

Research Article

Intelligent Noise Removal from EMG Signal Using Focused Time-Lagged Recurrent Neural Network

S. N. Kale and S. V. Dudul

Department of Applied Electronics, Sant Gadge Baba Amravati University, Amravati, 444602 Maharashtra, India

Correspondence should be addressed to S. N. Kale, sujatankale@rediffmail.com

Received 5 November 2008; Revised 6 February 2009; Accepted 30 March 2009

Recommended by Zhigang Zeng

Electromyography (EMG) signals can be used for clinical/biomedical application and modern human computer interaction. EMG signals acquire noise while traveling through tissue, inherent noise in electronics equipment, ambient noise, and so forth. ANN approach is studied for reduction of noise in EMG signal. In this paper, it is shown that Focused Time-Lagged Recurrent Neural Network (FTLRNN) can elegantly solve to reduce the noise from EMG signal. After rigorous computer simulations, authors developed an optimal FTLRNN model, which removes the noise from the EMG signal. Results show that the proposed optimal FTLRNN model has an MSE (Mean Square Error) as low as 0.000067 and 0.000048, correlation coefficient as high as 0.99950 and 0.99939 for noise signal and EMG signal, respectively, when validated on the test dataset. It is also noticed that the output of the estimated FTLRNN model closely follows the real one. This network is indeed robust as EMG signal tolerates the noise variance from 0.1 to 0.4 for uniform noise and 0.30 for Gaussian noise. It is clear that the training of the network is independent of specific partitioning of dataset. It is seen that the performance of the proposed FTLRNN model clearly outperforms the best Multilayer perceptron (MLP) and Radial Basis Function NN (RBF) models. The simple NN model such as the FTLRNN with single-hidden layer can be employed to remove noise from EMG signal.

Copyright © 2009 S. N. Kale and S. V. Dudul. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Biomedical signal means a collective electrical signal acquired from any organ that represents a physical variable of interest. This signal is normally a function of time and is describable in terms of its amplitude, frequency, and phase. The EMG signal is a biomedical signal that measures electrical currents generated in muscles during its contraction representing neuromuscular activities. The nervous system always controls the muscle activity (contraction/relaxation). Hence, the EMG signal is a complicated signal, which is controlled by the nervous system and is dependent on the anatomical and physiological properties of muscles. EMG signal acquires noise while traveling through different tissues. Moreover, the EMG detector, particularly if it is at the surface of the skin, collects signals from different motor units at a time which may generate interaction of different signals. Detection of EMG signals with powerful and advance methodologies is becoming a very important requirement in

biomedical engineering. The main reason for the interest in EMG signal analysis is in clinical diagnosis and biomedical applications. So far, research and extensive efforts have been made in the area, developing better algorithms, upgrading existing methodologies, and improving detection techniques to reduce noise and to acquire accurate EMG signals [1]. Noise removal from noisy EMG signal is a filtering problem. Here the Neural Network model is trained to separate known noise from EMG signal.

Literature survey [2–5] shows that Neural Networks (NNs) have been efficiently used for nonlinear multivariable function approximation. However, there is still enough scope to choose an appropriate NN model so that the performance measures are optimized to approach zero and unity for mean square error (MSE) and correlation coefficient (r), respectively. In function approximation, the goal is to find the parameters of the best linear approximation to the input and the desired response pairs. In nonlinear system identification, conventional techniques such as least square

approach, partial least square regression, principal components regression, ordinary least square regression, regression tree, Levenberg Marquardt algorithm, and multivariate adaptive regression splines algorithm generally do not work reasonably if the underlying problem is overly complex [6–8]. Therefore NN approach is worth considering for solving system identification problem [9]. A typical problem of noise removal in EMG signal is considered in this paper. This benchmark data for noise removal in EMG signal is taken from the companion CD of a book on neural network [10]. Data contains an electromyographic (EMG) signal and the interference (60 Hz) noise picked from the power supply. The two files are, respectively, “EMG with noise” and “noise” only. The goal is to obtain back the EMG using adaptive filtering techniques. The training file is used to train a neural network for noise removal from EMG signal.

Optimal Focused Time Lag Recurrent Neural Network (FTLRNN) is developed to remove noise effectively from EMG signal. Other classes of NN configuration such as Multilayer Perceptron Neural Network (MLP NN) and Radial Basis Function (RBF) have also been compared for such noise removal problem.

This paper deals with intelligent removal of noise from the EMG signal using FTLRNN-based model.

2. EMG and Sources of Noise

EMG stands for electromyography. It is the study of signals. EMG is sometimes referred to as myoelectric activity. Muscle tissue conducts electrical potentials similar to the way nerves do, and the name given to these electrical signals is the muscle action potential. Surface EMG is a method of recording the information present in these muscle action potentials. When detecting and recording the EMG signal, there are two main issues of concern that influence the fidelity of the signal. The first is the signal-to-noise ratio. That is, the ratio of the energy in the EMG signals to the energy in the noise signal. In general, noise is defined as electrical signals that are not part of the desired EMG signal. The other issue is the distortion of the signal, meaning that the relative contribution of any frequency component in the EMG signal should not be altered. There are many applications for the use of EMG. EMG is used clinically for the diagnosis of neurological and neuromuscular problems. It is used diagnostically by gait laboratories and by clinicians trained in the use of biofeedback or ergonomic assessment. EMG is also used in many types of research laboratories, including those involved in biomechanics, motor control, neuromuscular physiology, movement disorders, postural control, and physical therapy.

Electrical Noise and Factors Affecting EMG Signal. The amplitude range of EMG signal is 0–10 mV (+5 to –5) prior to amplification. EMG signals acquire noise while traveling through different tissues. It is important to understand the characteristics of the electrical noise. Electrical noise, which will affect EMG signals, can be categorized into the following types.

(1) *Inherent Noise in Electronics Equipment.* All electronics equipments generate noise. This noise cannot be eliminated; using high-quality electronic components can only reduce it.

(2) *Ambient Noise.* Electromagnetic radiation is the source of this kind of noise. The surfaces of our bodies are constantly inundated with electric-magnetic radiation, and it is virtually impossible to avoid exposure to it on the surface of earth. The ambient noise may have amplitude that is one to three orders of magnitude greater than the EMG signal.

(3) *Motion Artifact.* Motion artifact causes irregularities in the data. There are two main sources for motion artifact: (1) electrode interface and (2) electrode cable. Motion artifact can be reduced by proper design of the electronics circuitry and set-up.

(4) *Inherent Instability of Signal.* The amplitude of EMG is random in nature. EMG signal is affected by the firing rate of the motor units, which, in most conditions, fire in the frequency region of 0 to 20 Hz. This kind of noise is considered as unwanted, and the removal of the noise is important.

