Journal of Theoretical and Applied Information Technology
15th June 2017. Vol.95. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PARTICLE SWARM OPTIMIZATION FOR OPTIMIZING LEARNING PARAMETERS OF FUNCTION FITTING ARTIFICIAL NEURAL NETWORK FOR SPEECH SIGNAL ENHANCEMENT

OMAIMA N. A. AL-ALLAF

Assistant Prof., Dept. of CIS, Faculty of Sciences and Information Technology,

AL-Zaytoonah University of Jordan, P.O. Box 130, Amman (11733)- Amman- Jordan

Email: omaimaalallaf@zuj.edu.jo

ABSTRACT

Speech signals are effected by noise generated by various sources of interferences. Removing noise from speech signals can be regarded as an active research area in signal processing. Thus, we need powerful methods in this area. Therefore, Function Fitting (FitNet) Artificial Neural Networks model was used in this paper for enhancing speech signals. Particle Swarm Optimization (PSO) was used during FitNet learning process to optimize the FitNet learning parameters (such as learning rate, momentum variable and network weights) to achieve best results of speech signal enhancement. At the same time, different optimization techniques for optimizing the values of learning parameters were suggested in this work. This is done to improve the performance of FitNet model for signal enhancement. Better results (320 learning steps, PSNR equal 38 and mean square error (MSE) equal 0.0027) from experiments were achieved when adopting PSO with FitNet with swarm size equal 40 and PSO number of iterations equal 100. Good results (312 learning steps, PSNR equal 35.94 and MSE equal 0.00002) were obtained also when adopting the suggested optimization techniques (learning rate equal 0.00003, 5 hidden units in one hidden layer with the using of Levenberg-Marquardt (LM) as learning algorithm) for optimizing the learning parameters.

Keywords: *Signal Enhancement, Artificial Neural Networks, Function Fitting Model, Particle Swarm Optimization*

1. INTRODUCTION

Removing noise from speech signals can be regarded as a challenge problem in applications of speech enhancement, recognition and communication. Approaches related to noise reduction can be considered to reduce noise in speech signals under noisy environment. Many researches over past years were focused on this problem.

In the last years, the technologies related to multimedia applications have been increased. The capability of Artificial Neural Networks (ANNs) to approximate any non-linear function also makes them suitable for non-linear transformations commonly used in speech feature extraction [1]-[5]. Thus, ANNs had been widely used in the applications of image and signal processing [6], [7].

Many ANNs models that differ in architecture and training algorithms were suggested in the past. Multi layers ANNs and backpropagation neural network (BPNN) are largely used ANN models. Many literature studies were focused on approaches that speeding up the ANNs learning process [8], [9]. Therefore, we focus in this paper on suggesting approaches that reduce the learning time as possible as of the used ANN model.

Signal enhancement is process of performing filtering of a signal for noise reduction. Recently, ANNs are found to be an efficient tool for signal enhancement. ANNs can be used as a signal filtering approach to remove noise from any speech signal [10], [11].

An appropriate ANN approaches, architectures and learning parameters can be applied. Few literature researches suggested different ANNs models for removing noise from speech signal [11]-[19]. It is important to evaluate ANN approaches for solving signal processing problems [10].

Thus, we need to analyze these literature studies and suggest an approach for speech signal enhancement that leads to: less ANN learning time, less Mean Square Error (MSE) and best Peek Signal to Noise Ratio (PSNR).
Particle swarm optimization (PSO) is an optimization technique that is based on swarm intelligence to solve optimization problems. The algorithm is widely used [20]-[22]. PSO was used for solving different recognizing problems. We adopted PSO in our previous researches [23]-[25] for solving different problems. PSO is used with ANN in research [26] for noise reduction. And used also in [27] for speech recognition
In this paper, Function Fitting (FitNet) ANN model was constructed with different architectures and learning parameters for speech signal enhancement. Many optimization techniques for obtaining best values of learning parameters were suggested in this paper. This is done to improve the performance of FitNet ANN learning process and get best results.
At the same time, PSO was used in this work for optimizing the learning rate, momentum variable and network weights. This is done to increase the performance of speech signal enhancement system.
Mean Square Error (MSE) and Peek Signal to Noise Ratio (PSNR) were used to examine the performance of FitNet model.
In the remaining of this paper: related studies were included in section 2. Section 3 includes description of noise reduction whereas section 4 describes the steps of PSO algorithm. Section 5 describes the research methodology. Section 6 includes experiments and finally section 7 concludes this work.

