)		BIS (60- 85)
Moderate Anaesthesia (C3)	Suitable for Surgery	BIS (40 - 60)

The common anaesthetic level which classifies the patient depth of consciousness is detailed in Table 1. However, since all our patient’s records showed that *DoA* is between 45 and 100, we have reduced the levels to only three, as listed in Table-2.

3.2. Phase_2: Statistical Analysis of the Trained Set:

Statistical analysis was carried out on the selected training patterns. First, we extracted the standard deviation, mean, covariance, and the maximum with the minimum coverage of all the samples in each anaesthesia level.

The MSSE measurements were calculated on 6 patients’ EEG records at 5 scales for the DoA. Measurements were computed over a window of 10 seconds with no overlap with and without EMD filtering. Figure 4 shows the MSSE for 4 patients with and without EMD filtering.

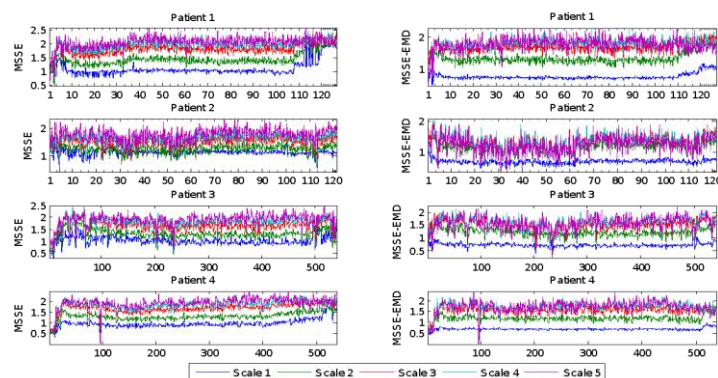


Figure4. Multi-scale sample entropy for four patients with and without EMD

To clearly investigate the trends in signal variation, median values of all the indices from the 6 patients in two states (Awake and Unconscious) were plotted together for comparison (Figure 5) and visualised by boxplot plotting. Figure 5 has 4 subplots; the first row has 2 plots of Awake and ‘Unconscious’ when no EMD filtering was applied. It is obviously seen that the mean for the ‘Conscious’ state was around 1.0 and almost exhibits no change when scales 2 to 5 are observed. The variation range (minimum to maximum) becomes relatively larger from scale 1 to 5.

The bottom part of the Figure shows a similar boxplot for the sample entropy for 5 scales with EMD filtering. It is seen that the trend of the 5 means are clearly different for the ‘Awake’ state: starts from $SampEn = 0.8$ for scale 1 and increases until scale 3; then, from scale 4 to scale 5, it decreases to around $SampEn = 1.5$. The ‘Unconscious’ mean level trend has a logarithmic shape; it starts from 0.70 for scale number 1; raises until it settles near a value of 1.5 with scale numbers from 2 to 5; and then it settles.

It is evident from the previous boxplot that the boundaries of the $SampEn$ for the 3 classes are distinguishable for the time scale 1 as well as indices 2 and 3. Table 3 outlines the means, standard deviations, covariance, and max-min values of the 3 state boundaries. It is also noticeable that the variation of $SampEn$ increases as the scale number increases.

Table 3. Different statistical properties outlining the 3 different classes of anaesthesia

Statistical Measurement Name	Class 1	Class 2	Class 3
Mean	1.5117	1.2545	0.9887
	1.4395	1.6493	1.2573
	1.6575	1.9802	1.5832
Standard Deviation	0.0864	0.0478	0.1188
	0.2209	0.1855	0.1120
	0.2980	0.1545	0.1420
Covariance	0.0075	0.0023	0.0141
	0.0488	0.0344	0.0125
	0.0888	0.0239	0.0202
Lower to Upper Bound	1.3959-1.7017	1.1455-1.3789	0.6964-1.2455
	1.1396-1.9828	1.4138-	0.8152-1.5135
	1.1876-2.1353	2.25441.5571-2.4343	0.8060-1.9213

This verified that entropy (or scale 1 MSSE) can make a low distinction between the ‘Awake’, ‘Light Anaesthesia’, and ‘Unconscious’ states, although the best distinction is noticeable between the ‘Awake’ and ‘Unconscious’ states. Nevertheless, MSSE has much better distinction ability between these two states when all five scales are observed. The neural network classifier (mentioned in the next few section) will play the main role in recognising different states of anaesthesia based on different scales of the MSSE.