3. Performance Measures

Assessment of the performance of various neural networks is done by visual inspection of EMG and noise signals from the graph as well as from the optimal values of Mean Square Error (MSE), and r (Correlation coefficient).

Mean Square Error (MSE). The formula for the mean square error is

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N (d_{ij} - y_{ij})^2}{NP}, \quad (1)$$

where P = number of output processing elements, N = number of exemplars in the dataset, y_{ij} = network output for exemplar i at processing element j , and d_{ij} = desired output for exemplar i at processing element j .

Correlation Coefficient (r). By definition, the correlation coefficient between a network output x and a desired output d is

$$r = \frac{\sum_i (x_i - \bar{x})(d_i - \bar{d})/N}{\sqrt{\sum_i (d_i - \bar{d})^2/N} \sqrt{\sum_i (x_i - \bar{x})^2/N}}, \quad (2)$$

where $\bar{x} = (1/N) \sum_{i=1}^N x_i$ and $\bar{d} = (1/N) \sum_{i=1}^N d_i$.

The correlation coefficient is confined to the range $[-1, 1]$. When $r = 1$, there is a perfect positive linear correlation between x and d , that is, they covary, which means that they vary by the same amount.

4. Computer Simulation

Here a dataset is chosen that can be used in removal of noise from EMG signal. There are 2000 training patterns. Training

TABLE 1: EMG signal datasets used for NN-based models. Total number of samples = 2000.

(a) Datasets based on forward and reverse tagging and % variation

S.N.	Dataset	Tagging	% variation (Train-CV-test)
1	Dataset1	Forward (training, cross validation, and testing)	10-15-75%
2	Dataset2		20-15-65%
3	Dataset3		30-15-55%
4	Dataset4		40-15-45%
5	Dataset5		50-15-35%
6	Dataset6		60-15-25%
7	Dataset7		70-15-15%
8	Dataset8		80-15-05%
9	Dataset9	Reverse (testing, cross validation, and training)	10-15-75%
10	Dataset10		20-15-65%
11	Dataset11		30-15-55%
12	Dataset12		40-15-45%
13	Dataset13		50-15-35%
14	Dataset14		60-15-25%
15	Dataset15		70-15-15%
16	Dataset16		80-15-05%

(b) Datasets based on multifold differential learning. Group I: (1–500 samples), Group II: (501–1000 samples), Group III: (1001–1500 samples), Group IV: (1501–2000 samples)

Dataset	Train on group	Test on group	Dataset	Train on group	Test on group
Dataset1	I	II	Dataset18	I + III	II + IV
Dataset2	I	III	Dataset19	I + IV	II
Dataset3	I	IV	Dataset20	I + IV	III
Dataset4	II	III	Dataset21	I + IV	II + III
Dataset5	II	IV	Dataset22	II + III	I
Dataset6	II	I	Dataset23	II + III	IV
Dataset7	III	IV	Dataset24	II + III	I + IV
Dataset8	III	I	Dataset25	II + IV	I
Dataset9	III	II	Dataset26	II + IV	III
Dataset10	IV	I	Dataset27	II + IV	I + III
Dataset11	IV	II	Dataset28	III + IV	I
Dataset12	IV	III	Dataset29	III + IV	II
Dataset13	I + II	III	Dataset30	III + IV	I + II
Dataset14	I + II	IV	Dataset31	I + II + III	IV
Dataset15	I + II	III + IV	Dataset32	II + III + IV	I
Dataset16	I + III	II	Dataset33	III + IV + I	II
Dataset17	I + III	IV	Dataset34	IV + I + II	III

of the neural network should be independent of dataset. Therefore different permutations and combinations of the dataset producing many independent datasets are used for training and testing of neural networks.

Table 1 depicts the various datasets on which the neural networks are trained. Once the data is randomized, the total samples are divided into three parts, namely, training, cross validation, and testing samples. If the samples are divided in the sequence of training, cross validation, and testing, it is a forward tagging. On the other hand the

sequence of testing, cross validation, and then training is termed as reverse tagging. Percentage of training and testing samples are varied, and cross validation samples are kept constant as shown in Table 1(a). Forward tagging and reverse tagging of dataset give total 16 different datasets to assess the performance of an estimated network model. This dataset is also tested for multifold differential learning. Multifold differential learning of neural network is carried out on the dataset, that is, the total samples are divided into four groups each containing 500 samples as given in

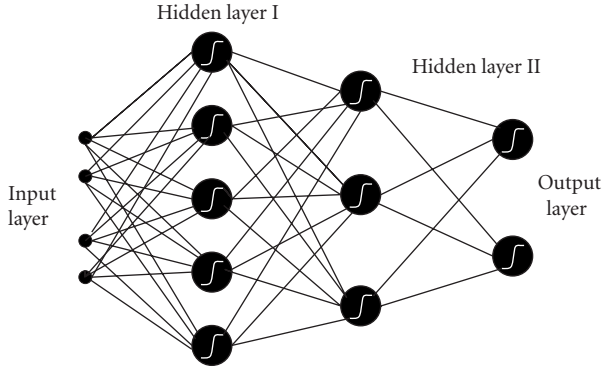


FIGURE 1: A simple multilayer perceptron.

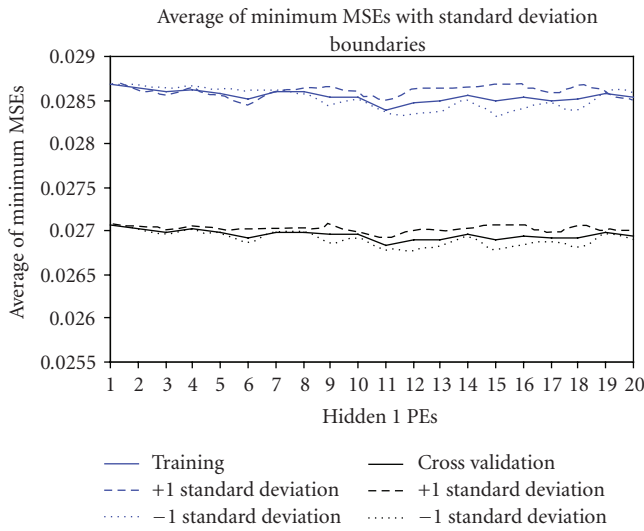


FIGURE 2: Variation of average of minimum MSE for 5 runs versus number of PEs in the first hidden layer.

Table 1(b). Sample numbers of each group is mentioned in Table 1(b). All possible combinations are used to train the neural network and assess the performance by testing. There are total 34 datasets formed for differential learning as described in Table 1(b). To assess the performance of neural network skeptically, total 50 different datasets are used. This is necessary because the estimated NN model should consistently work on the different datasets. This also ensures that the proposed NN model has truly learned meaningful information from the dataset and is free from biases.

