

Classification of Prostate Cancer using Wavelet Neural Network

Mohanad Najm Abdulwahed

Materials Department, University of Technology, Baghdad, Iraq

Article Info

Article history:

Received Jun 17, 2018

Revised Jul 27, 2018

Accepted Aug 20, 2018

Keywords:

Artificial neural network

Prostate cancer

Wavelet neural network

ABSTRACT

Prostate cancer is the century disease that endanger the life of men. The earlier to diagnose the disease, the probability of curing this disease is higher. Therefore, new approaches of diagnosis is required to effectively detect the prostate cancer in early stage compared to the traditional methods. Therefore, WNN is a new adopted approach in prostate cancer diagnosis. Morlet function is used as an activation function of wavelet neural network (WNN) and back propagation (BP) is applied to train the Wavelet network. WNN classifies prostate cancer according to three factors: patient age, PSA level, and prostate volume. WNN performance is evaluated based on the percentage of classification and the computational complexity of several cases. The results of the simulation show that WNN has lower mean squared error (MSE) than the Neural Network (NN).

Copyright © 2018 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

Mohanad Najm Abdulwahed,

Materials Department,

University of Technology,

Baghdad, Iraq.

Email: mohanad.najm@yahoo.com

1. INTRODUCTION

Today, prostate cancer has become one of the most dangerous diseases that can affect the health of men, and the detection of the disease in early stages can help to get rid of the disease easily. The traditional methods for diagnosis the prostate cancer depended on individual biomarkers and tested the tissue and cell specimens. The genetic and age are the coefficients that help to diagnosis the prostate cancer, and they are the first data must be available to determine the disease. Prostate-Specific Antigen (PSA) is the protein biomarker that usually used to detect the patient of prostate cancer [1]. The range of the PSA is recorded low (less than two ng/mL) in the blood of the healthy men, and it is raised for the prostate cancer patients more than two [2]. Although the PSA is an important indicator that can detect the prostate cancer disease, the other multiple factors should be taken into account to the diagnosis of the illness such as age and genetic [3].

There are several studies have discussed the diagnosis of the prostate cancer in early stages applying more recent and effective methods. Kallen et al. [4] used an example of Convolutional Neural Network (CNN) and he applied both Support Vector Machine (SVM) and a pre-trained CNN for classification and feature extraction respectively. In automated classification of Glason grading, Gorelik et al. [5] segmented the microscopic images into a meaningful pathologically segments such as, Lumen, stroma and lymphocyte etc. Doyle et al. [6] proposed a Bayesian multi-resolution method for classification using AdaBoost. In Torsten et al. [7] used a classification by learning vector quantization in which the Neural Networks training were adapted into different sample sizes per class. Another recent research article by Litjens et al. [8] using CNN for hematoxylin classification and eosin images of prostatic tissue.

Wavelet networks were firstly introduced by Zhang and Benveniste [9] as an alternative approach to FeedForward Neural Networks (FFNN) that fix the drawbacks of NNs and Wavelet Analysis(WA) while it has the best performance using both of these approaches [10]. Chen, Feng [10] used Wavelet Neural Network

(WNN) in system modelling and time series prediction based on multiresolution learning. Their results indicated that WNN has a powerful approximation capability and suitability in prediction and modeling. Therefore, WNN is a significant tool in classification and signals processing [11]-[12].

In this article, a new approach for classification of prostate cancer’s risk is proposed using WNN. The simulation of WNN method is performed to get high classified results. The new approach simulation result is compared with the traditional classification approaches such as Neural Networks (NN). This article is organized as follows. Part 2 introduces the data collection. Part 3 presents WNN with a description of its applied learning algorithm. In Part 4 the simulation results of both NN and WNN is summarized. Finally, Part 5 provides the conclusion of this paper.

2. DATA COLLECTION

Finding and collecting the data is very important for training and testing the WNN and NN. In this work, the data will be divided into three sections, Age, PSA and the Prostate volume. For human’s age, the risk of prostate cancer is highly expected starting from the age of 50 and above. PSA is a protein produced by the malignant and the normal cells of the prostate gland [13]. The range of PSA measurements ranges from 2 to 4 (ng/ml) in the blood for the healthy men. However, the indication of PSA varies with age as it shows a very high probability of cancer with PSA value more than 4 ng/ml and an age over 50. However, features of PSA diagnosing the disease decreases with patients of Prostate cancer with large volume of Prostate because the co-occurring decreases the effect of the cancer on PSA [14]. The volume of the prostate indicate an apparent symptom when it has enlarged affecting the urethra. That patients data have been provided by Baghdad Medical City hospital in April 2017.