3.3. Phase_3: Organising and Sorting the Extracted Epochs for Training

To accurately extract the best representation for DoA from EEG recordings, careful selection of the appropriate epochs should be undertaken. Sequences of EEG Blocks should be chosen for artefact-clean samples and those with good $SampEn$ readings that strongly correlate to the expert’s decision about the DoA .

As shown from Table 1, the data of the 6 selected patients (Table 6) had recordings that clearly cover the 100 to 40 BIS level readings with sufficient number of blocks covering the 3 chosen consciousness states, in order to ensure the general diversity needed for successful training. Besides, the chosen blocks for the same class showed some acceptable correlation in $SampEn$ values.

From the six patients’ EEG recordings, we obtained only 169 blocks of EEG recording, each of 10 seconds, for the first class (BIS 100 to 85); 224 blocks for the second class (BIS 85 to 62) and 770 blocks representing the third class (BIS 40 to 60). Then, all those selected epochs were inserted in 2-D arrays (3-time indexed scaled entropy readings for each block giving an array depth of 1163 lines by 3 columns), which is called the main training database (MTDB). A corresponding output vector has been prepared such that it gives a reading of 0.9 corresponding to input class one, a reading of 0.5 for a class two, and a reading of 0.25 for class three. This is the vector target that the training should reach within a small error boundary. That output vector will be added to the reconstructed MTDB table (Table 6).

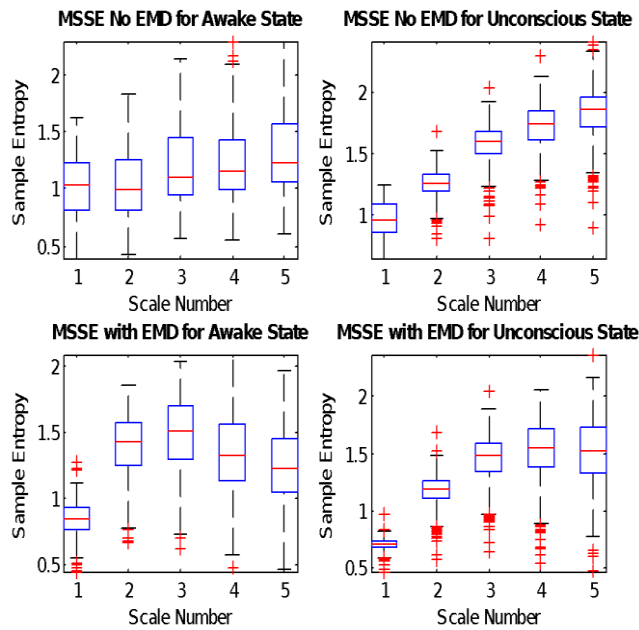


Figure5. The statistical box plots for MSSE with EMD and without EMD, at scales 1, 2, 3, 4, and 5 in the awake state (upper row) and unconsciousness state (lower row).

3.4. Phase_4: Training the network

From the MSE values, only 3 time indices levels were considered as inputs for the ANN(NN Tools the neural network toolbox in *Matlab*). The output target was the depth of anaesthesia score assigned by the experts in the plot after digitisation. Data from six patients out of 20 were collected to train the ANN with, and data from the remaining 14 patients were used to test the ANN.

A multi-layer perceptron network was chosen for its good generalisation effect, with 30 hidden neurons in its first layer employing TAN activation function and another 50 neurons in the second layer with TAN activation function. A single neuron in the output layer with LOG activation function was dedicated to give the output ranges between 0.1 to 0.9.

