

Wavelet Neural Network Initialization Using LTS for DNA Sequence Classification

Abdesselem Dakhli¹(✉), Wajdi Bellil², and Chokri Ben Amar²

¹ Department of Computer Science, REGIM,
University of Gabes, 6002 Gabes, Tunisia
abdesselemdakhli@gmail.com

² Department of Computer Science, REGIM,
University of Sfax, 3018 Sfax, Tunisia
{wajdi.bellil,chokri.benamar}@ieee.org

Abstract. In this paper, we present a new approach for DNA sequence classification. The proposed approach is based on using the Wavelet Neural Network (WNN) and the k-means algorithm. The satisfying performance of the Wavelet Neural Networks (WNN) depends on an appropriate determination of the WNN structure. Our approach uses the Least Trimmed Square (LTS) and the Gradient Algorithm (GA) to solve the architecture of the WNN. The initialization of the Wavelet Neural Network is solved by using the Least Trimmed Square (LTS) method, which is applied for selecting the wavelet candidates from the Multi Library of the Wavelet Neural Networks (MLWNN) for constructing the WNN. Besides, the Gradient Algorithm (GA) is implemented for training the WNN in our method. The GA is used to solve the structure and learning of the WNN. This algorithm is applied to adjust the parameters of WNN. The performance of the WNN is investigated by detecting the simulating and real signals in white noise. The proposed method has been able to optimize the wavelet neural network and classify the DNA sequences. In this study, the LTS model is compared to the two initialization algorithms: Residual Based Regressor Selection (RBRS) and Stepwise Regressor Selection by Orthogonalization (SRSO). The LTS algorithm is to find the regressors, which provide the most significant contribution to the approximation of error reduction. The advantage of the LTS algorithm is to select the candidate wavelet from the MLWNN. This wavelet can reduce the approximation error. Our aim is to construct classifier method that gives highly accurate results. This classifier permits to classify the DNA sequence of organisms. The classification results are compared to other classifiers. The experimental results have shown that the WNN-LTS model outperformed the other classifier in terms of both the running time and clustering. In this paper, our system consists of three phases. The first one, which is called transformation, is composed of three sub steps; binary codification of DNA sequences, Fourier Transform and Power Spectrum Signal Processing. The second section is the approximation; it is empowered by the use of Multi Library Wavelet Neural Networks (MLWNN). Finally, the third section, which is called the classification of the DNA sequences.

Keywords: LTS · Wavelet neural networks · DNA sequences · MLWNN · RBRS · SRSO

1 Introduction

The construction of the neural networks structure suffers from some deficiencies: the local minima, the lack of efficient constructive methods, and the convergent efficiency, when using ANNs. As a result, the researchers discovered that the WNN, is a new class of neural networks which joins the wavelet transform approach. The WNN were presented by Benveniste and Zhang. This approach is used to approximate the complex functions with a high rate of convergence [1]. This model has recently attracted extensive attention for its ability to effectively identify nonlinear dynamic systems with incomplete information [1–5]. The satisfying performance of the WNN depends on an appropriate determination of the WNN structure. To solve this task many methods are proposed to optimize the WNN parameters. These methods are applied for training the WNN such as the least-square which is used to train the WNN when outliers are present. These training methods are applied to reduce some function costs and improve performed the approximation quality of the wavelet neural network. On the other hand, the WNN has often been used on a small dimension [6]. The reason is that the complexity of the network structure will exponentially increase with the input dimension. The WNN structure has been studied by several researchers. Moreover, the research effort has been made to deal with this problem over the last decades [6–9]. The application of WNN is usually limited to problem of small dimension. The number of wavelet functions in hidden layer increases with the dimension. Therefore, building and saving WNN of large dimension are of prohibitive cost. Many methods are used to reduce the size of the wavelet neural networks to solve large dimensional task. Such as, magnitude approach is applied to eliminate of wavelets function with small coefficients [3]. The method, referred as Matching Pursuit (MP), was first introduced by Mallat. This method is used to determine a good wavelet basis in a dictionary [10]. Following this line of works, the Residual Based Regressor Selection (RBRS) algorithm is proposed for the synthesis of WNN. The Stepwise Regressor Selection by Orthogonalization (SRSO) and the Orthogonal Least Square (OLS) suggested are both the popular approaches [8]. These methods are used to reduce the complexity of the selected subset models. The number of regressors is smaller than that of the inputs [27, 28]. In addition, Bellil et al. applied a new initialization by selecting the method of the library WNN training. This approach is used to approximate a small number of inputs [22]. In this study, we use the Least Trimmed Square (LTS) method to select a little subset of wavelet candidates from MLWNN constructing the WNN structure in order to build a method to classify a collection containing a dataset of DNA sequences. This method is used to optimize an important number of inputs of DNA sequences. The Beta wavelet function is used to build the WNN. This wavelet makes the WNN training very efficient a reason of adjustable parameters of this function. Various approaches are used for clustering the DNA sequences such as the WNN, which is applied to construct a classification system. Wu et al. usedii an artificial neural network to classify the DNA sequences [11]. Moreover, Jach and Marin are proposed a method to classify the mitochondrial DNA Sequences. This approach joins the WNN and a Self-organizing map method. The feature vector sequences constructed by using the WNN [12]. Yang et al. used a Wavelet packet analysis to extract features of DNA sequences, which are

