Combining Genetic Algorithm and Levenberg-Marquardt in Neural Network Training for Hypoglycemia Detection using EEG Signals

Lien B. Nguyen, Anh V. Nguyen, Sai Ho Ling, *Senior Member, IEEE* and Hung T. Nguyen, *Senior Member, IEEE*

***Abstract*— Hypoglycemia is the most common but highly feared complication induced by the intensive insulin therapy in patients with type 1 diabetes mellitus (T1DM). Nocturnal hypoglycemia is dangerous because sleep obscures early symptoms and potentially leads to severe episodes which can cause seizure, coma, or even death. It is shown that the hypoglycemia onset induces early changes in electroencephalography (EEG) signals which can be detected non-invasively. In our research, EEG signals from five T1DM patients during an overnight clamp study were measured and analyzed. By applying a method of feature extraction using Fast Fourier Transform (FFT) and classification using neural networks, we establish that hypoglycemia can be detected efficiently using EEG signals from only two channels. This paper demonstrates that by implementing a training process of combining genetic algorithm and Levenberg-Marquardt, the classification results can be improved markedly up to 75% sensitivity and 60% specificity on a separate testing set.**

I. Introduction

For patients with type 1 diabetes mellitus (T1DM), the insulin therapy is crucial because the hormone is no longer produced by patients' bodies internally. However, this therapy induces an increased risk of hypoglycemia by three times [1]. Hypoglycemia is the medical term for the state produced by an abnormally low blood glucose level (BGL). This is considered as the most common but highly severe complication for patients with T1DM and a limiting factor of the intensive insulin therapy.

Symptoms of hypoglycemia vary from mild to severe episodes. Mild hypoglycemia normally leads to early warning symptoms such as sweating, shaking, nervousness, heart plumping, confusion, etc. It can be easily alleviated by taking glucose-rich drink or food. If left untreated, it can become a severe episode which may progressively results in cognitive impairments, seizures, coma, and even death. A study in 2004 reported that severe hypoglycemia happens in one third of 1076 self-reported participants with an incidence rate of 1.3 episodes/patient-year [2]. Nocturnal hypoglycemia is especially dangerous because sleep reduces and obscures warning symptoms, so that an initially mild episode may become severe. It was reported previously that almost 50% of all episodes of severe hypoglycemia occur at night during sleep [3]. Because of its severity, intensive research has been devoted to the development of systems that can detect the onset of hypoglycemic episodes, and then give an alarm to provide enough time for patients and carers to take action.

Regarding correlation between hypoglycemia and electroencephalography (EEG) signals, since the EEG is directly related to the metabolism of brain cells, a failure of cerebral glucose supplying can cause early changes in EEG signals. A number of studies have reported important traces in EEG signals induced by hypoglycemia episodes in T1DM patients [4-6]. Recent studies also lead to acceptable results which show the potential ability to detect hypoglycemia from EEG signals [7-9]. Nevertheless, all of these results need to be improved further in order to be applied into the real clinical environment.

In term of classification algorithm, artificial neural networks have been employed popularly in biomedical area as a powerful tool [10]. It has been recognized that neural network can successfully classify complex situations and effectively model non-linear relationships between inputs and outputs. One of the most popular training techniques is Levenberg-Marquardt (LM) which is based on the second-order gradient information of an error function in order to direct the training process to a local optimal. Genetic algorithm (GA) is a derivative-free global search optimization which is inspired by the natural evolution. This technique has been applied widely in evolving neural network models which can efficiently drive the training process to the global optimal.

In our previous works, EEG signals of 5 T1DM children from a glucose clamp study were analyzed to find important spectral features to be used as inputs for a neural network based classification unit [9, 11]. This paper aims to propose a training strategy for neural network in order to enhance the performance of the developed classification algorithm. To do this, a combination of GA and LM is explored to utilize advantages as well as avoid limitations of each algorithm in training neural network. GA is used to locate the region of the global optimal consistently. LM algorithm acts as a fine tuner to help the training process quickly converge toward the global solution. Our main objective is to demonstrate that by applying a properly combined strategy to train neural network, the performance of hypoglycemia detection using only 2 EEG channels can be improved markedly. Section II provides an overview of the methodology used in our study. Results of the study will be mentioned in Section III. A conclusion for this study is drawn in Section IV.