The network was trained using back propagation method with a TRAINRP training function and LEARNGDM for its adaptation learning function. The performance function was MSE (Mean Square Error). The epochs reached and the training goal for the three networks is listed in Table 4.

Table4. Feed-Forward Back-Propagation Neural Network training parameters

Description	Neural Network 1	Neural Network 2	Neural Network 3
Learning Criteria	3 Levels, MSSampEn, EMD	2 levels, MSSampEn, No EMD	2 levels, MSSampEn, EMD
Anaesthetic Levels	3	2	2
Number of Layers	1+2+1	1+2+1	1+2+1
Neurons in Input Layer	3	3	3
Neurons in Hidden Layer 1 and 2	30, 50 (<i>Tan</i>)	30, 50 (<i>Tan</i>)	30, 50 (<i>Tan</i>)
Neurons in Output Layers	1(<i>Log</i>)	1(<i>Log</i>)	1(<i>Log</i>)
Training Epochs	75,776	108	100,000
Training Goal Reached	0.00406	0.00048	0.0021

3.5. Phase_5: Testing the network on the remaining patients’ records

The neural network was trained with the MSE entropy values of 3 time scales obtained from pre-processed EEG segments (epochs of 1280 sample of 10 s duration). A total number of 2639 prepared testing patterns of 10-second duration corresponding to the remaining 15 patients were used as a test data set for evaluating the performance of the neural network.

The classification results are shown in Table-5 for three networks: NW1 is a model for recognising 3 levels of anaesthesia when no EMD filtering is enabled; NW2 is a model for recognising 2 levels of

anaesthesia without embedding EMD filtering; and NW3 is the same as NW2 but EMD filtering is implemented. Figure 6 shows the output of the neural network model corresponding to the 2 and 3 levels of anaesthesia when tested with the unseen test data of 2639 test patterns. Figure 7 shows the first 400 patterns of each output.

In addition, Table 5 shows the accuracy of the recognition for each anaesthetic level. The performance of the ANN is evaluated using two parameters: sensitivity for each anaesthetic level and overall accuracy. The following equations show the performance evaluation parameters formulae.

$$Sensitivity_{Awake} = \frac{\text{Correctly detected Awake anaesthesia}}{\text{Total of actual Awake anaesthesia}} \cdot 100 \tag{11}$$

$$Sensitivity_{Low\ Anaesthesia} = \frac{\text{Correctly detected Low anaesthesia}}{\text{Total of actual Low anaesthesia}} \cdot 100 \tag{12}$$

$$Sensitivity_{Moderate\ Anaesthesia} = \frac{\text{Correctly detected Moderate anaesthesia}}{\text{Total of actual Moderate anaesthesia}} \cdot 100 \tag{13}$$

$$Overall\ Accuracy_{Total} = \frac{\text{Total Correctly detected anaesthesia levels}}{\text{Total of applied anaesthesia levels}} \cdot 100 \tag{14}$$

Table 5. Results of classification of depth of anaesthesia

Anaesthetic Level Class	Total Applied Patterns	Correctly Detected Patterns with Class Sensitivity		
		NW1	NW2	NW3
Wake BIS100-85	98	85 (86%)	73 (74%)	90(92%)
Light Anaesthesia BIS 85-62	68	18 (26%)	---	---
Surgery Anaesthesia BIS 60-40	2473	1801 (72%)	2250 (91%)	2347 (95%)
Over all Accuracy with Sensitivity	2639	1904 (72%)	2323 (90%)	2437 (95%)