3. WAVELET NEURAL NETWORKS (WNNs)

Wavelet Analysis is the result of both Fourier Transform (FT) and Window Fourier Transform (WFT) to improve the shortcomings of FT. Fourier transform is a common method in digital signal processing. However, when it was reported in the time series prediction, it showed the time limitations. WA is an analytical and mathematical tool for a wide range of research applications. Wavelet Analysis (WA) is used presently for time position, analysis of time series and intensity [15] The wavelet is a special waveform with a finite duration at the average zero point. The WA deploys a unique function introduced as the mother wavelet. This wavelet unique function is the result of a series of basic orthogonal sets composed of a father wavelet and a mother wavelet which satisfies:

$$\int \varphi(t) dt = 1 \tag{1}$$

$$\int \psi(t)dt = 0 \tag{2}$$

The wavelet family has what is called as the children of wavelet, which is a translated and dilated form of wavelet mother:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a_j}} \psi\left(\frac{t-b}{a}\right) \tag{3}$$

4. WAVELET NEURAL NETWORK CONFIGURATION (WNNC)

The WNN configuration is presented and displayed in Figure 1. The significant structure of WNN is introduced as the following [16]:

1. WNN Parameters initialization The matrix that exists between the input and the hidden layer is $W = (W_{jk})_{p \times m}$

The matrix that exists between the hidden and the output layer: $V = (V_{ij})_{n \times p}$

Hidden layer neuron’s Dilation vector: $a_j = (a_1, a_1, \dots a_p)$

Hidden layer neuron’s Translation vector: $b_j = (b_1, b_1, \dots b_p)$

Where m, p, n, stands for the input layer, hidden layer and the input layer of WNN respectively. The parameters initialization is done arbitrarily.

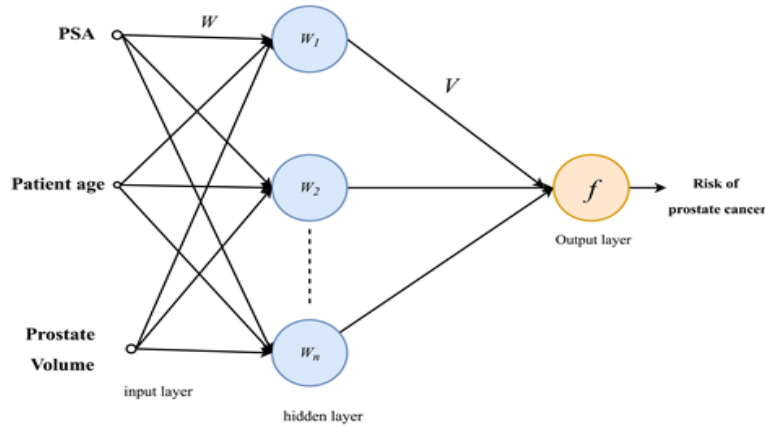


Figure 1. Wavelet Neural Network configuration

2. Generate the output of the network by computing the forward pass by applying the following Equation:

$$\psi_{a_j,b_j}(\sum_{k=1}^m W_{jk} x_k) = \frac{1}{\sqrt{a_j}} \psi\left(\frac{net_j^{(1)} - b_j}{a_j}\right) \tag{4}$$

Where:

$$net_j^{(1)} = \sum_{k=1}^m W_{jk} x_k \tag{5}$$

The *i*th node output at the output layer is expressed as:

$$y_i = f\left(\sum_{j=1}^p V_{ij} * net_j^{(1)}\right) = f(net_j^{(2)}) \tag{6}$$

where *X* is the WNN input vector and it is expressed as $X = (x_1, x_2, \dots, x_m)$. when *X* is achieved, the *j*th output node of the hidden layer can be found out.

3. Compute the output total error by applying the following equation:

$$E = \frac{1}{2} \sum_{q=1}^Q \sum_{i=1}^n (d_{qi} - y_{qi})^2 \tag{7}$$

Where *E* stands for the total error, *Q* stands for training samples number for each sample *q*, *t_q* stands for the desire vector. The desire vector is expressed as $T_q = (t_{q1}, t_{q2}, \dots, t_{qn})$, and *y_{qi}* stands for the output vector and it is expressed by $Y_q = (y_{q1}, y_{q2}, \dots, y_{qn})$. To be able to achieve the output minimum total error is considered a significant advantage of WNN .