Evaluation of NN is done by a standard method in statistics called *independent validation* where the available data are divided into a training set, a cross validation (CV) set, and a test set. The entire dataset is usually randomized first. The training data is used to update the weights in the network. The test data is then used to assess how well the network has generalized. The learning and generalization ability of the estimated NN model is assessed on the basis of performance measures such as MSE, correlation coefficient r , and visual inspection of desired and actual graphs of EMG signal.

The network has been trained at least 5 times starting from different random initial weights so as to avoid local

TABLE 2: Variable parameters of MLP NN model.

S.N.	Parameter	Typical range	Optimal parameter
1	Hidden layer	1 to 4	2
2	PE	H I-1 to 20 H II-1 to 10 H III-10 H IV-10	Hidden layer I- 15 Hidden layer II- 10
3	Learning rule	Momentum (Mom), Conjugate gradient (CG), Levenberg Marquardt (LM), Quick propagation (QP), Step, Delta bar delta	<i>Momentum</i>
4	Transfer function in output layer	Linear, Lineartanh, Tanh	<i>Tanh</i>

TABLE 3: Number of hidden layer and r .

S.N.	No. of hidden layer	Correlation coefficient	
		Noise	EMG
1	1	0.76966	0.56711
2	2	0.77212	0.57113
3	3	0.77107	0.56936
4	4	0.76574	0.56861

minima. Neurodimension NeuroSolutions (version 5) is specifically used for obtaining results. System with 512 MB RAM, 40 GB hard disk, 2 MB cache, and 1.6 GHz clock is used to carry out this simulation.

Various neural networks are used to compare the performance, and FTLRNN is the best in removal of noise from EMG signal.

4.1. MLP NN. MLP-based NN model is used in this study because it has solid theoretical foundation [11]. MLPs are feedforward neural networks trained with the standard back-propagation algorithm [12]. They are supervised networks, so they require a desired response to be trained. Figure 1 shows the architecture of MLP NN.

An exhaustive and careful experimental study has been carried out to determine the optimal configuration of MLP NN model. All possible variations such as number of hidden layers, number of PEs (processing elements) in each hidden layer, different transfer functions in the output layer, and different supervised learning rules are investigated in simulation.

Table 2 shows various parameters of the MLP NN model which are varied for obtaining optimal parameters.

Supervised learning epochs = 1000, error threshold = 0.01, transfer function in hidden layer = tanh, number of PEs in input layer = 1, and number of PEs in output layer = 2.

The number of hidden layers is varied from 1 to 4, and performance measures of the MLP NN model are found better for two hidden layers as shown in Table 3. With

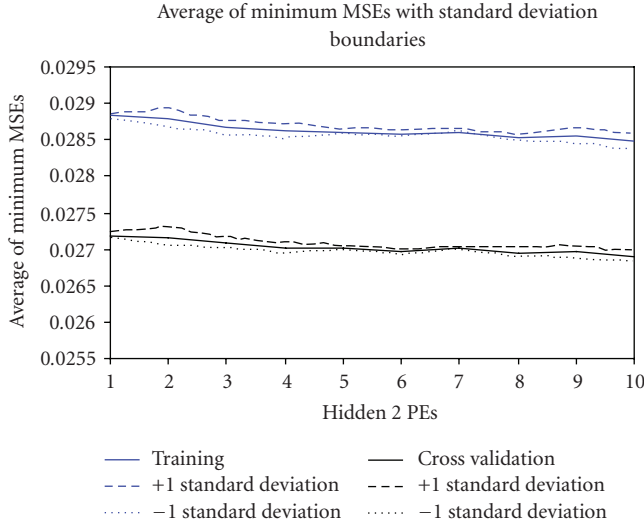


FIGURE 3: Variation of average of minimum MSE for 5 runs versus number of PEs in the second hidden layer.

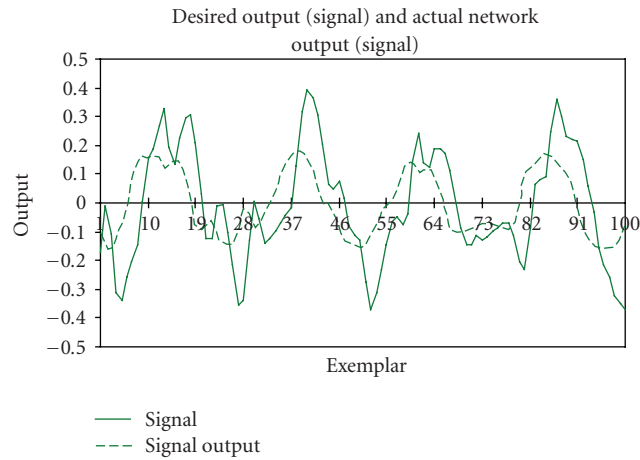


FIGURE 4: Modeling capability of MLP NN on testing dataset for EMG signal.

increase in number of hidden layers, the performance of the network has not improved significantly.

It is found from Figures 2 and 3 that the optimal performance of the model is obtained for 15 neurons in the first hidden layer and 10 neurons in the second hidden layer with regard to MSE minimum, r -correlation coefficient. Figures 2 and 3 portray average MSE with respect to the number of PEs in the first and second hidden layers, respectively.

Figures 4 and 5 depict modeling capability of MLP NN on test dataset which portrays desired output and actual output of the MLP NN on test dataset. It is seen that actual outputs of EMG signal and noise signal do not follow the desired output closely. There has been a lot of deviations between the output of the NN and the desired output.

For the datasets MLP NN model is trained for five times. The performance measures such as MSE and r on training dataset and testing dataset are obtained. Optimal

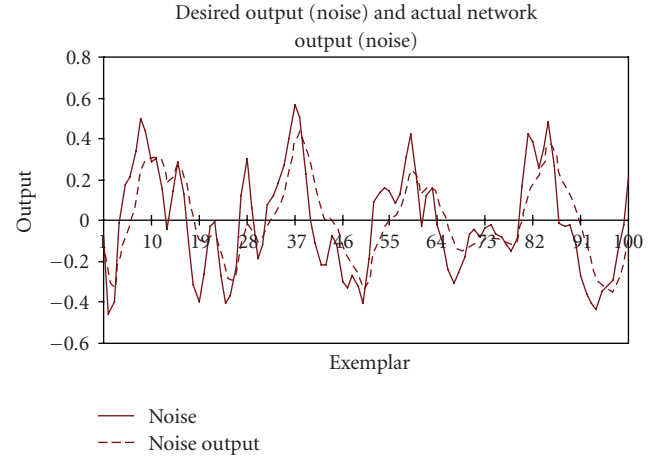


FIGURE 5: Modeling capability of MLP NN on testing dataset for EMG noise signal.