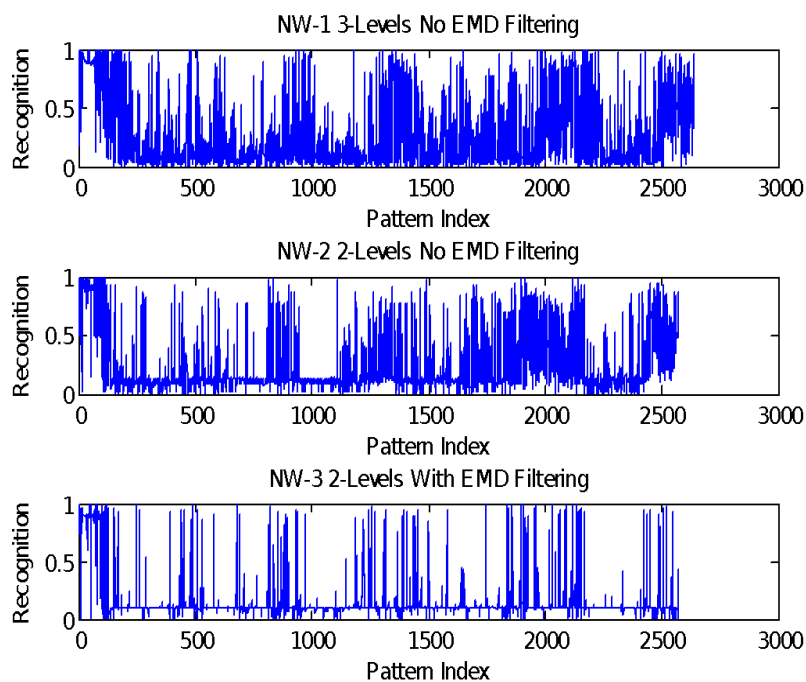


Figure 6. Neural network output for the three networks: 3 levels without filtering and 2 levels without and with filtering.

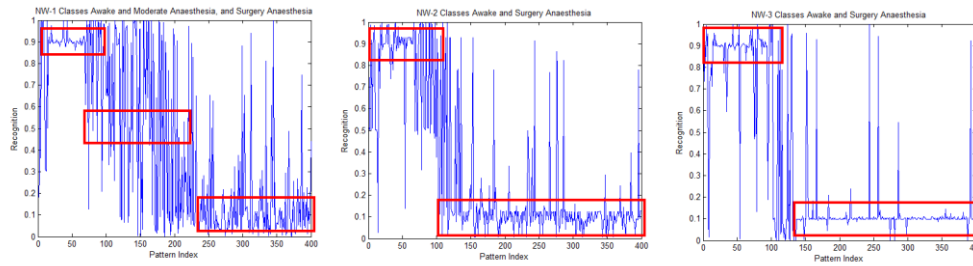


Figure7. The 3 network outputs for the 400 test patterns: patterns [1-98] for class ‘Awake’, and patterns [99-400] for class ‘Moderate Anaesthesia’, while those for the first subplot 3 classes are represented; and finally the previous two beside ‘Light Anaesthesia’ [99-167]. The red rectangles shows the three classes for each network.

Table6. Records for the six selected patients that were used for training, with their IDs, total epochs, surgery time, and the chosen epochs indices. Numbers (1, 2, and 3) in small font indicate the DoA three levels.