applied to recognize the types of other sequences [13]. Wu et al. applied the neural network to classify the nucleic acid sequence. This classifier used three-layer and feed-forward networks that employ back-propagation learning algorithm [14]. Since a DNA sequence can be converted into a sequence of digital signals, the feature vector can be built in time or frequency domains. However, most traditional methods, such as k-tuple and DMK,... models build their feature vectors only in the time domain, i.e., they use direct word sequences. This paper contains three sections: in Sect. 2, we present our proposed approach. Section 3 presents the wavelet theory used to construct the WNN of our method. Section 4 shows the simulation results of our approach and Sect. 5 ends up with a conclusion.

2 Methods

This paper presents a new approach based on the wavelet neural network, which is constructed by using the Multi-library Wavelet Neural Networks (MLWNN). The WNN structure is solved by using the LTS method. Our approach is divided into two stages: approximation of the input signal sequence and clustering of feature extraction of the DNA sequences using the WNN and k-means algorithm.

2.1 Conversion of the DNA Sequence into a Genomic Signal

The species are classified in class by using the DNA sequence, which is composed of four basic nucleotides A(Adenine), G(Guanine), C(Cytosine) and T(Thymine), where each organism is identified by its DNA sequence [21, 22]. The feature extraction vector presents the DNA sequence. This vector is used to classify the species [23]. In this study, this vector coded by using a digital format, which can be used for DNA signal spectrum indicates 1 or 0 for the existence or not of a specific nucleotide at the DNA sequence level [24, 25]. For example, if $x[n] = [A A A T \dots]$, we obtain: $x[n] = [1000 1000 1000 0001 \dots]$

The indicator sequence is manipulated with mathematical methods. The sequence of complex numbers, called $f(x)$ (1), is obtained by using the discrete Fourier Transform:

$$f(x) = \sum_{n=0}^{N-1} X_e(n) e^{-j\pi nk/N}, k = 0, 1, 2, \dots, N-1 \quad (1)$$

The Power Spectrum is applied to compute the $Se[k]$ (2) for frequencies $k = 0, 1, 2, \dots, N-1$ is defined as,

$$Se[k] = |f(x)|^2 \quad (2)$$

2.2 Wavelet Neural Network

The wavelet neural network is defined by the combination of the wavelet transform and the artificial neuron networks [33, 34]. It is composed of three layers. The salaries of the weighted outputs are added. Each neuron is connected to the other following layer.

The WNN (Fig. 1) is defined by pondering a set of wavelets dilated and translated from one wavelet candidate with weight values to approximate a given signal f . The response of the WNN is:

$$\hat{y} = \sum_{i=1}^{N_w} w_i \Psi\left(\frac{x - b_i}{a_i}\right) + \sum_{k=0}^{N_i} a_k x_k \tag{3}$$

where $(x_1, x_2, \dots, x_{N_i})$ is the vector of the input, N_w is the number of wavelets and y is the output of the network. The output can have a component refine in relation to the variables of coefficients a_k ($k = 0, 1 \dots N_i$) (Fig. 1). The wavelet mother is selected from the MLWNN, which is defined by dilation (a_i) which controls the scaling parameter and translation (b_i) which controls the position of a single function ($\Psi(x)$). A WNN is used to approximate an unknown function:

$$y = f(x) + \varepsilon, \tag{4}$$

Where f is the regression function and ε is the error term.