4. Find each parameter partial derivatives as expressed by $\frac{\partial E}{\partial W_{jk}}$, $\frac{\partial E}{\partial V_{ij}}$, $\frac{\partial E}{\partial a_j}$ and $\frac{\partial E}{\partial b_j}$

$$\frac{\partial E}{\partial V_{ij}} = - \sum_{q=1}^Q (t_{qi} - y_{qi}) \cdot y_{qi} \cdot (1 - y_{qi}) \cdot \psi_{a_j,b_j}(net_{qj}^{(1)}) \tag{8}$$

$$\frac{\partial E}{\partial W_{jk}} = -a_j^{-1} \sum_{q=1}^Q \left[\psi'_{a_j,b_j}(net_j^{(1)}) \cdot x_{qk} \cdot \sum_{i=1}^n (t_{qi} - y_{qi}) \cdot y_{qi} \cdot (1 - y_{qi}) V_{ij} \right] \tag{9}$$

$$\frac{\partial E}{\partial a_j} = -a_j^{-1} \sum_{q=1}^Q \left\{ \left[\frac{a_j^{-1}}{2} \psi_{a_j,b_j}(net_{qj}^{(1)}) + \frac{net_{qj}^{(1)} - b_j}{a_j^2} \psi'_{a_j,b_j}(net_{qj}^{(1)}) \right] \cdot \sum_{i=1}^n (t_{qi} - y_{qi}) \cdot y_{qi} \cdot (1 - y_{qi}) V_{ij} \right\} \tag{10}$$

$$\frac{\partial E}{\partial b_j} = -a_j^{-1} \sum_{q=1}^Q [\psi'_{a_j, b_j}(net_j^{(1)}) \cdot \sum_{i=1}^n (t_{qi} - y_{qi}) \cdot y_{qi} \cdot (1 - y_{qi}) V_{ij}] \tag{11}$$

where

$$\psi'_{a_j, b_j}(net_{qj}^{(1)}) = \frac{1}{\sqrt{a_j}} \psi' \left(\frac{net_j^{(1)} - b_j}{a_j} \right)$$

5. Define the learning rate and momentum values as $\alpha = 0.99$ and $\eta = 0.2$ to update WNN parameters as shown in Equations (12-15) [15, 16]

$$V_{ij}(t + 1) = V_{ij}(t) - \eta \frac{\partial E}{\partial v_{ij}} + \alpha [V_{ij}(t) - V_{ij}(t - 1)] \tag{12}$$

$$v_{ij}(t + 1) = v_{ij}(t) - \eta \frac{\partial E}{\partial v_{ij}} + \alpha [v_{ij}(t) - v_{ij}(t - 1)] \tag{13}$$

$$a_j(t + 1) = a_j(t) - \eta \frac{\partial E}{\partial a_j} + \alpha [a_j(t) - a_j(t - 1)] \tag{14}$$

$$b_j(t + 1) = b_j(t) - \eta \frac{\partial E}{\partial b_j} + \alpha [b_j(t) - b_j(t - 1)] \tag{15}$$

5. RESULTS AND DISCUSSION

This part discusses WNN in diagnosis of prostate cancer. Three different items were used to measure prostate cancer risk, patient age, PSA, and prostate volume. In this article, WNN is applied as a new method: it consists of an input layer (m), hidden layer (p) and an output layer (n). The input layer, output layer and the hidden layer consists of three nodes, one node and seven nodes respectively as shown in the following equation [16].

$$N_{hidden} \geq (2 * N_{input}) + 1 \tag{16}$$

Where N_{hidden} and N_{input} denotes the neurons number in the hidden and the input layer respectively. Morlet function is deployed in the hidden layer as an activation function, and sigmoid function is deployed in output layer [11]. Equation 17 below indicates the Morlet function, as displayed in Figure 2. Table 1 presents the indication of th applied paparameters in this work.

$$\psi(t) = \cos(1.75 * t) \exp\left(-\frac{t^2}{2}\right) \tag{17}$$

Where, t is the summation output function in hidden layer.