2. RELATED STUDIES

ANN includes many neurons that operate in parallel and connected via weights. Speech signal enhancement includes processing of noisy speech to improve the quality of speech signal. Tlucak, 1999 [12] listed ANNs approaches for speech enhancement.
Whereas, Badri, 2010 [13] used BPNN and Recurrent network to enhance the noisy signals. And, Kevin (1988) [11] used BPNN approach to filter noise from signals. Experiments were conducted in this study using speech inputs. In order to compute the performance of ANN, the ANN outputs were compared with original speech signals.
And, Miry, et. al., 2011 [14] Adopted noise removing approach from speech signals using fuzzy and ANN
And, Pankaj and Anil (2012) [15] listed adaptive techniques for removing noise from speech signals such as: Adaptive Noise Cancellation (ANC)
Whereas, Chatterjee, et, al. (2013) [17] described Backpropagation ANN (BPNN) technique for enhancing the ECG signal. This is done using Recursive Least Square (RLC) algorithm. At the same time, Debananda, et, al. (2012) [16] described a noise filtering techniques using BPNN for signal enhancement. And, Andrew, et, al. (2012) [18] presented a recurrent auto encoder neural network for signal enhancement. This was used for automatic speech recognition (ASR).
Most of these literature studies includes limitations in performance (PSNR and MSE) in speech signal enhancement. We need to suggest ANN methodology that overcome as possible as these limitations and speed up the convergence time of ANN model.
We had been suggested in previous research (O. Al-Allaf (2015) [19] to use ANN models with different architectures for increasing the performance of speech signal enhancement. These models are: Function Fitting (FitNet), Nonlinear Autoregressive (NARX), Recurrent (RNNs), and Cascaded-ForwardNet. The four ANN models had been constructed and implemented using MathLab 2013a for signal enhancement. The training process of the four ANN models was conducted separatelly based on three optimization algorithms: Gradient descent (GD); Gradient descent with momentum (GDM); *and* Levenberg-Marquardt (LM). Experiments were executed based on different ANN architecture and learning parameters to determine the ANN model that lead to best performance. We got some limitations related to long ANN learning time.
Therefore, in this work, PSO is used to optimize the ANN weights, learning rate and momentum variable. This is done to speed up learning process and enhance the speech signals by obtaining high PSNR, less MSE and best values of R^2 (coefficient of determination of a linear regression).
At the same time, we suggested in this paper many optimization techniques related to learning parameters. This is done to obtain best values of FitNet ANN learning parameters (learning rate, momentum variables, number of hidden layers, and number of hidden units) to speed up FitNet learning process and get best results.

3. NOISE REDUCTION IN SPEECH SIGNALS

The techniques of speech enhancement focus on reducing noise from speech signal. Different techniques such as temporal filtering; spatial

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

temporal filtering; beamforming adaptive filtering; and so on, can be used for speech enhancement [4].

Many sytems related to speech signal enhancement had been suggested and experimented. These systems were based on cleaning speech signals from noise. This is based on improving the value of PSNR to improve the quality of speech signals. The performance of these systems were based on noise type and noise information [28].

The problem of noise reduction focus on recovering a speech signal x(n) from noisy speech signal y(n) that is corrupted using noise v(n) as shown in (1) [4]:

$$y(n) = x(n) + v(n) \qquad (1)$$

The additive noise is represented by Gaussian random process. The noise is assumed to be uncorrelated with the clean speech signal x(n). The relationship of speech signal model in Discrete Frequency Domain (DFD) can be obtained by applying Discrete Fourier transform (DFT) to (1) [4]. Speech enhancement methods can be classified into four main techniques: Filtering [29], [30], [17]; Noise subtraction [31]; Space Mapping [32]; and Hidden Markov Models (HMM) [33]. All of these techniques were described in details in our previous research [19].

4. PARTICAL SWARM OPTIMIZATION

PSO concepts were described in details in references [20]-[25]. Figure 1 describes the main steps of PSO [25]. Any individual in a population can be represented by a particle as N dimensional vector solution. The fitness function is computed for each particle. The particle position is adjusted according to (2) (to calculate the new velocity of this particle) to move to optimal position. This is done by comparing with other particles' neighbors [20]-[25].

$$\begin{aligned} v_i(t+1) &= w \times v_i(t) + \\ &(c1 \times rand \times (Pbest(t) - x_i(t))) + \\ &(c2 \times rand \times (Gbest(t) - x_i(t))) \end{aligned} \qquad (2)$$

Where,

W: weight that is ranged from 0 to 1,
V[]: particle velocity,
C1, C2: speeding factors with value 2,
Pbest: best value of particle i,
Xi: ith particle of swarm,
Gbest: best value that one of swarm particle reach it,

The new fitness of particle is calculated using (3) [20]-[25]:

$$x_i(t+1) = x_i(t) + v_i(t+1) \qquad (3)$$

5. PRESEARCH METHODOLOGY

Different ANN models such as: Cascaded Forward Neural Networks [34]-[36]; Function Fitting ANN [34], [35]; Nonlinear Auto-Regressive eXogenous (NARX) [34], [35]; Recurrent Neural Network (RNN) [34]-[38]; were constructed, implemented and discussed to remove noise from any signal (in details) in our previous research [19].

Many training algorithms for ANN models: Gradient descent (GD); Gradient descent with momentum (GDM); and Levenberg-Marquardt (LM) were adopted to train these models [19]. Experiments were conducted with good results. In many experiments, the training process may require long time. Therefore, we focus in this paper on selecting best values of ANN learning parameters for speech signal enhancement. This is done by suggesting many techniques and using PSO as will be shown in next sub sections 5.4 and 5.5.

5.1 Architecture and Parameters

FitNet ANN model that includes input, hidden and output layers was constructed in this work. Two neurons in input layer (One for signal and other for the noise) and 15 hidden units in hidden layer. One output unit that represent the signal after removing noise.

Three optimization training algorithms (GD; GDM and LM) were used to train the FitNet model separately.