Event	Patient EEG Record ID	Total number of Epochs	Surgery Time	Actual Epochs Range*	Appropriate Epochs Extracted (1280 Samples each)
Patient 1	1017-61-1106	918	11:07-13:40	[1 to 60] ¹ [61 to 75] ² [780 to 840] ² [76 to 285] ³ [286 to 675] ³	[23, 33, 41, 55, 58, 56, 57, 60, 61] ¹ [65, 69, 81] ² , [728 to 751] ² [767 to 772] ² , [802 to 839] ² [90 to 95] ³ , [120 to 123] ³ [109 to 112] ³ , [128 to 141] ³ [145 to 188] ³ , [192 to 243] ³ [274 to 303] ³ , [308 to 361] ³ [370 to 412] ³ , [370 to 412] ³ [431 to 478] ³ , [486 to 515] ³ [548 to 560] ³ , [579 to 597] ³ [625 to 637] ³
Patient 2	1017-61-1348	954	13:51 – 16:30	[1 to 30] ¹ [35 to 45] ² [48 to 894] ³	[14, 17, 21, 23, 31] ¹ [278 to 302] ³ , [308 to 346] ³ [452 to 481] ³ , [500 to 541] ³ [552 to 601] ³ , [611 to 659] ³ [704 to 721] ³ , [728 to 753] ³
Patient 3	1104-23-1345	816	13:46 – 15:30	[1 to 38] ¹ [42 to 63] ² [588 to 630] ² [72 to 573] ³	[48, 50, 62, 63, 67] ² [576, 578, 582, 585, 586] ² [587, 600] ² [322 to 376] ³ , [392 to 425] ³ [456 to 490] ³ , [517 to 535] ³
Patient 4	1017-65-1436	594	14:36 – 16:15	[1 to 12] ¹ [426 to 594] ² [24 to 407] ³	[24] ¹ [21, 22, 23, 27, 28, 29] ² [490, 497, 501, 507] ² [510 to 513] ² , [178 to 227] ³ , [240 to 272] ³ [300 to 346] ³ , [356 to 376] ³ [392 to 407] ³
Patient 5	1029-65-1010	576	10:09 – 11:45	[1 to 115] ¹ [120 to 132] ² [520 to 570] ² [137 to 509] ³	[29, 77, 116, 121] ¹ [[142] ² [285 to 303] ³ , [311 to 342] ³ [368 to 381] ³ , [430 to 482] ³
Patient 6	1107-17-1334	486	13:34 – 14:55	[1 to 65] ¹ [72 to 77] ² [[85 to end] ³	[33, 49] ¹ , [51 to 54] ¹ [145 to 219] ³ , 267 to 283] ³ [310 to 349] ³

As shown in Table 5, NW-1 showed weak recognition (26%) for the Class-2 ‘Light Anaesthesia’, but showed much better accuracy for Classes 1 and 3, ‘Awake’ and ‘Moderate Anaesthesia’, respectively. The reason comes from the close values for the 3 scales of the *SampEn* measures. The difference between the ‘Awake’ and the ‘Moderate Anaesthesia’ is wide. Therefore, the accuracy for Class-2 was 18% and that for the remaining two classes, ‘Awake’ and ‘Moderate anaesthesia’, were 86% and 72%, respectively.

For the other two networks under study, NW-2 and NW-3, which were trained for only two levels without EMD filtering, i.e., the 'Awake' and the 'Moderate anaesthesia', the results showed more accuracy. In particular, accuracy for Class-1 and Class-3 was 74% and 91%, respectively. For NW-3, which was trained for 2 levels with EMD filtering, the accuracy of classification reached 92% to 95%.

The overall accuracy is valued as 72%, 90%, and 95% for NW-1, NW-2, and NW-3, respectively.

4. CONCLUSION

In summary, we propose a new method to assess the DoA. The main feature of this method is the use of EMD to filter the EEG set in its generic form along with the use of MSE with *SampEn* and three time scales to obtain additional dimensions for better recognition of the three anaesthetic levels. Statistical analysis showed a distinct correlation between samples in the same class and distinction between the three different levels of consciousness. ANN was the sole classifier used in this method and its effectiveness was evaluated by testing on the EEG data of 14 of the 20 patients. An accuracy rate of over 95% was obtained using this method of using MSSE and EMD filtering with ANN classifier in detecting the 3 levels of DoA.