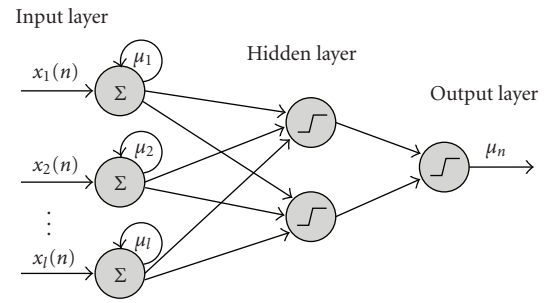


FIGURE 6: MLP with an input layer of context PEs (focused architecture).

TABLE 4: Number of hidden layers and r .

S.N.	No. of hidden layer	Correlation coefficient	
		Noise	EMG
1	1	0.99533	0.99471
2	2	0.98993	0.99302

performance is obtained when 80% of the entire dataset is used for training, 15% for cross validation, and 5% for testing. The *correlation coefficient* on test dataset is found as high as **0.78113** and $MSE = 0.02501$ for EMG signal and for noise signal $r = 0.5843$ and $MSE = 0.02485$.

4.2. Focused Time Lag Recurrent Neural Network (FTL-RNN). Time-lagged recurrent networks (TLRNs) are MLPs extended with short-term memory structures. Most real-world data contains information in its time structure, that is, how the data changes with time. TLRNs are the state of the art in nonlinear time series prediction, system identification, and temporal pattern classification.

Recurrent networks are neural networks with one or more feedback loops. The TDNN memory structure is simply a cascade of ideal delays (a delay of one sample). The gamma memory is a cascade of leaky integrators. The Laguerre memory is slightly more sophisticated than the

TABLE 5: Variable parameters of FTLRNN model.

S.N.	Parameter	Typical range	Optimal value
1	Learning rule	Momentum, Conjugate Gradient, Levenberg Marquardt (LM), Quick propagation (qp), step, delta bar delta (dbd)	<i>Momentum</i>
2	Transfer function in output layer	Linear, lineartanh, Tanh	<i>Linear</i>
3	Memory structure	TDNNAxon, Gamma, Laguarre	<i>Laguarre</i>
4	Depth in samples	2 to 12	4
5	Length of trajectory	10 to 60	50

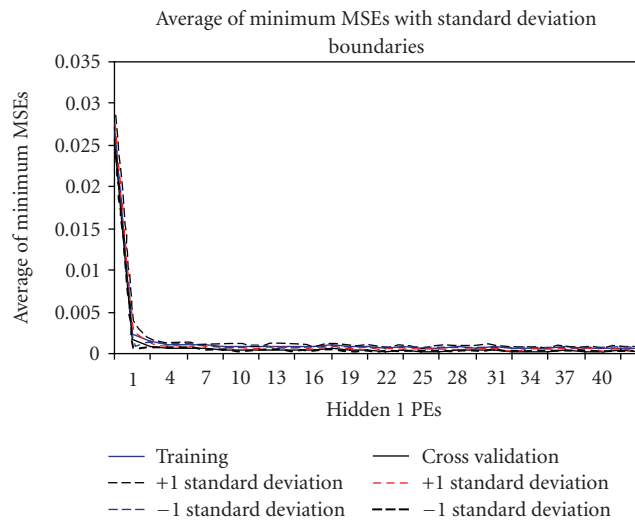


FIGURE 7: Variation of average of minimum MSE for 5 runs versus number of PEs in the first hidden layer.

gamma memory in that it orthogonalizes the memory space. This is useful when working with large memory kernels [10].

The input PEs of an MLP are replaced with a tap delay line. It is called the focused time delay neural network (TDNN). The topology is called focused because the memory is only at the input layer [13].

The delay line of the focused TDNN stores the past samples of the input. The combination of the tap delay line and the weights that connect the taps to the PEs of the first hidden layer is simply linear combiners followed by a static nonlinearity. The first layer of the focused TDNN is therefore a filtering layer, with as many adaptive filters as PEs in the first hidden layer.

The focused TDNN topology has been successfully used in nonlinear system identification, time series prediction, and temporal pattern recognition. Figure 6 shows architecture of FTLRNN. The focused topology of Figure 6 is a recurrent neural network and the recurrency is local to the PE. One of the advantages of locally recurrent neural networks is that the stability of the system can be judged by constraining the value of the local feedback parameters so

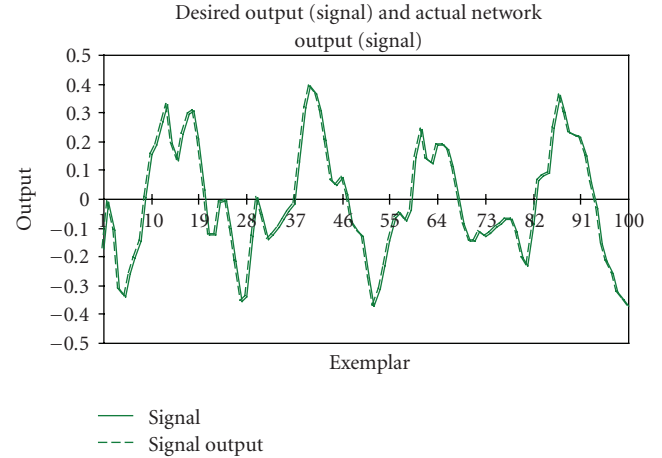


FIGURE 8: Modeling capability of FTLRNN on testing dataset for EMG signal.

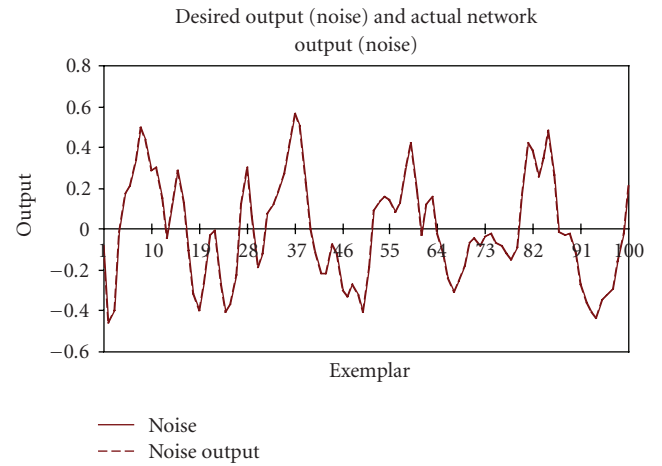


FIGURE 9: Modeling capability of FTLRNN on testing dataset for EMG noise signal.

that the local PE is stable. If local stability is enforced, the global system will be stable.