The weights between input layer and hidden layer, and also the weights between hidden layer and output layer were initialized randomly within a certain range.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

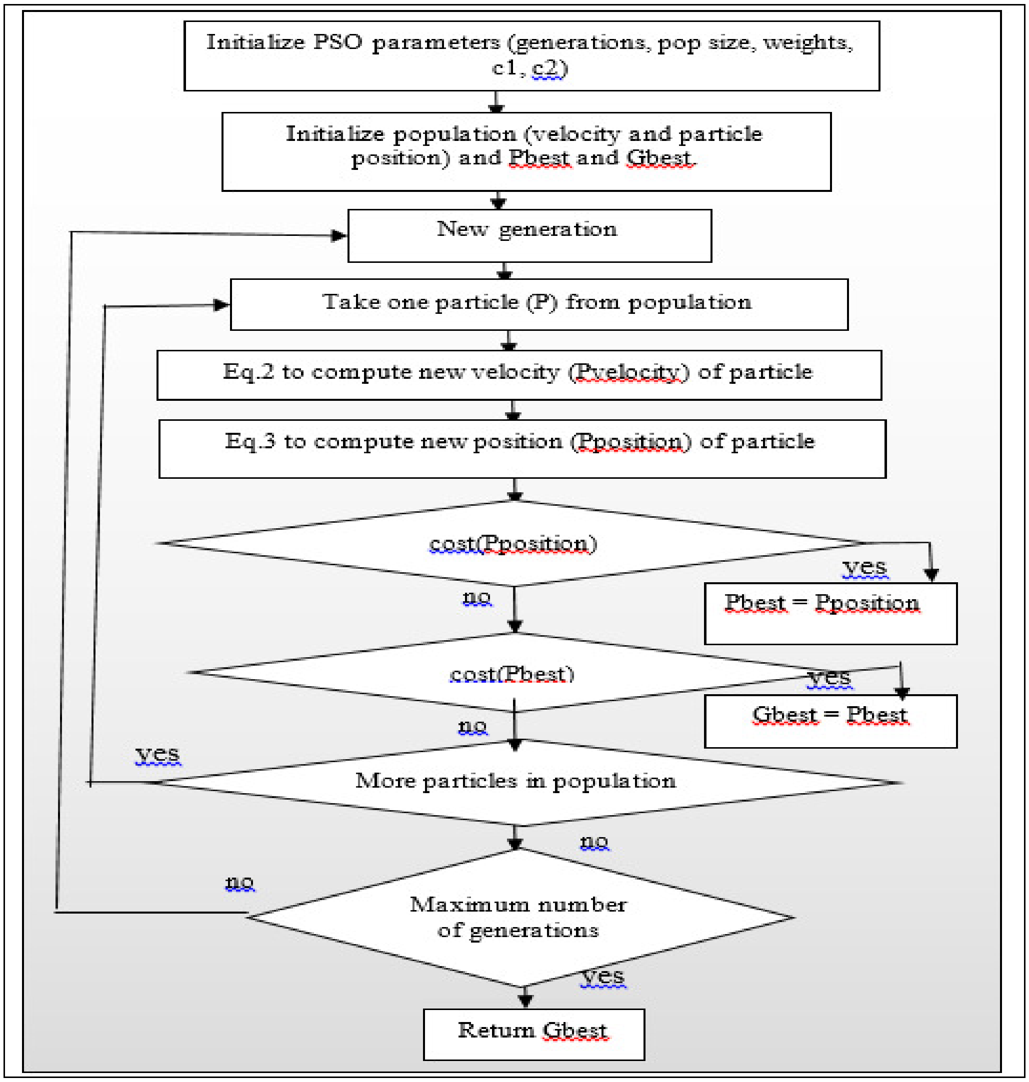


Figure.1: PSO algorithm [25]

Figure 2 shows the ANN FitNet model for speech signal enhancement. The training and testing processes were discussed later in the next sub sections 5.2 and 5.3. Simulation programs were written to implement the learning and testing processes of FitNet model.
The MathLab 2013a software was used to write these programs of speech signal enhancement system.
The signal sample size equal 50000 and the size of noise sample equal 50000. Audio files including music with song from MathLab Library (2013a) were used in training process.
Training samples were taken from MathLab library. Testing samples were downloaded from website "Odyssey FX | Wav Sound Effects" [39].

The FitNet ANN shown in Figure 2 consists of three layers. Each layer includes many neurons that are connected via weights. The sigmoid activation function is calculated for each neuron as shown in (4) and (5) [6], [8], [10].

$$net_p(j) = \sum_{i=1}^{N+1} w(j,i) \cdot x_p(i), 1 \le j \le N_h \qquad (4)$$

$$f(net_p(j)) = \frac{1}{1+e^{-net_p(j)}} \qquad (5)$$

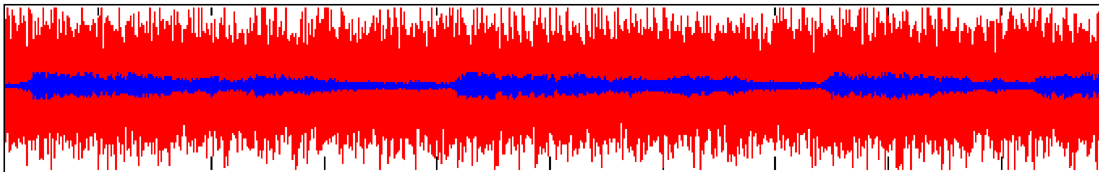
Whereas FitNet ANN error can be calculated using (6) to find the performance of network [6], [8], [10].

$$E(k) = \frac{1}{N_v} \sum_{p=1}^{N_v} [t_{pk} - y_{pk}]^2 \qquad (6)$$

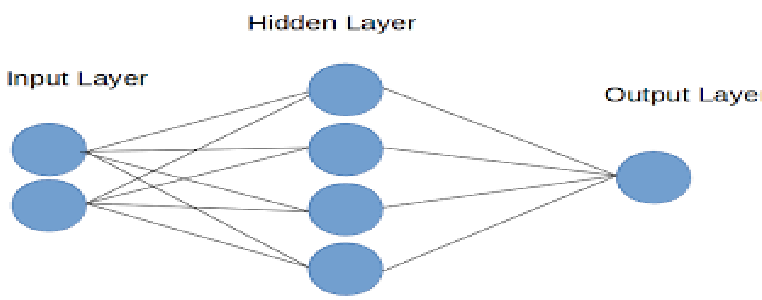
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Many techniques were suggested later in sub section 5.4 to get best values of learning parameters (learning rate, momentum variables and number of hidden units) to speed up the convergence time and avoid local minima of ANN model.