REFERENCES

- [1] C. D. Kent and K. B. Domino, "Depth of anaesthesia," *Current Opinion in Anaesthesiology* 22.6 (2009) 782-787.
- [2] Shieh, Jiann Shing, D.a. Linkens, and J.e. Peacock. "Hierarchical rule-based and self-organizing fuzzy logic control for depth of anaesthesia." *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 29.1 (1999): 98-109.
- [3] M. Jeanne, R. Logier, J. De Jonckheere, and B. Tavernier, "Heart rate variability during total intravenous anesthesia: effects of nociception and analgesia," *Autonomic Neuroscience: Basic and Clinical*, vol. 147, no. 1-2, pp. 91-96, 2009.
- [4] Sigl, Jeffrey C., and Nassib G. Chamoun. "An introduction to bispectral analysis for the electroencephalogram." *Journal of Clinical Monitoring* 10.6 (1994): 392-404.
- [5] C. Rosow and P. J. Manberg, "Bispectral index monitoring," *Anesthesiology Clinics of North America*, 19.4(2001) 947-966.
- [6] Punjasawadwong, Y., A. Phongchiewboon, N. Bunchungmonkol, and Dn Braaksma. "Bispectral index for improving anaesthetic delivery and postoperative recovery." *The Cochrane Database of Systematic Reviews* (2002).
- [7] Su, Cui, Zhenhu Liang, Xiaoli Li, Duan Li, Yongwang Li, and Mauro Ursino. "A Comparison of Multiscale Permutation Entropy Measures in On-Line Depth of Anesthesia Monitoring." *Plos One* 11.10 (2016).
- [8] Benzy, V. K., and E. A. Jasmin. "A Combined Wavelet and Neural Network Based Model for Classifying Depth of Anaesthesia." *Procedia Computer Science* 46 (2015): 1610-1617.
- [9] Lalitha, V., and C. Eswaran. "Automated detection of anesthetic depth levels using chaotic features with artificial neural networks." *Journal of Medical Systems* 31.6 (2007): 445-452.
- [10] Srinivasan, V., C. Eswaran, and N. Sriraam. "EEG based automated detection of anesthetic levels using a recurrent artificial neural network." *International Journal of Bioelectromagnetism* 7.2 (2005): 267-270.
- [11] Coşkun, Mustafa, et al. "Determining the appropriate amount of anesthetic gas using DWT and EMD combined with neural network." *Journal of Medical Systems* 39.1 (2015): 1-10.
- [12] Kreuzer, Matthias, Eberhard F. Kochs, Gerhard Schneider, and Denis Jordan. "Non-stationarity of EEG during wakefulness and anaesthesia: advantages of EEG permutation entropy monitoring." *Journal of Clinical Monitoring and Computing* 28.6 (2014): 573-80.
- [13] Costa, Madalena, Ary L. Goldberger, and C-K. Peng. "Multiscale entropy analysis of complex physiologic time series." *Physical Review Letters* 89.6 (2002): 068102.
- [14] Ortolani, O. "EEG signal processing in anaesthesia. Use of a neural network technique for monitoring depth of anaesthesia." *British Journal of Anaesthesia* 88.5 (2002): 644-648. Web.
- [15] Lavrov, N. G., V. V. Bulaev, E. N. Solouhin, S. A. Taratuhin, and A. V. Chistyakov. "Bispectral Analysis of Electroencephalogram Using Neural Networks to Assess the Depth of Anesthesia." *Biomedical Engineering* 49.6 (2016): 380-84.