A thorough experimental study has been carried out to determine optimal parameters of FTLRNN model. Here the number of hidden layers is varied from 1 to 2, and performance measures of the FTLRNN model are found better for single hidden layer as shown in Table 4. With increase in the number of hidden layers, the performance of the network has not improved significantly.

Figure 7 portrays average MSE with respect to the number of PEs in the first hidden layer. 27 neurons are selected for optimal performance.

Table 5 shows various parameters of the FTLRNN model which are varied for obtaining optimal parameters. For momentum learning rule, the results are optimum. The *Momentum* provides the gradient descent with some inertia, so that it tends to move along a direction, that is, the average estimate for down. The amount of inertia (i.e., how much of the past to average over) is dictated by the momentum parameter, ρ . The higher the momentum is, the more it

TABLE 6: Selection of memory structure for optimal performance of FTLRNN model.

S.N.	Memory structure	Mean square error		Correlation coefficient	
		Noise	EMG	Noise	EMG
1	TDNNAxon	0.000455	0.000308	0.99719	0.996754
2	Gamma	0.000301	0.000281	0.99827	0.997204
3	Laguarre	0.0002624	0.000280	0.998427	0.997305

TABLE 7: Selection of depth of samples for optimal performance of FTLRNN model.

S.N.	Depth in samples	Mean square error		Correlation coefficient	
		Noise	EMG	Noise	EMG
1	2	0.02982	0.02982	0.7709	0.5628
2	4	0.00024	0.00024	0.9985	0.9975
3	6	0.00037	0.00031	0.9977	0.9969
4	8	0.00031	0.00036	0.99806	0.9962
5	12	0.00047	0.00028	0.9972	0.9971

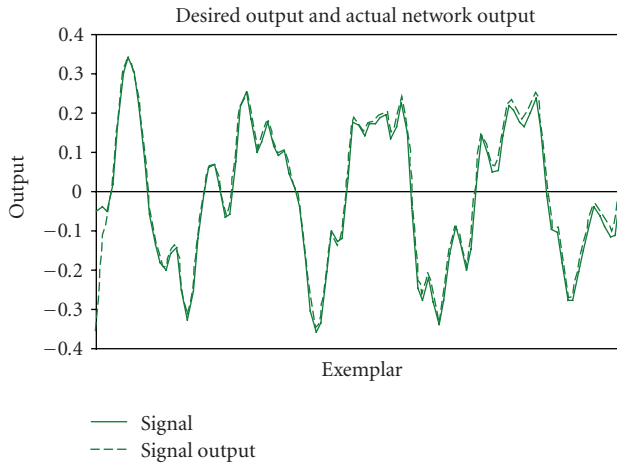


FIGURE 10: Modeling capability of FTLRNN on training dataset.

TABLE 8: Selection of length of trajectory for optimal performance of FTLRNN model.

S.N.	Length of trajectory	Mean square error		Correlation coefficient	
		Noise	EMG	Noise	EMG
1	05	0.00023	0.000228	0.9986	0.9974
2	10	0.00024	0.000236	0.9985	0.9976
3	20	0.00024	0.00025	0.9985	0.9974
4	30	0.00024	0.00025	0.9985	0.9975
5	40	0.00023	0.00022	0.9985	0.9975
6	50	0.00023	0.00024	0.9985	0.9976
7	60	0.00024	0.00025	0.9984	0.9974

smoothes the gradient estimate and the less effect a single change in the gradient has on the weight change. Linear transfer function has optimal results.

Supervised learning epochs = 1000, error threshold = 0.01, transfer function in hidden layer = tanh, number of PEs

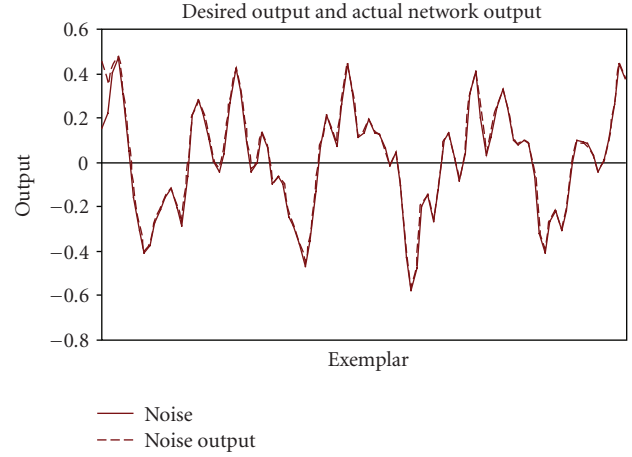


FIGURE 11: Modeling capability of FTLRNN on training dataset.

TABLE 9: Variable parameters of RBF NN model.

S.N.	Parameter	Typical range	Optimal parameter
1	Cluster centers	5 to 100	5
2	Unsupervised learning rule	Conscience-full, standard full	<i>Conscience-full</i>
3	Supervised learning rule	Momentum, Conjugate gradient, Levenberg Marquardt, Quick Propagation, Step, delta bar delta	<i>Levenberg Marquardt</i>
4	Metric	Euclidean, Dot Product, Box car	<i>Euclidean</i>
5	Transfer function in output layer	Linear, lineartanh, Tanh	<i>Linear</i>

in input layer = 1, number of PEs in hidden layer 1 = 27, and number of PEs in output layer = 2.

For the various datasets, FTLRNN model is trained for five times with different random initializations of connection weights. The performance measures like MSE and r on training dataset, cross validation dataset, and testing dataset are obtained. Optimal performance is obtained for training 80%, cross validation 15%, and 5% testing. The correlation coefficient on test dataset is found as 0.9984 and 0.9973 for noise signal and EMG, respectively. MSE for EMG signal and noise is obtained as 0.0002.

Table 6 depicts that the Laguarre memory structure leads to the optimal performance. Laguarre is a local recurrent

TABLE 10: Selection of unsupervised learning rule for RBF NN model.

S.N.	Unsupervised learning rule	Mean square error		Correlation coefficient	
		Noise	EMG	Noise	EMG
1	Conscience full	0.02484	0.024993	0.7827	0.5825
2	Standard full	0.02954	0.02950	0.77338	0.5698

TABLE 11: Selection of competitive learning metric for RBF NN model.