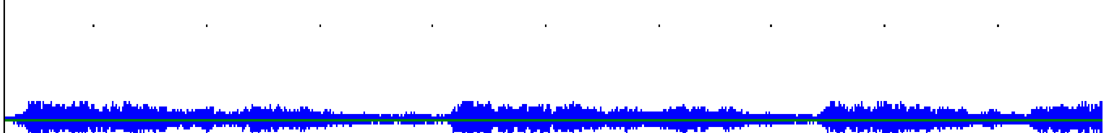
We will discussed also in sub section 5.5 the adoption of PSO to find the best values of learning parameters of FitNet ANN model for speech signal enhancement system to get better results.



(a) Input (noisy speech signal)



(b) FitNet ANN with 2 inputs and one output



(c) Output (Speech signal after removing noise)

Figure 2: (a)(b)(c) FitNet model for signal enhancement

5.2 Training Process of FitNet ANN Model

The training process of FitNet model includes many steps [19]:

1) Initializing the FitNet ANN weights with random numbers in the rage [0.001..0.009].
2) Initializing the learning parameters of training algorithm that is used to train the FitNet model. The values of learning rate (η) and momentum variable (α) are initialized with values in the range [0.1..0.9].
3) Set the variable "iterations" to zero. This variable is used to count the number of steps required to train the FitNet ANN model
4) Set the variable "Threshold error" with 0.0000001. This variable is used to test if the FitNet model is trained or not.
5) Set the variable "Total_error" to zero. This variable is used to compute the overall error of FitNet ANN.

6) Increment variable "iterations" with one.
7) Set variable "i" with 0. Where variable "i" represent signal sequence)
8) Open the file that includes the sound signal (s_i) that is selected to be used in FitNet ANN training, and then read the signal (s_i). This sound signal (s_i) is use as "input signal" to input layer neuron. Add noise to this sound signal (s_i) by adding random noise to second input neuron.
9) Initialize "target output" of the FitNet ANN model to be same as "input signal".
10) Calculate the outputs of hidden layer neurons and the outputs of output layer neurons using (4) and (5).
11) Calculate the $\text{error}_{(i)}$ of the signal (s_i): target output – actual output. Also, calculate Total_Error of FitNet using (6).
12) Increment "i" variable with one. If there is another sound signal to be used in training, then go to step8, otherwise, go to step 13.
13) Update weights between output and hidden layer units and also weights between hidden and input layer neurons using (8) and (9).
14) If Total_error <= Threshold_error then stop. Otherwise , then go to step 5.

5.3 Testing Process of FitNet ANN Model

The training process of FitNet model includes many steps [19]:

1) Open the file that includes the speech signal to be used in testing process. Read the sound signal to be used as "input signal" to input layer neuron.
2) Add noise to this sound signal by adding random noise to second input neuron.
3) Calculate outputs of hidden layer neurons using and also outputs of output layer neurons using (4) and (5).
4) Calculate the PSNR of input noisy signal.Then calculate the PSNR, MSE and R^2 of output clean signal.

5.4 Optimizing Learning Parameters

The performance of FitNet ANN model can be improved by applying many techniques to learning parameters used in training algorithm:

✓ Add bias unit (θ) of value equal 1 for each FitNet layer except the output layer. This is to provide a constant term in weighted sum of next layer units as in (7) to improve FitNet convergence.

$$NET_j = \sum_{i=1}^{N} X_i W_{ji} + \theta_j \qquad (7)$$

✓ Use momentum variable (α) ranged from 0.0 to 1.0 to reduce FitNet training time and enhance stability of learning process. The momentum variable is added to the weight adjustment equation that is proportional to the amount of previous weight change as in (8) and (9) [8].

$$W_{ji}^{new} = W_{ji}^{old} + [\Delta W_{ji}^{q}]^{new} \qquad (8)$$

$$[\Delta W_{ji}^{q}]^{new} = \eta\, \delta_i^q\, O_j^{q-1} + \alpha\, [\Delta W_{ji}^{q}]^{old} \qquad (9)$$

✓ Use beta (β) ranged from 0.1 to 1 in the sigmoidal function during FitNet training to determine steepness of sigmoid function shape as shown in (10) and (11). When β is set to 0.1, the learning is slowly converge, but when the value of β is set to 1, instability may occur.

$$OUT = F(NET_j) = \frac{1}{1 + e^{\beta \times NET_j}} \qquad (10)$$

$$F'(NET_j) = \beta \times (OUT\,(1 - OUT)) \qquad (11)$$

✓ Adding random Noise during the training process.to avoid the local minimum problem in FitNet training. This is to shift the location of error function by permitting the descent process to escape from local minimum and continue on its downward search for a global minimum. Adding random noise can improve the network's ability to generalize.

✓ Changing learning rate value (η) by allowing the value of η to begin at a high value and to decrease during the learning process. This is for efficient FitNet learning.

5.5 PSO for Selecting Best Learning Parameters of ANN Training Algorithm

PSO is adopted in this paper to select best values of: networks weights, learning rate and momentum variables. These learning parameters are used in training algorithm for FitNet ANN learning process for speech signal enhancement system. This is done to get better PSNR, lowest MSE and reduce the learning time as possible as.