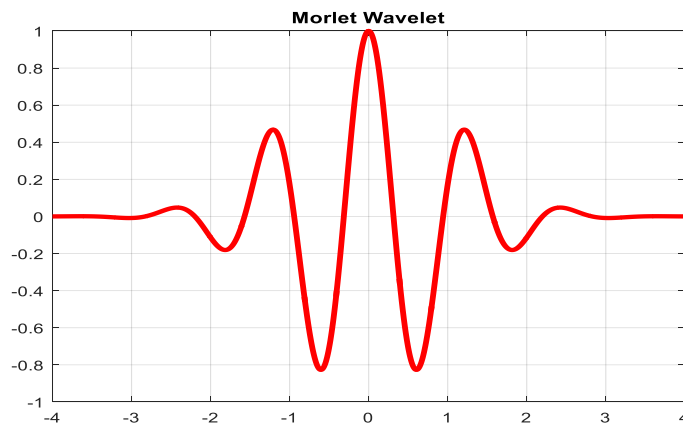


Figure 2. Wavelet Morelt Function [11]

Table 1. Parameters Indication

No.	Parameter's Name	WNN	NN
1	Number of neurons Input layer	3	3
2	Number of neurons Hidden layer	7	7
3	Number of neurons Output layer	1	1
4	Activation function in hidden layer	Morlet	sigmoid
5	Activation function in output layer	sigmoid	tansig
6	Learning rate value (η)	0.2	0.5
7	Momentum value (α)	0.99	0.99
8	iteration number	200 epoch	200 epoch
9	Training method	backpropagation	

Each network forms 200 epochs based on the data element network of Backpropagation's BP algorithm. Once the training is over, WNN is directly applied for the classification of prostate cancer risk. WNN performance can be estimated depending on the following indicators.

5.1. Convergence of Performances

Figures 3 and 4 show the MSE of WNN and NN, respectively. Figure 3 indicates that WNN has a lower MSE than NN. It is obviously demonstrated that WNN has a lower MSE value than neural network, which indicates the high performance of the classification, which means a high performance in the diagnosis of prostate cancer. As a result, WNN has a higher convergence rate compared to NN.

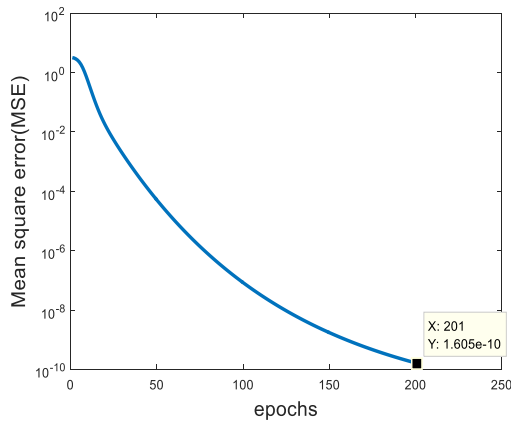


Figure 3. MSE for WNN

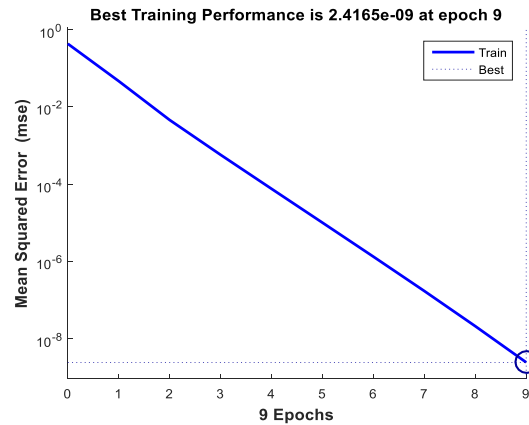


Figure 4. MSE for NN

5.2. Classification Experimental Results

Table 2 shows that the classification results of WNN outperform the classification results obtained by NN based on several cases of Prostate cancer patients. Furthermore, the results of the cases with ages above 50 and PSA over 4 ng/ml indicate a higher classification results of WNN compared to NN.

Table 2. Classification of WNN and NN

Name	Age	(PSA)	Volume (cm)	Output WNN. %	Output NN. %	Case
Case 1	60	1.06	8*6	3.6	6.6	normal
Case 2	84	0.39	10*7	30.61	45.61	normal
Case 3	72	95.11		93.69	90.69	Abnormal
Case 4	60	161.9		92.6	82.6	Abnormal
Case 5	70	11.3	(7-15)	99.99	93.99	Abnormal
Case 6	85	54.9	*	99.69	92.69	Abnormal
Case 7	50	7.48	(10-13)	24.6	24.6	Abnormal
Case 8	75	16.2		99.99	99.99	Abnormal
Case 9	67	14.11		99.88	96.88	Abnormal
Case 10	67	6.9		99.99	92.99	Abnormal
Case 11	76	100		92.6	90.6	Abnormal
Case 12	61	8.28		99.9	90.9	Abnormal