S.N.	Metric	Mean square error		Correlation coefficient	
		Noise	EMG	Noise	EMG
1	Euclidean	0.02484	0.024993	0.7827	0.5825
2	Dot product	0.02985	0.02967	0.7713	0.5664
3	Box car	0.02955	0.029496	0.77329	0.56994

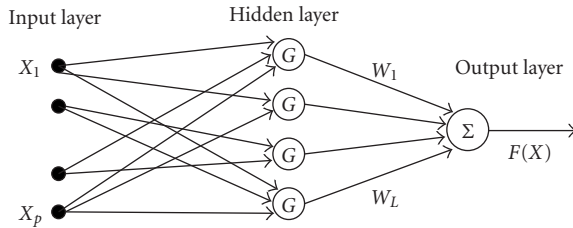


FIGURE 12: Radial basis function (RBF) network.

memory structure. It has internal feedback loops with an adaptable weight. The Laguerre memory is slightly more sophisticated than the gamma memory in that it orthogonalizes the memory space. This is useful when working with large memory kernels. The Laguerre memory is based on the Laguerre functions. The Laguerre functions are an orthogonal set of functions that are built from a low-pass filter followed by a cascade of all pass functions.

Depth of samples parameter (D) is used to compute the number of taps (T) contained within memory structure of the network. Optimal value of D is 4 as shown in Table 7.

The trajectory length corresponds to the samples setting within the dynamic controller. It specifies how many samples to read before backpropagation occurs. Table 8 shows the length of trajectory selected as 50 for optimal performance.

Figures 8 and 9 display modeling capability of FTLRNN, which shows desired output and actual output of the FTLRNN on test dataset for EMG and noise, respectively. It is seen that the output of the NN follows the desired output very closely.

Figures 10 and 11 display modeling capability of FTLRNN, which shows desired output and actual output of the FTLRNN on training dataset for signal and noise, respectively. It is seen that actual output follows the desired output closely.

4.3. Radial Basis Function (RBF). RBF was first introduced in the solution of the real multivariate interpolation problem [14, 15]. The construction of an RBF network, in its most basic form, involves three layers. The input layer is made up of source nodes (sensory units) that connect the network to

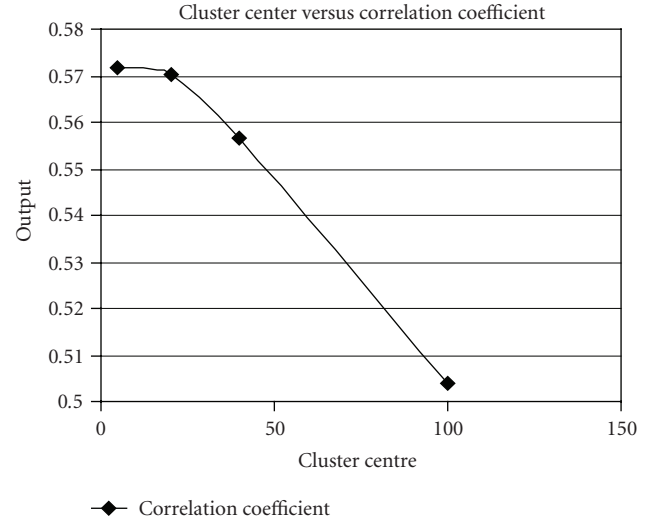


FIGURE 13: Variation of correlation coefficient for optimal number of cluster centers for RBF NN model.

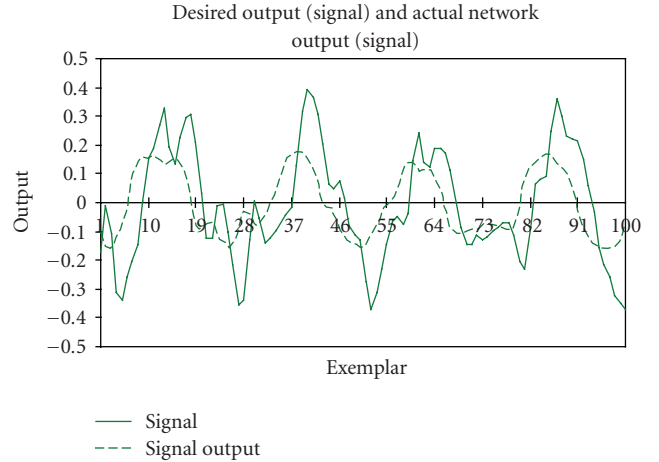


FIGURE 14: Modeling capability of RBF on test dataset of EMG signal.

its environment. The second layer, the only hidden layer in the network, applies a nonlinear transformation from the input space to the hidden space. The output layer is linear, supplying the response of the network to the activation pattern (signal) applied to the input layer [16]. Architecture of RBF NN model is shown in Figure 12.

A rigorous experimental study has been undertaken to determine optimal performance of RBF NN model. The variable parameters of RBF NN are listed in Table 9.

From Figure 13, it is seen that the optimal performance is obtained with 5 cluster centers.

TABLE 12: Selection of transfer function for various NN models on test dataset.

S.N.	NN	Signal	Transfer function					
			Tanh		Linear tanh		Linear	
			MSE	r	MSE	r	MSE	r
1	MLP	N	0.02501	0.78114	0.02557	0.77557	0.02521	0.7792
		EMG	0.02485	0.58433	0.0251	0.5779	0.02513	0.5781
2	FTLRNN	N	0.00027	0.99797	0.00015	0.99882	0.000067	0.99843
		EMG	0.00023	0.99695	0.00013	0.9981	0.000048	0.99731
3	RBF	N	0.2473	0.78385	0.02482	0.78297	0.02477	0.7839
		EMG	0.02481	0.5852	0.02481	0.5851	0.02480	0.5854

TABLE 13: Selection of optimal learning rule for various NN models on test dataset.

S.N.	NN	Signal	Learning rule					
			Momentum	Conjugate gradient	Levenberg Marquardt	Quick propagation	Step	Delta- bar- delta
1	MLP	Noise	0.7786	0.76672	0.77414	0.77388	0.77365	0.76857
		EMG	0.5840	0.57205	0.57182	0.56866	0.57026	0.57125
2	FTLRNN	Noise	0.99843	0.99617	0.99781	0.99806	0.99842	0.99792
		EMG	0.99731	0.99572	0.99684	0.99756	0.99695	0.99787
3	RBF	Noise	0.77314	0.7737	0.77382	0.77335	0.77335	0.7724
		EMG	0.57059	0.5712	0.57186	0.5718	0.57184	0.5711

TABLE 14: Regression performance of NN models on EMG noise signal.

S.N.	NN model	Performance measures					
		Training dataset		CV dataset		Test dataset	
		MSE	r	MSE	r	MSE	r
1	MLP	0.02688	0.77322	0.03070	0.77021	0.02501	0.78114
2	FTLRNN	0.000254	0.99809	0.00201	0.98701	0.000067	0.99950
3	RBF	0.02664	0.77544	0.03036	0.77309	0.024704	0.784138

TABLE 15: Regression performance of NN models on EMG signal.