Many steps will be executed to combine PSO with ANN training algorithm that is used to train FitNet ANN model:

- ✓ The PSO algorithm, its parameters, variables and equations were disscused earlier in details in section 4 and Figure 1.
- ✓ The PSO algorithm is called at the beginning of executing the training algorithm of FitNet ANN model. This is done after initialization of learning parameters of ANN training algorithm.
- ✓ The randomly initialized FitNet ANN learning rate, momentum variable, weights between input and hidden layers' neurons and the weights between hidden and output layers' neurons are used as particles of PSO population in the first PSO generation.
- ✓ PSO is executed according to its initalized variables, parameters and equations as discussed in Figure 1. This execution is repeated until reach the best values of ANN learning parameters and required number of PSO generations.
- ✓ After that, the particles in population of last PSO generation represent the best values of: FitNet ANN learning rate, momentum variable, weights between input and hidden layers' neurons and weights between hidden and output layers' neurons. These values are used to begin the execution of FitNet ANN training algorithm.

6. EXPERIMENTAL RESULTS

The architecture of FitNet ANN model and training algorithms were discussed earlier in sub section 5.1. And the speech signal samples that were used for training and testing processes were discussed also in sub section 5.1.

Also the research parameters such as learning rate, momentum variable, ANN weights, threshold error and other variables used in ANN training algorithm were discussed also in sub section 5.2.

The constructed FitNet model consists of 3 layers (input layer, hidden layer and output layer). The input layer includes Two neurons: one represent speech signal and other neuron to represent noise. The hidden layer includes 10 neurons. Finally output layer includes only one neuron that represent the clean speech signal after removing noise from it as shown in Figure 2.

Number of iterations, MSE, R^2 and PSNR are used to check the performance of the signal enhancement system. PSNR will be calculated between the output clean signal and input noisy signal to check the quality of signal enhancement system. The efficiency of this FitNet ANN model was tested by several speech signals.

The first Three experiments were based on training the FitNet model separately using the LM, GM and GDM training algorithms respectively. One simulation program for each experiment. Table 1 shows: MSE, PSNR, R^2 and the number of iteration required to train FitNet model.

Table 1: FitNet (with 10 hidden units) results

Algorithm	Itera-tions	MSE	PSNR	R2
LM	414	0.00005	35.34	0.99
GD	475	0.0011	34.23	0.97
GDM	460	0.0006	34.75	0.98

We can note from Table 1 that, best results (smallest number of iterations, high value of PSNR, least MSE) were obtained using the Levenberg-Marquardt (LM) training algorithm.

6.1 Optimization of Learning Rate

Other Four experiments were conducted for training the FitNet model using Levenberg-Marquardt (LM) algorithm with 10 hidden units using different values of learning rate (0.03, 0.003, 0.0003 and 0.00003) separately. Table 2 shows: number of iterations, MSE, PSNR and R2 for each FitNet experiment.

Table 2: Effects of learning rate

learning rate	R^2	PSNR	MSE	Iter – ations
0.03	0.98	35.37	0.00006	413
0.003	0.94	35.76	0.00003	397
0.0003	0.994	35.96	0.00011	365
0.00003	0.997	36.01	0.00001	311

We can note from Table 2 that there are no notable differences between the values of MSE, PSNR and R2 for these Four experiments when decreasing the value of learning rate. But the number of iterations were decreased when decreasing learning rate.

We suggest in this paper to use PSO algorithm to select best value of learning rate that is appropriate to the constructed FitNet ANN architecture.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

6.2 Optimization of Number of Hidden Layers and Hidden Neurons in Hidden Layer

To select the best value of number of hidden neurons in hidden layer, other Three experiments were executed. This is done based on training the FitNet ANN model using LM algorithm with different number of hidden layer neurons (5, 15 and 25) separately. The value of learning rate is selected to be 0.00003. Table 3 shows the impact of number of hidden neurons on: number of iterations, MSE, PSNR and R^2 factors.

Table 3: Different Number of hidden units in hidden layer

Number of hidden neurons	R^2	PSNR	MSE	Itera-tions
5	0.994	35.94	0.00002	312
15	0.996	35.97	0.00001	397
25	0.999	36.98	0.00001	423

We can note from Table 3 that increasing the number of hidden layer neurons for this FitNet ANN model will lead to increasing the number of iterations required for learning process. At the same time, there is no big differences in values of MSE, PSNR and R^2.

To decide the best value of number of hidden layers in the constructed FitNet ANN model, another Four experiments were executed using different number of hidden layers (1, 2, 3, 4) separately each with 15 hidden units. Table 4 shows the results of these Four experiments.

Table 4: Different Number of hidden layers

Number of hidden layers	R^2	PSNR	MSE	Itera-tions
1	0.996	37.97	0.00001	397
2	0.997	36.98	0.00004	423
3	0.998	35.97	0.00059	498
4	0.999	34.97	0.00098	560

We can note from Table 4 that increasing the number of hidden layers will increase the learning time with little different in FitNet performance. FitNet ANN model with 4 hidden layers require greater value of number of iterations.

6.3 Testing Process and Untrained Signals

The testing process of the FitNet ANN model was based on using LM training algorithm and 10 hidden units in one hidden layer.

The quality of output clean speech signal is examined using PSNR. The PSNR of the input noisy speech signal is calculated. This is done by applying the noisy speech signal to the input layer of FitNet model and executing the testing algorithm (sub section 5.3). After that the PSNR of the output speech signal is calculated to determine the ability of FitNet model to enhance noisy speech signals.

As results of testing process: MSE is 0.00007; PSNR of input noisy signal is equal to 16.22, and PSNR of output clean signal is equal to 34.22. As a result, FitNet model has the ability to remove noise from any speech signal with small differences between PSNR values. At the same time, the values of PSNR of output signal is high according to the values of PSNR of corresponding input noisy signal.