Lien B. Nguyen, Anh V. Nguyen, Sai Ho Ling and Hung T. Nguyen are with the Centre for Health Technologies, Faculty of Engineering and Information Technology, University of Technology Sydney, Broadway, NSW 2007, Australia, (e-mail: BichLien.Nguyen@student.uts.edu.au; Anh.Nguyen-3@student.uts.edu.au; Steve.Ling@uts.edu.au; Hung.Nguyen@uts.edu.au).

II. METHODS

A. Study

Five T1DM adolescents (between the ages of 12 and 18 year old) volunteered for an overnight hypoglycemia study at the Princess Margaret Hospital for Children in Perth, Australia. During the study, EEG signals were continuously recorded and stored by a Compumedics system with the sampling rate of 128 Hz. The EEG electrodes were positioned at 4 channels O1, O2, C3 and C4 according to the International 10/20 system, referenced to A1 and A2. Also, we placed 2 electrodes at patients' chins to acquire the electro-myogram (EMG) signals and 2 electrodes near patients' eyes to measure the electro-oculogram (EOG) signals. The actual BGLs were routinely collected to be used as reference using Yellow Spring Instruments with the general sampling period of 5 minutes. Data were collected with the approval of the Women's and Children's Health Service, Department of Health, Government of Western Australia, and with informed consent.

B. Signal Processing and Feature Extraction

Raw EEG signals are filtered to get rid of unwanted artifacts by using IIR filters as well as a visually artifact-rejecting method based on the corresponding EMG and EOG signals. After being filtered, non-artifact EEG signals are segmented into 5-second non-overlapping epochs. These epochs are labeled as hypoglycemia or non-hypoglycemia according to the corresponding BGLs. Epochs which correspond with BGL lower than 3.3mmol/l are defined as hypoglycemia. Conversely, they are labeled as non-hypoglycemia. By using FFT, each epoch is transformed into the frequency domain which results in the power spectrum $P(f_i)$, with frequency resolution of 0.2 Hz. The power spectrum is then subdivided into 3 frequency bands: theta (θ: 4-8Hz), alpha (α: 8-13 Hz) and beta (β: 13-30Hz). From the power spectrum of each frequency band, two EEG features including power level and centroid frequency are extracted. Details about processing signals and extracting iEEG parameters were presented in previous papers [9, 11].

For comparison and classification purposes, two sets of data are extracted. The hypoglycemia set includes 188 data points; the non-hypoglycemia set includes 240 data points. Each final data point in each set is estimated as the average of two consecutive non-overlapping points. Statistical analysis is applied to compare and determine the significance of differences in EEG features between the hypoglycemia set and non-hypoglycemia set.

C. Classification

Considering the final aim of developing a real-time system which requires reducing the computational cost, in our study, a feed-forward three-layer neural network is developed as classification unit. The structure of the neural network is shown in Fig. 1. The input layer includes features extracted from EEG signals. The output layer consists of one output node which indicates the state of hypoglycemia or non-hypoglycemia. The desired output is set at 1 in case of hypoglycemia and -1 in the other case.

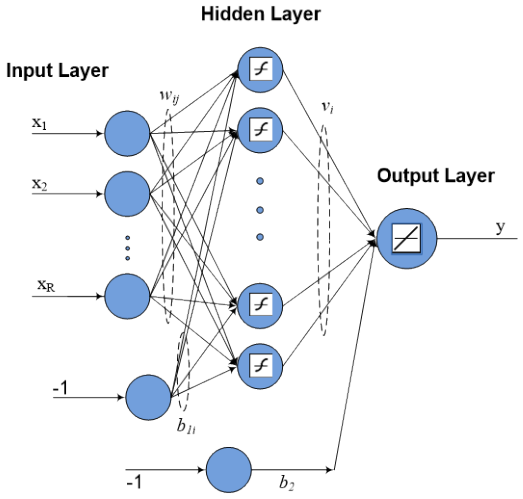

Fig. 2. Neural network structure

From Fig.1, the input-output relationship of the developed neural network can be written as follows:

$$y = \sum_{i=1}^{S} v_i \, tansig\left[\sum_{j=1}^{R}\left(w_{ij} x_j - b_{1i}\right)\right] - b_2 \qquad (1)$$

where w_{ij}, $i = 1, 2, ..., S$; $j = 1, 2, ..., R$ is the weight of the link between i-th hidden node and the j-th input; v_i is the weight of the link between i-th hidden node and the output; b_{1i}, b_2 are the biases for the input layer and hidden layer respectively; S is the number of hidden nodes; R is the number of inputs; $tansig$ is the hyperbolic tangent sigmoid transfer function of hidden layer.