5.3. Computational complexity

Computational complexity is considered as a significant factor for evaluating the performance of the network. Table 3 illustrates, the CPU time utilization in the same processor during the training for each approach. The installed memory (RAM) has 3GB using Intel Core i3 M CPU 370 2.40 GHz processor. The time utilization results of the CPU indicate the speed efficiency of WNN than NN. Furthermore, it has been found that WNN has fewer operations than NN method. Therefore, WNN is not as complex as NN.

Table 3. CPU utilization time

Training Approach	WNN	NN
CPU time(sec)	10.190	18.22

6. CONCLUSION

In this article, WNN is used for the diagnosis of prostate cancer. The Morlet function is deployed in the hidden layer as an activation function while BP method is applied for training the network. The results of the simulation show that WNN performance works much better than neural networks in the classification of prostate cancer. The results show that WNN converges faster than NN, and the convergence rate of MSE is slower at $1.605e-10$. In addition, WNN introduces better classification values and exhibits less computational complexity than NN. However, compared to NN, the WNN needs more time to achieve the best value of MSE because it has a lower learning rate to obtain higher convergence rate.

REFERENCES

- [1] Wolf, A., et al., American Cancer Society guideline for the early detection of prostate cancer: update 2010. CA: a cancer journal for clinicians, 2010. 60(2): p. 70-98.
- [2] Van Cangh, P.J., et al., *Free to total prostate-specific antigen (PSA) ratio improves the discrimination between prostate cancer and benign prostatic hyperplasia (BPH) in the diagnostic gray zone of 1.8 to 10 ng/mL total PSA*. Urology, 1996. 48(6): p. 67-70.
- [3] Brawer, M.K., Prostate-specific antigen: Current status. CA: a cancer journal for clinicians, 1999. 49(5): p. 264-281.
- [4] Källén, H., et al. Towards grading gleason score using generically trained deep convolutional neural networks. in *Biomedical Imaging (ISBI)*, 2016 IEEE 13th International Symposium on. 2016. IEEE.
- [5] Gorelick, L., et al., Prostate histopathology: Learning tissue component histograms for cancer detection and classification. *IEEE transactions on medical imaging*, 2013. 32(10): p. 1804-1818.
- [6] Doyle, S., et al., A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE transactions on biomedical engineering*, 2012. 59(5): p. 1205-1218.
- [7] Mattfeldt, T., et al., Classification of prostatic carcinoma with artificial neural networks using comparative genomic hybridization and quantitative stereological data. *Pathology-Research and Practice*, 2003. 199(12): p. 773-784.
- [8] Litjens, G., et al., *Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis*. Scientific reports, 2016. 6: p. 26286.
- [9] Zhang, Q. and A. Benveniste, Wavelet networks. *IEEE transactions on Neural Networks*, 1992. 3(6): p. 889-898.
- [10] Chen, Z., T. Feng, and Q. Meng. *The application of wavelet neural network in time series prediction and system modeling based on multiresolution learning*. in Systems, Man, and Cybernetics, 1999. IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on. 1999. IEEE.
- [11] Ghawbar, F.M., et al., *PAT and Px Code Sidelobe Reduction Using Wavelet Neural Network*, in Advances in Machine Learning and Signal Processing. 2016, Springer. p. 117-128.
- [12] Ahmed, M.S., N.S.M. Shah, and S.I. Anka, Sidelobe reduction using wavelet neural network for binary coded pulse compression. 2016.
- [13] Thompson, I.M., et al., Prevalence of prostate cancer among men with a prostate-specific antigen level ≤ 4.0 ng per milliliter. *New England Journal of Medicine*, 2004. 350(22): p. 2239-2246.
- [14] Hu, X., et al., Artificial neural networks and prostate cancer—tools for diagnosis and management. *Nature Reviews Urology*, 2013. 10(3): p. 174.
- [15] Alexandridis, A.K. and A.D. Zaprani, Wavelet neural networks: with applications in financial engineering, chaos, and classification. 2014: John Wiley & Sons.
- [16] Zhao, J., W. Chen, and J. Luo. *Feedforward wavelet neural network and multi-variable functional approximation*. in International Conference on Computational and Information Science. 2004. Springer.