The FitNet model was tested using trained and untrained speech signals. Figure 2 shows the input noisy speech signal which was already used in training process. Figure 2 shows also the trained speech signal after removing noise from it using FitNet model. Whereas, Figure 3 shows the untrained speech signal after removing noise using FitNet model. The MSE value is equal to 0.000081; PSNR of input noisy Signal is equal to 20.45; and PSNR of output clean Signal is equal to 34.55.

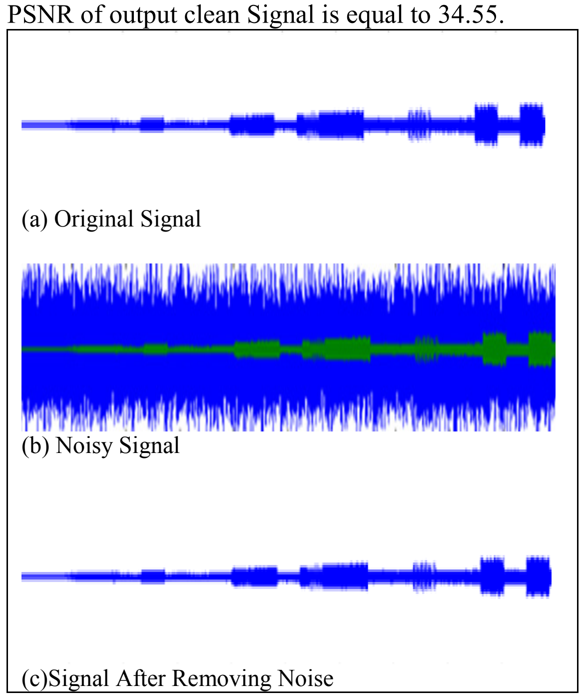


Figure 3: (a) Original untrained speech signal, (b) Untrained speech signal with noise, (c) Speech signal after filtering using FitNet

Different untrained noisy speech signals were used also in testing process of FitNet ANN model with 10 hidden units in one hidden layer that is trained using LM algorithm.

Figure 4 shows the results of FitNet ANN testing process of Five different untrained speech signals.

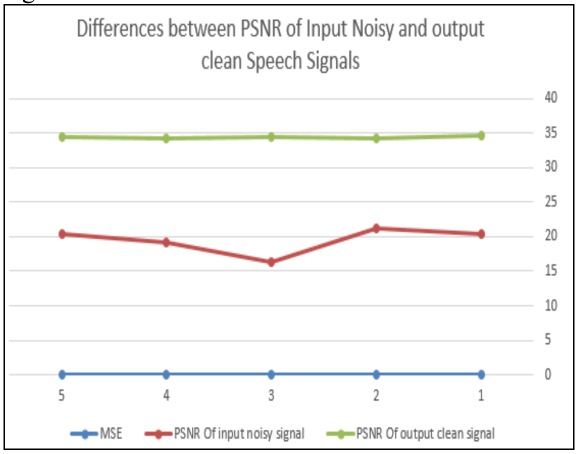


Figure 4: Differences of PSNR of noisy and clean untrained speech signals

6.4 Different Swarm size and Iterations in PSO

Five experiments were executed based on different PSO swarm size (20, 30 and 40) and different number of iterations (100 and 150) for optimizing the learning parameters of FitNet. FitNet model with 10 hidden units in one hidden layer. This model is trained using LM algorithm. Table 5 shows the MSE, PSNR and number of iterations required for training process.

Table 5: FitNet Results when Using PSO

	FitNet Leaning Iterations	MSE	PSNR
PSO parameters			
Swarm =40 Iterations=100	320	0.0000 27	38
Swarm=40 Iterations=150	368	0.0096	37
Swarm=30 Iterations=100	397	0.0031	37
Swarm=30 Iterations=150	420	0.0055	33
Swarm=20 Iterations=150	463	0.0089	30

We can note from Table 5 that the best results (better PSNR is equal 38 and lowest MSE is equal 0.0027) were obtained when increasing the PSO swarm size to 40 and decreasing number of PSO iterations to 100. Also, the number of iterations required to train the FitNet ANN is reduced to 320 iterations.

6.5 Comparisions with other Works

The results of literature studies related to speech signal enhancement using ANN showed that, performance of suggested ANN models to enhance untrained signal is low in comparision with the the trained speech signals. At the same time, many of the suggested ANN models in literature studies require long time for convergence. Table 6 shows the comparisons between the related studies and our work.

7. CONCLUSION

In this paper, FitNet ANN model was constructed with 3 layers for speech signal enhancement to remove noise from any speech signal. The input layer includes one neuron for speech signal and other neuron for noise. The output layer consists of one neuron that represent clean speech signal. Different number of neurons in hidden layer were used for FitNet model (5, 15 and 25) respectively. The FitNet ANN was trained separately using three different training algorithms (GD; GDM and LM). PSO algorithm was used to optimize the FitNet learning parameters and weights. The FitNet training and testing programs were written using MathLab 2013a software. Training samples: stereo speech signals (noisy and clean) were taken from MathLab 2013a. Testing samples obtained from Odyssey FX | Wav Sound Effects [39].

Many experiments were conducted for training/testing FitNet model. This is done separately with different: number of hidden layers, number of hidden layer neurons, learning parameters to get best results.