- [16] Li, Tianning, and Peng Wen. "Depth of anaesthesia assessment based on adult electroencephalograph beta frequency band." *Australasian Physical & Engineering Sciences in Medicine* 39.3 (2016): 773-781.
- [17] Güntürkün, Rüştü. "Determining the amount of anesthetic medicine to be applied by using Elman's recurrent neural networks via resilient back propagation." *Journal of Medical Systems* 34.4 (2010): 493-497.
- [18] Srinivasan, V., C. Eswaran, and N. Sriraam. "EEG based automated detection of anesthetic levels using a recurrent artificial neural network." *International Journal of Bioelectromagnetism* 7.2 (2005): 267-270.
- [19] Liang, Zhenhu, Yinghua Wang, Xue Sun, Duan Li, Logan J. Voss, Jamie W. Sleight, Satoshi Hagihira, and Xiaoli Li. "EEG entropy measures in anesthesia." *Frontiers in Computational Neuroscience* 9 (2015).
- [20] Ghanatbari, M., A. R. Mehri Dehnavi, H. Rabbani, and A. R. Mahoori. "Estimating the depth of anesthesia by applying sub parameters to an artificial neural network during general anesthesia." *2009 9th International Conference on Information Technology and Applications in Biomedicine* (2009).
- [21] Musizza, Bojan, and Samo Ribaric. "Monitoring the depth of anaesthesia." *Sensors* 10.12 (2010): 10896-10935.
- [22] Whyte, Simon David, and Peter Driscoll Booker. "Monitoring depth of anaesthesia by EEG." *BJA CEPD Reviews* 3.4 (2003): 106-110.
- [23] Costa, Madalena, Ary L. Goldberger, and C-K. Peng. "Multiscale entropy analysis of biological signals." *Physical Review E* 71.2 (2005): 021906.
- [24] Kreuzer, Matthias, Eberhard F. Kochs, Gerhard Schneider, and Denis Jordan. "Non-stationarity of EEG during wakefulness and anaesthesia: advantages of EEG permutation entropy monitoring." *Journal of Clinical Monitoring and Computing* 28.6 (2014): 573-80.
- [25] Palendeng, Mario Elvis, Peng Wen, and Yan Li. "Real-time depth of anaesthesia assessment using strong analytical signal transform technique." *Australasian Physical & Engineering Sciences in Medicine* 37.4 (2014): 723-730.
- [26] Manilo, L. A., and S. S. Volkova. "Recognition of the deep anesthesia stage from parameters of the approximated entropy of EEG signal." *Pattern Recognition and Image Analysis* 23.1 (2013): 92-97.
- [27] Schmidt, Gunter N., Petra Bischoff, Thomas Standl, Angelika Hellstern, Olaf Teuber, and Jochen Schulte Esch. "Comparative Evaluation of the Datex-Ohmeda S/5 Entropy Module and the Bispectral Index?? Monitor during Propofol???Remifentaniil Anesthesia." *Anesthesiology* (2004): 1283-290.
- [28] Viertio-Oja, H., V. Maja, M. Sarkela, P. Talja, N. Tenkanen, H. Tolvanen-Laakso, M. Paloheimo, A. Vakkuri, A. Yli-Hankala, and P. Merilainen. "Description of the Entropytm algorithm as applied in the Datex-Ohmeda S/5tm Entropy Module." *Acta Anaesthesiologica Scandinavica* 48.2 (2004): 154-61.
- [29] Ferenets, Rain, Ann Vanluchene, Tarmo Lipping, Bj??rn Heyse, and Michel M. R. F. Struys. "Behavior of Entropy/Complexity Measures of the Electroencephalogram during Propofol-induced Sedation." *Anesthesiology* 106.4 (2007): 696-706.
- [30] Sleight, J. W., and J. Donovan. "Comparison of bispectral index, 95% spectral edge frequency and approximate entropy of the EEG, with changes in heart rate variability during induction of general anaesthesia." *British Journal of Anaesthesia* 82.5 (1999): 666-71.
- [31] Richman, J. S., and J. R. Moorman. "Physiological time-series analysis using approximate and sample entropy." *American Journal of Physiology—Heart and Circulatory Physiology* 278.6 (2000): H2039-2049.
- [32] Bruhn, Jürgen, Heiko Röppcke, and Andreas Hoeft. "Approximate Entropy as an Electroencephalographic Measure of Anesthetic Drug Effect during Desflurane Anesthesia." *Anesthesiology* 92.3 (2000): 715-26.
- [33] Takahashi, Tetsuya, Raymond Y. Cho, Tomoyuki Mizuno, Mitsuru Kikuchi, Tetsuhito Murata, Koichi Takahashi, and Yuji Wada. "Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: A multiscale entropy analysis." *NeuroImage* 51.1 (2010): 173-82.