S.N.	NN model	Performance measures					
		Training dataset		CV dataset		Test dataset	
		MSE	r	MSE	r	MSE	r
1	MLP	0.026537	0.57309	0.03041	0.56829	0.02485	0.58433
2	FTLRNN	0.00024	0.99692	0.00216	0.97584	0.000048	0.99939
3	RBF	0.02654	0.57122	0.03037	0.56915	0.024817	0.585088

TABLE 16: Comparison of all the NN architectures on test dataset.

NN model	Transfer function	Learning rule	Mean square error MSE	Correlation coefficient r	Epochs	Time elapsed/ epoch/exemplar in microseconds	% error
MLP (1-15-10-2)	Tanh	Momentum	0.02501 (noise) 0.02482 (EMG)	0.78114 (noise) 0.58433 (EMG)	1000	19.16	253
FTLRNN (1-27-2)	Linear	Momentum	0.000067 (noise) 0.000048 (EMG)	0.99950 (noise) 0.99939 (EMG)	1000	14	10
RBF (CC-05)	Linear	LM	0.02470 (noise) 0.02482 (EMG)	0.78414 (noise) 0.58509 (EMG)	1000	8.3	293

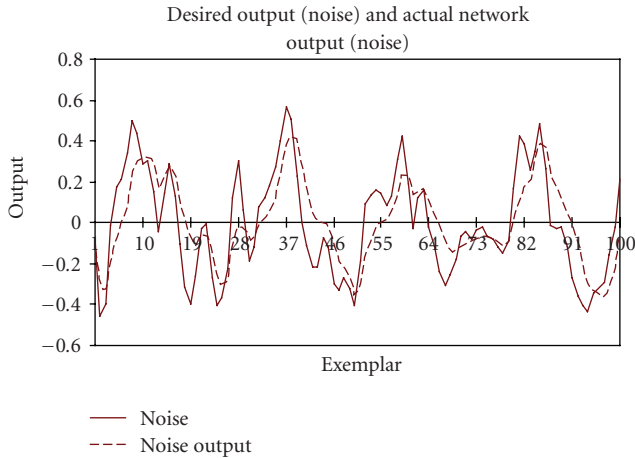


FIGURE 15: Modeling capability of RBF on test dataset of EMG noise signal.

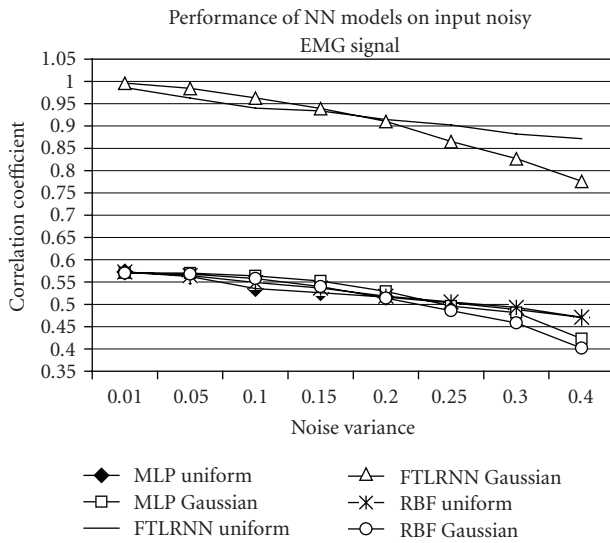


FIGURE 16: Effect of noise on EMG signal.

Tables 10 and 11 depicts the optimal performance of RBF NN. Conscience full-unsupervised learning rule and Euclidean competitive learning metric are selected for optimal performance.

Figures 14 and 15 give modeling capability of RBFNN, which show desired output and actual output of the RBF NN on test dataset for EMG signal. It is seen that actual output follows the desired output distantly.

5. Results and Comparison

Table 12 depicts the performance parameters for variation in learning rules for MLP NN, FTLRNN, and RBFNN on test dataset. From Table 12, it is observed that focused time-lagged recurrent neural network gives optimal performance for linear transfer function.

Table 13 depicts the selection of learning rule for optimal performance of each NN. In FTLRNN momentum learning rule is selected for the best performance.

Tables 14 and 15 display the regression performance of NN models. It shows performance parameters, MSE and r on training, cross validation, and test dataset for MLP NN, FTLRNN, and RBF NN for noise and EMG signal. From the observation, it clear that for FTLRNN model the lowest MSE and the highest correlation coefficient are obtained. FTLRNN is the best neural network to remove noise from EMG signal.

Table 16 displays the comparison of the MLP NN, FTLRNN, and RBF NN. For all the three NNs, the number of epochs is kept 1000. MSE for FTLRNN model is 0.0027 times less than that of MLP and RBF NN models. Correlation coefficient for FTLRNN is 1.71 times higher than that of MLP and RBF NN models. Percentage error for FTLRNN is minimum. It is 0.04 times and 0.034 times smaller than MLP and RBF NN, respectively. Time elapsed per epoch per exemplar for FTLRNN is 0.73 times and 1.67 times to that of MLP and RBF NN, respectively. As compared to RBF NN, FTLRNN model requires more time for training but from MSE and r , and by visual inspection of modeling characteristics, the FTLRNN model is definitely superior to other two NNs.

Effect of Noise on EMG Signal. The estimated MLP NN, FTLRNN, and RBF NN are checked for their robustness by adding uniform and Gaussian noise in input as well as in output of NNs. Figure 16 portrays the performance of NNs with uniform and Gaussian noise. Noise variance is varied from 0.01 to 0.4. In FTLRNN uniform noise tolerance for EMG signal 0.4-noise variance is obtained whereas when Gaussian noise is introduced, the noise variance 0.3 is detected. In MLP NN and RBF NN as noise variance is increased, the performance parameters are reduced to very low values.

Learning Ability of FTLRNN on Different Data Partitions. The learning of NN models for independent of datasets is tested. MLP, FTLRNN, and RBF NN models are trained on various datasets as shown in Table 1(a) (forward tagging and reverse tagging). Figure 17 displays the performance of these NN models for filtered EMG signal. Performance of FTLRNN-based model is found to be almost the same for all the datasets as compared to MLP and RBF NN models.

Multifold Differential Learning. The total samples are divided into four groups each containing 500 samples as described in Table 1(b). Performance of FTLRNN, MLP NN, and RBF NN models is displayed in Figure 18. It is observed that the performance FTLRNN-based model is consistent. It is also observed that correlation coefficient is the highest for FTLRNN for all datasets.