We can note from experimental results that decreasing the value of learning rate will decrease the number of iterations required for FitNet training. At the same time, increasing the number of hidden layers and also number of hidden neurons will increase number of iterations for learning process. FitNet model has the ability to remove noise from trained and untrained speech signals.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 6: Comparisons Between Different Studies

Research	ANN Model	Architecture	Training algorithms	Performance Measures	Testing ability
Kevin S. Cox (1988) [11]	BPNN		BP	MSE=0.1 Frequency	Satisfied only trained samples
J. Tlucak, et al [12]	BPNN	3 layers 5 input neurons Many hidden neurons One output neurons	BP	Error=0.07 SNR=36.6	Satisfied only trained samples
Lubna Badri, 2010 [13]	Multilayer Neural Network (MLNN) and RNN	4layers 2 input neurons 2 hidden layers (7 units, 3units) 1 output neuron	GD, GDM, GDA, LM, and GDX	MSE = 0.0112285 MSE (MLNN)= 0.0153542 MSE (RNN)= 0.0151797 Time	Satisfied only trained samples
Debananda Padhi, et, al (2012) [16]	BPNN	4layers Input layer 2 hidden layers One output layer	BP	MSE PSNR (not mentioned)	Satisfied only trained samples
O. Al-Allaf (2015) [19]	FitNet NARX RNN Cascaded	3 layers 2 input neurons 10 hidden neurons 1 output neurons	LM, GD, and GDM	MSE=0.00005 PSNR=35.35 R^2=0.99 Iterations=413	Satisfy trained/ untrained samples
Our work O. Al-Allaf (2017)	FitNet With PSO	3 layers 1,2,3 and 4 hidden layers 2 input neurons 10 hidden neurons 1 output neuron	LM, GD, and GDM	MSE=0.000027 PSNR=38 R^2=0.994 Iterations=320	Satisfy trained/ untrained samples

Finally, we suggested many recommendations to optimizing the FitNet learning parameters to improve the performance of learning process and get best results. These recommendations are: adding bias unit; using momentum variable (α); changing learning rate (η) during learning process; using the beta term (β); and adding small random noise to FitNet weights. We suggested to use other optimization algorithms in future research to improve our results as possible as.

ACKNOWLEDGMENT

The author would like to acknowledge Al-Zaytoonah University of Jordan for supporting the Scientific Research.

REFERENCES

[1] S.Y. Kung, J.N. Hwang, Neural networks for intelligent multimedia processing, Proc. IEEE 86 (6), (1998) 1244–1272.

[2] Aurelio Uncini∗, Audio signal processing by neural networks, Neurocomputing 55 (2003) 593 – 625.

[3] V.Saradhadevi and Dr.V.Sundaram, A Novel Two-Stage Impulse Noise Removal Technique based on Neural Networks and Fuzzy Decision, International Journal of Computer Applications (0975 – 8887), Volume 21– No.3, May 2011.

[4] Yang Liu, Mingli Xiao and Yong Tie, A Noise Reduction Method Based on LMS Adaptive Filter of Audio Signals, the 3rd international conference on multimedia technology

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

(ICMT2013), november, DOI: 10.2991/icmt-13.2013.123

[5] S. Parveen and Phil Green, Speech Enhancement With Missing Data Techniques Using Recurrent Neural Networks, ICASSP, pp: I – 733 - I – 736, 2004 IEEE.

[6] R. P. Lippmann. "An Introduction to Computing with Neural Nets," IEEE ASSP Magazine, vol.4, no.2, April 1987, pp.4-22.

[7] N. K. Ibrahim, R.S.A. Raja Abdullah and M.I. Saripan. "Artificial Neural Network Approach in Radar Target Classification," Journal of Computer Science, vol. 5, no.1, 2009, pp.23-32, ISSN: 1549-3636, Science Publications.

[8] P. D. Wasserman. Neural Computing: Theory and Practice, Van Nostrand Reinhold Co. New York, USA, 1989, ISBN: 0-442-20743-3.

[9] Ra´ul Rojas, Neural Networks: A Systematic Introduction, Springer, Berlin Heidelberg NewYork, 1996.

[10] Yu Hen Huand and Jenq-Neng Hwang, handbook Of Neural Network Signal Processing, Electrical Engineering And Applied Signal Processing (Series), Crc Press Llc, 2002

[11] Kevin S. Cox, An Analysis Of Noise Reduction Using Back-Propagation Neural Networks, Thesis, Faculty Of The School Of Engineering Of The Air Force Institute Of Technology, Air University, Master of Science In Computer Eng. Captain, Usaf, Afit/Gce/Eng/88d-3, 1988.

[12] J. Tlucak, et al, Neural Network Based Speech Enhancement.,Radioengineering. Vol. 8, No. 4, Dec1999.

[13] L. Badri, Development of Neural Networks for Noise Reduction, The International Arab Journal of Information Technology, Vol. 7, No. 3, pp:289-294, July 2010.

[14] M. Miry, et. al. Adaptive Noise Cancellation for speech Employing Fuzzy and Neural Network, Iraq J. Electrical and Electronic Eng, Vol.7, No.2, 2011, pp: 94-101.

[15] Pankaj Bactor and Anil Garg, Different Techniques for the Enhancement of the Intelligibility of a Speech Signal, International Journal of Engineering Research and Development,Vol.2, Issue.2, July 2012, pp:57-64.

[16] Debananda Padhi, et, al. Filtering Noises from speech Signal : A BPNN approach, International Journal of Advanced Research in CS and Software Engineering, Vol.2, Issue.2, Feb 2012.

[17] Kalyan Chatterjee, et, al. Adaptive Filtering and Compression of Bio- Medical Signals Using Neural Networks, International Journal of Engineering and Advanced Technology (IJEAT),Vol.2 Issue.3, Feb2013, pp:323-327, ISSN: 2249 – 8958.