- [34] G. Ouyang, C. Dang, and X. Li, "Multiscale entropy analysis of EEG recordings in epileptic rats," *Biomedical Engineering: Applications, Basis and Communications*, vol. 21, no. 3, pp. 169–176, 2009.
- [35] Liu, Quan, Qin Wei, Shou-Zen Fan, Cheng-Wei Lu, Tzu-Yu Lin, Maysam F. Abbod, and Jiann-Shing Shieh. "Adaptive Computation of Multiscale Entropy and Its Application in EEG Signals for Monitoring Depth of Anesthesia During Surgery." *Entropy* 14.12 (2012): 978-92.
- [36] R. Nayak, L. Jain, and B. Ting, "Artificial neural networks in biomedical engineering: a review," in *Proceedings of the Asia-Pacific Conference on Advance Computation*, 2001.
- [37] Bruhn, Jürgen, Thomas W. Bouillon, Lucian Radulescu, Andreas Hoeft, Edward Bertaccini, and Steven L. Shafer. "Correlation of Approximate Entropy, Bispectral Index, and Spectral Edge Frequency 95 (SEF95) with Clinical Signs of "Anesthetic Depth" during Coadministration of Propofol and Remifentanyl." *Anesthesiology* 98.3 (2003): 621-27.
- [38] Wei, Qin, Quan Liu, Shou-Zhen Fan, Cheng-Wei Lu, Tzu-Yu Lin, Maysam Abbod, and Jiann-Shing Shieh. "Analysis of EEG via Multivariate Empirical Mode Decomposition for Depth of Anesthesia Based on Sample Entropy." *Entropy* 15.9 (2013): 3458-470.
- [39] Huang, Jeng-Rung, Shou-Zen Fan, Maysam Abbod, Kuo-Kuang Jen, Jeng-Fu Wu, and Jiann-Shing Shieh. "Application of Multivariate Empirical Mode Decomposition and Sample Entropy in EEG Signals via Artificial Neural Networks for Interpreting Depth of Anesthesia." *Entropy* 15.9 (2013): 3325-339.
- [40] Wei, Qin, Yang Li, Shou-Zen Fan, Quan Liu, Maysam F. Abbod, Cheng-Wei Lu, Tzu-Yu Lin, Kuo-Kuang Jen, Shang-Ju Wu, and Jiann-Shing Shieh. "A critical care monitoring system for depth of anaesthesia analysis based on entropy analysis and physiological information database." *Australasian Physical & Engineering Sciences in Medicine* 37.3 (2014): 591-605.
- [41] Liu, Quan, Qin Wei, Shou-Zen Fan, Cheng-Wei Lu, Tzu-Yu Lin, Maysam F. Abbod, and Jiann-Shing Shieh. "Adaptive Computation of Multiscale Entropy and Its Application in EEG Signals for Monitoring Depth of Anesthesia During Surgery." *Entropy* 14.12 (2012): 978-92.
- [42] Shalbaf, R., H. Behnam, J. Sleight, and L. Voss. "Measuring the effects of sevoflurane on electroencephalogram using sample entropy." *Acta Anaesthesiologica Scandinavica* 56.7 (2012): 880-89.
- [43] S. Samara Singhe, *Neural Networks for Applied Sciences and Engineering: Form Fundamental to Complex Pattern Recognition*, Taylor & Francis, Boca Raton, Fla, USA, 2007.
- [44] https://en.wikipedia.org/wiki/Sample_entropy.
- [45] Su C, Liang Z, Li X, Li D, Li Y, Ursino M. "A comparison of multiscale permutation entropy measures in on-line depth of anesthesia monitoring". *PloS one*. 2016 Oct 10;11(10):e0164104.
- [46] Trnka P, Hofreiter M. "The empirical mode decomposition in real-time", In *Proceedings of the 18th international conference on process control, Tatranská Lomnica, Slovakia* (2011), 14-17.

AUTHOR'S BIOGRAPHY

Dr. Mohamad Almeer, is a Qatari scholar. He received his PhD degree from the University of Wales, UK in 2000 and his Master degree from Wayne State University, Michigan, U.S.A 1996. He joined Qatar University in 2000 and chaired several committees at Qatar University and outside. He is consultant for several entities in Doha, Qatar. Dr. Almeer worked on digital systems, map and data reduction and management and published several journals papers in this area. Dr. Almeer's research interest includes, neural networks, Digital Systems, security, high performance computing.