In developing neural network, training algorithm plays the most important role in order to achieve a desired classification performance. In this paper, we explore a 2-step training process, which combines advantages of GA and LM. It is noted that the error function used for training is defined as the mean squared error (mse) of the output and its corresponding target. The number of hidden nodes is selected as the one which gives the best classification performances by the trials and errors method.

GA is used to evolve neural network parameters by searching over the whole domain and direct the training process to the global optimal region. First, a population of chromosomes or individuals is initialized. Each chromosome is expressed by $[w_{ij}\ v_i\ b_{1i}\ b_2]$. Thus, the length of chromosome is equal to the number of neural network parameters. During the evolution, each chromosome is evaluated by a fitness function which is defined as follows:

$$f(chromosome) = \frac{1}{1 + mse} \qquad (3)$$

At each iteration (or generation) of the training process, the population is updated through a process of selection, crossover and mutation. The selection chooses some chromosomes out of the population for reproduction based on fitness values of each chromosome in the population. The selected chromosomes undergo two genetic operators of crossover and mutation. Basically, the crossover operation helps combine information while the mutation operation helps change characteristics of two selected chromosomes to

generate offspring. The offspring after that are evaluated by the fitness function to replace the worst chromosomes in the old population to generate a new population. This process is repeated until one of terminating conditions is met. The best chromosome in the final population will be employed as the initial neural network parameters for the next step of local searching by the LM algorithm.

To overcome GA drawbacks of inefficient fine tuning and slow convergent rate, LM algorithm is implemented on the parameters sets obtained by GA. In brief, the LM algorithm estimates the second directional derivative of the performance function (mse function), in order to ensure the training process direct to a local optimal. To avoid overtraining which may cause bad generalization, the cost function on a validation set is also monitored during the training process. When the validation error keeps increasing for a given number of iterations, the training is stopped. The parameters at that stopped iteration will be used as final neural network weights and biases.

After determining the final parameters for the neural network, a Receiver Operating Characteristic (ROC) curve will be plotted for the training set. By definition, the ROC curve presents the tradeoff between the true positive rate versus the false positive rate (equivalently, sensitivity versus 1–specificity) for different thresholds of the classifier output. We utilize this characteristic to find the neural network output threshold that can improve the classification performance. It is noted that in the application of hypoglycemia detection, the sensitivity, which represents the rate of correctly classifying hypoglycemia episodes, is more important than the specificity. Therefore, based on the ROC curve, a criterion for output threshold is set at the point producing classification sensitivity of 80%. This procedure leads to a desired sensitivity and a relatively low but reasonable specificity. Details about the definition of sensitivity and specificity in our application can be found in [11]. The area under the ROC curve (*AuR*) is also estimated as a measure of classification performance. It is noted that the higher the *AuR* produces the better the classification.

III. Results

The actual BGL profiles from five T1DM patients which were collected during the study are shown in Fig. 2. The BGL threshold to distinguish between hypoglycemia and non-hypoglycemia is set at 3.3 mmol/l.

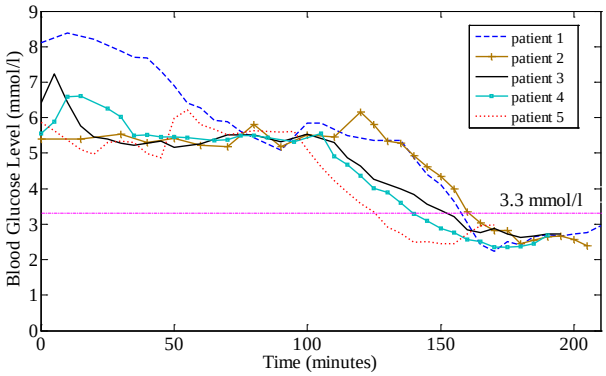

Fig. 2. Actual blood glucose level profiles in 5 T1DM children

Statistical analysis yields significant changes in the centroid theta frequency and centroid alpha frequency. Under hypoglycemia conditions, the decrease in centroid alpha frequency is the most significant and consistent feature ($p < 0.0001$ at all 4 channels). The results also indicate a slight increase in centroid theta frequency ($p < 0.05$). The data from 5 patients show no significant change in the power level of all 3 frequency bands, as well as in the centroid beta frequency. The study shows that there is no significant difference in responses of EEG channels in the same brain areas (i.e. O1 and O2 in the occipital lobe; C3 and C4 in the central lobe). Therefore, in this paper, we use EEG data from only two channels C3 and O2, which are in two different sides and areas of the brain. As a result, centroid theta frequency and centroid alpha frequency from two channels C3 and O2 are employed as inputs for classification.