[18] Andrew Maas, et, al. Recurrent Neural Networks for Noise Reduction in Robust ASR, INTERSPEECH 2012, 13th Annual Conference of the International Speech Communication Association, Portland, Oregon, USA, September 9-13, 2012. ISCA 2012.

[19] Omaima N. A. AL-Allaf , Removing Noise from Speech Signals Using Different Approaches of Artificial Neural Networks, I.J. Information Technology and Computer Science, 2015, Vol.7, No.7, pp:8-18.

[20] Qinghai Bai, Analysis of Particle Swarm Optimization Algorithm, Computer and Info Science, Vol.3, No.1, 2010.

[21] H. Kuo, J. Chang and C. Liu, Particle Swarm Optimization For Global Optimization Problems, Journal of Marine Science and Technology, Vol. 14, No. 3, pp:170-181, 2006.

[22] Dian P.R., Siti M.S. and Siti S.Y., Particle Swarm Optimization: Technique, System and Challenges, International Journal of Computer Applications, Vol.14, No.1, pp:19-27, Jan2011.

[23] Shahlla A. AbdAlKader and Omaima N. Ahmad AL-Allaf, Particle Swarm Optimization Based Discrete Cosine Transform for Person Identification by Gait Recognition, The 7th International Conference on Information Technology (ICIT 2015), AlZaytoonah University of Jordan, Amman, Jordan, pp:156-163, 12 May 2015, doi:10.15849/icit.2015.

[24] Omaima N. Ahmad AL-Allaf, Shahlla A. AbdAlKader, Performance Analysis of Different Feature Extraction Algorithms Used with Particle Swarm Optimization for Gait Recognition System, International Journal of Recent Technology and Engineering (IJRTE), ISSN: 2277-3878, Volume-4, Issue-2, pp:23-30, 19 May 2015, Blue Eyes Intelligence Engineering & Sciences Publication Pvt. Ltd.

[25] Omaima N. Ahmad AL-Allaf, Improving the Performance of Particle Swarm Optimization for Iris Recognition System Using Independent Component Analysis, The 2015 World Congress in Computer Science, Computer Engineering, and Applied Computing (WORLDCOMP'15) The 17th International Conference on Artificial Intelligence, (ICAI'15), July 27-30, 2015, Las Vegas, USA).

[26] Manpreet Kaur, Danvir Mandal, Speckle Noise Reduction in Medical Ultrasound Images using Particle Swarm Optimization with Artifical Neural Networks, International Journal of Electronics & Communication Technology Vol. 2, Issue 3, pp:147-150, Sept. 2011, ISSN : 2230-7109(Online) | ISSN : 2230-9543(Print)
[27] Rita Yadav and Danvir Mandal Speaker Recognition using Particle Swarm Optimization (PSO), International Journal of Electronics & Communication Technology, Vol. 2, ISSN : 2230-7109(Online) | ISSN : 2230-9543(Print) Issue 3, pp:167-171, Sept. 2011
[28] J.P. Haton, Problems and solutions for noisy speech Recognition, Journal De Physique Iv, Colloque C5, supplement au Journal de Physique 111, Vole 4, May 1994.
[29] Widrow, B. et, al. Adaptive Noise Cancelling: Principles and Applications, Proc. IEEE, 63(12), pp: 1692-1716, 1975.
[30] Koo, B., Gibson, J.D., Gray, S.D. Filtering of Colored Noise for Speech Enhancement and Coding, International Conference on Acoustics, Speech, and Signal Processing, 1989. ICASSP-89, pp:349-352, Glasgow, 1989.
[31] Hermansky H. An Efficient Speaker-independent Automatic Speech Recognition by Simulation of Some Properties of Human Auditory Perception, Proc. of: Acoustics, Speech, and Signal Proc., IEEE Int. Conf. on ICASSP '87, Vol.12, DOI:10.1109/ICASSP.1987. 1169803, pp:1159-1 162, Dallas, 1987.
[32] Tamura, S., Waibel, A.: Noise Reduction using Connectionist Models, Proc. ICASSP-88, 553-556, New York, 1988.
[33] Varga, A.P., Moore, R.K.: Hidden Markov Model Decomposition of Speech and Noise, Proc. ICASSP-90, 845-848, Albuquerque, 1990.
[34] O. De Jesus and M. Hagan, "Backpropagation Algorithms for a Broad Class of Dynamic Networks," IEEE Transactions on Neural Networks, vol.18, No.1, pp.14 -27, Jan.2007.
[35] [35] MathWorks, Neural Network Toolbox 7.0, MathWorks Announces Release 2010a of the MATLAB and Simulink Product Families, 2010, MathWorks, Inc.
[36] O. De Jesus and M.T. Hagan, "Backpropagation Through Time for a General Class of Recurrent Network," in Proc. of the International Joint Conference on Neural Networks, Washington, DC, vol.4, pp.2638–2643, 2001, ISBN: 0-7803-7044-9, DOI: 10.1109/IJCNN.2001.938786.
[37] James Martens and Ilya Sutskever, Learning Recurrent Neural Networks with Hessian-Free Optimization, Proceedings of the 28 th International Conference on Machine Learning, Bellevue,WA, USA, 2011.
[38] Alex. Graves, et, al. A Novel Connectionist System for Improved Unconstrained Handwriting Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 5, 2009.
[39] Odyssey FX | Wav Sound Effects, http://www.wavealchemy.co.uk/odyssey_fx/pid53#purchase, 2014 Wave Alchemy Limited, Company No: 07797886, 20 Burns Street, Ilkeston, Derbyshire, United Kingdom, DE78AA, Downloaded:21 Jun 2015.