6. Conclusion

EMG signal carries valuable information regarding the nerve system. Noise removal in EMG signal using ANN is studied in this paper. Authors demonstrate that FTLRNN-based filter elegantly removes noise from the EMG signal. Compact

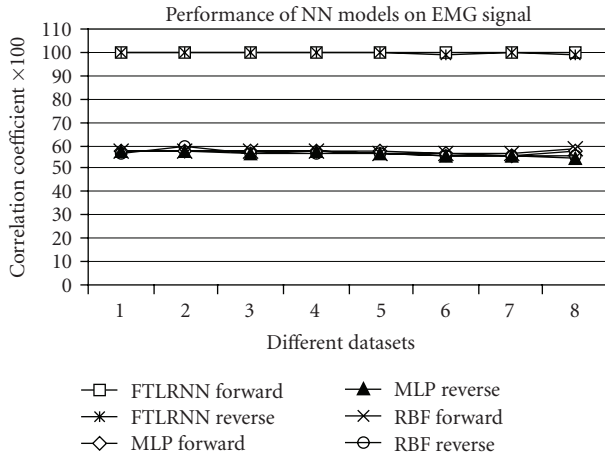


FIGURE 17: Correlation coefficient with various datasets.

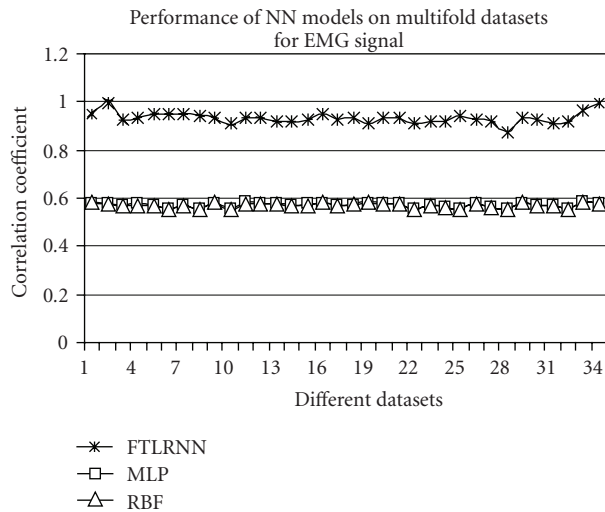


FIGURE 18: Performance of NNs on differential learning.

FTLRNN with only one hidden layer having architecture (1-27-2) is able to remove noise with reasonable accuracy. When the performance of MLP and RBF neural network-based models is carefully examined for dataset, FTLRNN based model has clearly outperformed its MLP NN and RBF NN counterparts with respect to the performance measures such as MSE and r as well as the visual inspection of graphs of actual and desired output of filtered EMG signal. For FTLRNN-based filter correlation coefficient is obtained as high as 0.99939, and MSE is found to be as low as 0.000048 for filtered EMG signal. Also for noise signal the correlation coefficient and MSE are optimally found as 0.99950 and 0.000067, respectively. Moreover, the actual output of the estimated FTLRNN model follows the desired output more closely than that of other NN models. In case of learning ability of FTLRNN-based model, the performance parameters are found consistent, and hence learning is almost independent of specific partitioning of the dataset. It is also seen that the time elapsed per epoch per exemplar required to train the network is considerably low for FTLRNN-based model. The least percentage error

equal to 10% for FTLRNN on test dataset is obtained. It is also observed that when uniform and Gaussian noise is introduced in EMG signal, the network sustains reasonable level of noise. For uniform noise, 100% tolerance is observed, and for Gaussian noise, it is 75%. This confirms the noise immunity of the proposed FTLRNN-based model. The estimated FTLRNN is a robust network developed to detect EMG signal from noisy EMG signal.

Proposed FTLRNN-based model with Laguarre memory is able to filter noise from a typical EMG signal contaminated by noise.

References

- [1] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," *Biological Procedures Online*, vol. 8, no. 1, pp. 11–35, 2006.
- [2] S.-I. Wu and H. Zheng, "Stock index forecasting using recurrent neural networks," in *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications (AIA '06)*, Innsbruck, Austria, February 2006.
- [3] R. D. de Veaux, J. Schumi, J. Schweinsberg, and L. H. Ungar, "Prediction intervals for neural networks via nonlinear regression," *Technometrics*, vol. 40, no. 4, pp. 273–282, 1998.
- [4] R. Barman, B. Prasad Kumar, P. C. Pandey, and S. K. Dube, "Tsunami travel time prediction using neural networks," *Geophysical Research Letters*, vol. 33, no. 16, Article ID L16612, 6 pages, 2006.
- [5] S. M. Sadat-Hashemi, A. Kazemnejad, C. Lucas, and K. Badie, "Predicting the type of pregnancy using artificial neural networks and multinomial logistic regression: a comparison study," *Neural Computing & Applications*, vol. 14, no. 3, pp. 198–202, 2005.
- [6] H. M. Fredric and I. Kostanic, *Principles of Neurocomputing for Science & Engineering*, Tata McGraw-Hill, New Delhi, India, 2000.
- [7] P. T. Boggs, R. H. Byrd, and R. B. Schnabel, "A stable and efficient algorithm for nonlinear orthogonal distance regression," *SIAM Journal on Scientific Computing*, vol. 8, no. 6, pp. 1052–1078, 1987.
- [8] P. A. Lewis and J. G. Stevens, "Nonlinear modeling of time series using Multivariate Adaptive Regression Splines (MARS)," Tech. Rep. ADA222710, Defense Technical Information Center, Fort Belvoir, Va, USA, 1990.
- [9] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Pearson Education, Delhi, India, 1999.
- [10] J. C. Principe, N. Euliano, and W. C. Lefebvre, *Neural & Adaptive Systems: Fundamentals through Simulations*, John Wiley & Sons, New York, NY, USA, 2000, noise reduction dataset.
- [11] H. Demuth and M. Beale, *Neural Network Toolbox: User's Guide, Version 3.0*, The MathWorks, Natick, Mass, USA, 1998.
- [12] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals, and Systems*, vol. 2, no. 4, pp. 303–314, 1989.
- [13] S.-Z. Qin, H.-T. Su, and T. J. McAvoy, "Comparison of four neural net learning methods for dynamic system identification," *IEEE Transactions on Neural Networks*, vol. 3, no. 1, pp. 122–130, 1992.
- [14] M. J. D. Powell, "Radial basis functions for multivariable interpolation: a review," in *Proceedings of the IMA Conference*

- on Algorithms for the Approximation of Functions and Data*, pp. 143–167, Shrivenham, UK, July 1985.
- [15] W. Light, “Ridge functions, sigmoidal functions and neural networks,” in *Approximation Theory VII*, E. W. Cheney, C. K. Chui, and L. L. Schumaker, Eds., pp. 163–206, Academic Press, Boston, Mass, USA, 1992.
 - [16] T. M. Cover, “Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition,” *IEEE Transactions on Electronic Computers*, vol. 14, no. 3, pp. 326–334, 1965.