A neural network is developed with 4 input nodes (2 features x 2 channels), S hidden nodes and 1 output nodes. S is varied from 1 to 16 to select the one that give the best performance. As a result, it's recognized that for our application with 4 input nodes and 1 output nodes, $S = 9$ yields the best classification results. The following results are corresponding with a neural network of 9 hidden nodes.

For GA training, the overall data set is separated into a training set and a testing set. The training set is formed from the data of 3 patients, named patient A, B and C. This set consists of 284 data points which includes 112 points of hypoglycemia. The testing set is formed from data of 2 patients, named patient D and E. This set includes 144 data points which include 76 points of hypoglycemia. For LM training, in order to implement the early stopping, the above training set is subdivided into a LM-training set and a validation set with a ratio of 3:1.

To implementing GA-based neural network training, the following parameters and operators are implemented:

- Selection method: Normalized Geometric Ranking
- Crossover operator: Blend-α crossover
- Mutation operator: Non-uniform mutation
- Chromosome length: 55
- Maximum number of generations: 2000
- Population size: 50
- Parameter range: $-3 \le w_{ij}\ v_i\ b_{1i}\ b_2 \le 3$

The classification results are presented in Table I and II. For comparison, three different training strategies are implemented, including LM, GA and combination of GA and LM (GA+LM). Based on the ROC curve of each case, the corresponding *AuR* and output threshold to distinguish between hypoglycemia and non-hypoglycemia are selected. The reported results are the mean and best performance of 20 run times. The best results in Table I showed that by combining GA and LM algorithms to train neural network, classification results are enhanced markedly up to 75% sensitivity and 60% specificity. The LM algorithm produces comparable best classification results with the GA+LM algorithm, and better than the GA algorithm. However, the mean results in Table I demonstrate the positives as well as drawbacks of each algorithm. The LM algorithm produces the worst mean classification performance due to the

inherence of trapping in local optimal. The GA algorithm produces consistent results in all running times which can be explained by the ability of directing the optimization process to the region of optimal solution. Taking this advantage of the GA algorithm, combining with the fine tuning capability of the LM algorithm, the GA+LM algorithm produces better classification results compared to the other two algorithms.

TABLE I. MEAN CLASSIFICATION RESULTS

Training method	AUR	Training		Testing	
		Sen	**Spe**	**Sen**	**Spe**
LM	0.73	80%	52%	67%	42%
GA	0.70	80%	37%	83%	40%
GA+BP	0.79	80%	57%	74%	52%

Sen: sensitivity ; Spe: specificity

TABLE II. BEST CLASSIFICATION RESULTS

Training method	AUR	Training		Testing	
		Sen	**Spe**	**Sen**	**Spe**
LM	0.75	80%	53%	71%	54%
GA	0.74	80%	42%	73%	41%
GA+BP	0.82	80%	61%	75%	60%

Sen: sensitivity ; Spe: specificity

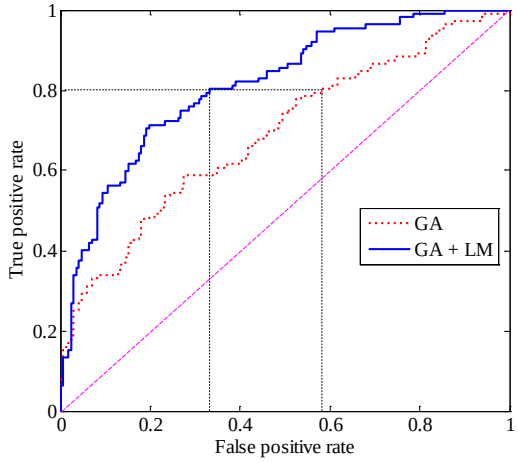

Fig. 3. ROC curve

IV. CONCLUSIONS

In this paper, a strategy of combining GA and LM to train neural network is demonstrated with the aim of improving the classification performance for the hypoglycemia detecting algorithm. Four EEG parameters from two non-invasive EEG channels C3 and O2 are used as inputs for a neural network classification unit. By utilizing the global search ability of GA and the local search ability of LM in training neural network, it is showed that classification results can be enhanced remarkably up to 80% sensitivity and 61% specificity on the training set, and 75% sensitivity and 60% specificity on the testing set. Based on achieved results, a more advanced algorithm of neural network will be investigated in future research to improve the study's overall performance.

One of the limitations of this current study is the shortage of data. The data set from five participated patients is sufficient to establish that hypoglycemia induces changes in EEG signals which can be detected by using neural network. However, in order to apply these results to develop a hypoglycemia detecting system that can perform in real clinical environments, other studies with more participants need to be implemented in future. Furthermore, it should be noted that this is a glucose clamp study, which is not spontaneous hypoglycemia. In the future work, a natural hypoglycemia study would be carried out to explore the difference in EEG responses to glucose-induced hypoglycemia and spontaneous hypoglycemia as well as to validate the performance of the developed algorithm.

With the proposed methodology, we are continuing to pursue our final purpose of developing the real-time system that can efficiently and continuously monitor patients' conditions and alert them and their carers when hypoglycemia is detected nocturnally during sleep.

REFERENCES

[1] D. R. Group, "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus," *New England Journal of Medicine,* vol. 329, pp. 977-986, 1993.

[2] U. Pedersen-Bjergaard, S. Pramming, S. R. Heller, T. M. Wallace, Å. K. Rasmussen, H. V. Jørgensen, D. R. Matthews, P. Hougaard, and B. Thorsteinsson, "Severe hypoglycaemia in 1076 adult patients with type 1 diabetes: influence of risk markers and selection," *Diabetes/Metabolism Research and Reviews,* vol. 20, pp. 479-486, 2004.

[3] D. R. Group, "Epidemiology of severe hypoglycemia in the diabetes control and complications trial," *The American journal of medicine,* vol. 90, pp. 450-459, 1991.

[4] S. Pramming, B. Thorsteinsson, B. Stigsby, and C. Binder, "Glycaemic threshold for changes in electroencephalograms during hypoglycaemia in patients with insulin dependent diabetes," *British Medical Journal (Clinical research ed.),* vol. 296, pp. 665-667, March 5, 1988.

[5] K. Howorka, G. Heger, A. Schabmann, P. Anderer, G. Tribl, and J. Zeitlhofer, "Severe hypoglycaemia unawareness is associated with an early decrease in vigilance during hypoglycaemia," *Psychoneuroendocrinology,* vol. 21, pp. 295-312, 1996.

[6] L. Hyllienmark, J. Maltez, A. Dandenell, J. Ludvigsson, and T. Brismar, "EEG abnormalities with and without relation to severe hypoglycaemia in adolescents with type 1 diabetes," *Diabetologia,* vol. 48, pp. 412-419, 2005.

[7] F. Laione and J. Marques, "Methodology for hypoglycaemia detection based on the processing, analysis and classification of the electroencephalogram," *Medical and Biological Engineering and Computing,* vol. 43, pp. 501-507, 2005.

[8] C. B. Juhl, K. Højlund, R. Elsborg, M. K. Poulsen, P. E. Selmar, J. J. Holst, C. Christiansen, and H. Beck-Nielsen, "Automated detection of hypoglycemia-induced EEG changes recorded by subcutaneous electrodes in subjects with type 1 diabetes--The brain as a biosensor," *Diabetes Research and Clinical Practice,* vol. 88, pp. 22-28, 2010.

[9] L. B. Nguyen, S. S. H. Ling, T. W. Jones, and H. T. Nguyen, "Identification of hypoglycemic states for patients with T1DM using various parameters derived from EEG signals," in *Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE*, pp. 2760-2763, 2011.

[10] H. T. Nguyen, "Intelligent technologies for real-time biomedical engineering applications," *International Journal of Automation and Control,* vol. 2, pp. 274-285, 2008.

[11] L. B. Nguyen, A. V. Nguyen, S. H. Ling, and H. T. Nguyen, "An adaptive strategy of classification for detecting hypoglycemia using only two EEG channels," in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 3515-3518, 2012.