2.3 Multi Library Wavelet Neural Network (MLWNN)

Many methods are used to construct the Wavelet Neural Network. Zhang applied two stages to construct the Wavelet neural Network [2, 3]. First, the discretely dilated and translated version of the wavelet mother function Ψ is used to build the MLWNN.

$$W = \left\{ \psi_i : \psi_i(x) = \alpha_i \psi\left(\frac{(x_k - b_i)}{a_i}\right), \alpha_i = \left(\sum_{k=1}^n \left[\psi\left(\frac{(x_k - b_i)}{a_i}\right)\right]^2\right)^{\frac{1}{2}}, i = 1, \dots, L \right\}, \tag{5}$$

Where L is the number of wavelets in W and x_k is the sampled input. Then the best M wavelet mother function is selected based on the training sets from the wavelet library W , in order to construct the regression:

$$f_M(x) = \hat{y} = \sum_{i \in I} w_i \psi_i(x), \tag{6}$$

Where $M \leq L$ and I is a subset wavelet from the wavelet library. Secondly, the minimized cost function:

$$j(I) = \min_{w_i, i \in I} \frac{1}{n} \sum_{k=1}^n \left(y_k - \sum_{i \in I} w_i \psi_i(x_k) \right)^2, \tag{7}$$

The stepwise selection by orthogonalization and the backward elimination are used by Zhang; the first is applied to select the appropriate wavelet in the hidden units while the second is used to choose the number of the hidden units, and of wavelets M , which

are selected as the minimum of the so-called Akaike’s final prediction error criterion (FPE) [2, 3]:

$$j_{FPE}(\hat{f}) = \frac{1 + n_{pa}/n}{1 - n_{pa}/n} \frac{1}{2n} \sum_{k=1}^n (\hat{f}(x_k) - y_k)^2 \tag{8}$$

Where the estimator have n_{pa} parameters.

The gradient algorithms used to train the WNN, like least mean squares to reduce the mean-squared error:

$$j(w) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}(w))^2, \tag{9}$$

Where $j(w)$ is the output of the Wavelet neural networks. The time-frequency locality property of the wavelet is used to give a signal f , a candidate library w of wavelet basis can be constructed.

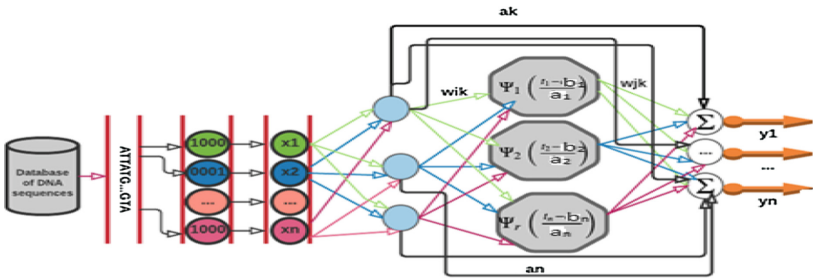


Fig. 1. The three layer wavelet network

2.4 Wavelet Network Construction Using the LTS Method

The set of training data $TN = \{x_1, x_2, \dots, x_k, f(x_k)\}_{k=1}^N$ is used to adjust the weights and the WNN parameters, and the output of the three layers of the WNN in Fig. 2 can be expressed via (7). The model selection is used to select the wavelet candidates from the Multi Library Wavelet Neural Networks (MLWNN). These wavelet mothers are used to construct the wavelet neural network structure. In this study, the Least Trimmed Squares estimator (LTS) is proposed to select a little subset of wavelet candidates from the MLWNN. These wavelet candidates are applied to construct the hidden layer of the WNN [35]. Furthermore, the Gradient Algorithm is proposed to optimize the wavelet neural networks parameter. The residual (or error) e_i at the i th output of the WNN due to the i th example is defined by:

$$e_i = y_i - \hat{y}_i, \quad i \in n \tag{10}$$

The Least Trimmed Square estimator is used to select the WNN weights that minimize the total sum of trimmed squared errors:

$$E_{total} = \frac{1}{2} \sum_{k=1}^p \sum_{i=1}^l e_{ik}^2 \tag{11}$$

The Gradient Algorithm used to optimize the parameters (a_i, b_i, w_i) of the WNN is as follows:

$$\frac{\partial c}{\partial \hat{y}} = e_k \tag{12}$$

$$\frac{\partial c}{\partial w_i} = e_k \psi \left(\frac{x_i - b_i}{a_i} \right) \tag{13}$$

$$\frac{\partial c}{\partial b_i} = -e_k w_i \frac{1}{a_i} \psi' \left(\frac{x_i - b_i}{a_i} \right) \tag{14}$$

$$\frac{\partial c}{\partial a_i} = -e_k w_i \left(\frac{x_k - b_i}{a_i^2} \right) \psi' \left(\frac{x_k - b_i}{a_i} \right) \tag{15}$$

where

$$e = \hat{y}(x_k) - f(x_k) \tag{16}$$

2.5 Approximation of DNA Sequence Signal

The classification of DNA sequences is an NP-complete problem; the alignment is outside the range of two sequence of DNA, the problem rapidly becomes very complex because the space of alignment becomes very important. The recent advance of the sequence technology has brought about a consequent number of DNA sequences that can be analyzed. This analysis is used to determine the structure of the sequences in homogeneous groups using a criterion to be determined. In this paper, the Power Spectrum is used to process the signal of the DNA sequence. These signals are used by the wavelet neural networks (WNN) to extract the signatures of DNA sequences, which are used to match the DNA test with all the sequences in the training set [29]. Initially, the signatures of DNA sequences developed by the 1D wavelet network during the learning stage gave the wavelet coefficients which are used to adapt the DNA sequences test with all the sequences in the training set. Then, the DNA test sequence is transmitted onto the wavelet neural networks of the learning DNA sequences and the coefficients specific to this sequence are computed. Finally, the coefficients of the learning DNA sequences compared to the coefficients of the DNA test sequences by computing the Correlation Coefficient. In this stage, the k-means clustering is used to classify the signatures of the DNA sequences [27].

2.6 Learning Wavelet Network Using Gradient Algorithm and LTS Methods

In this section, we show how the library wavelet is used to learn a wavelet neural network [15, 16, 26, 27].

- **Learning approach**

Step 1: The data set of DNA sequence is divided into two groups: training and testing dataset. These groups are applied to train and test the wavelet neural network.

Step 2: Conversion of DNA sequence to a genomic signal using a binary indicator and Power Spectrum Signal Processing

Step 3: the discretely dilated and translated version are used to construct the library W . The training data are proposed to create this library wavelet, apply the Least Trimmed Square (LTS) algorithm to select the optimal mother wavelet function (10), (11) and choose, from the library, the N wavelet candidate that best matches an output vector.

Step 3.1: Initializing of the mother wavelet function library

Step 3.2: Randomly initialize w_{jk} and v_{ij} .

Step 3.3: For $k = 1, \dots, m$

- Calculate the predicted output \hat{y}_i via (3).
- Compute the residuals $e_{ik} = y_i - \hat{y}_i$ via (10).
- the algorithm is stopped when the criteria diverged, then stop; otherwise, go to the next step
- Find the arranged values $e_{ik}^2 \leq \dots \leq e_{im}^2$. Choosing the N best mother wavelet function to initialize the WNN.

Step 4: The values of w_{ij}^{opt} , a_i^{opt} and b_i^{opt} are computed using the Gradient algorithm via (13), (15), (14) and (16) go to step 3.3.

- **Clustering using K-means**

Step 1: Generate a matrix $M_signature$ of DNA sequences ($(w_{ij}^{opt}, a_i^{opt} \text{ and } b_i^{opt})$).

Step 2: Let $M_signature$ $s_i = \{w_{ij}^{opt}, a_i^{opt}, b_i^{opt}\}$ be the set of data points and $V = \{v_1, v_2, \dots, v_c\}$ be the set of centers.

Step 3: choose the Number of groups (k)

Step 4: Assume the k training instance of signature of DNA as lonely-unit groups.

Step 5: Affect each of $(n-k)$ training instance of signature of DNA to the group with the proximate centroid and Recalculate the centroid of the winning group.

Step 6: Compute the distance of each signature of DNA from the centroid of each group, switch the instance if it is not in the group and update the centroid of the winning and losing group.

Step 7: Go to Step 6 if the convergence is not achieved

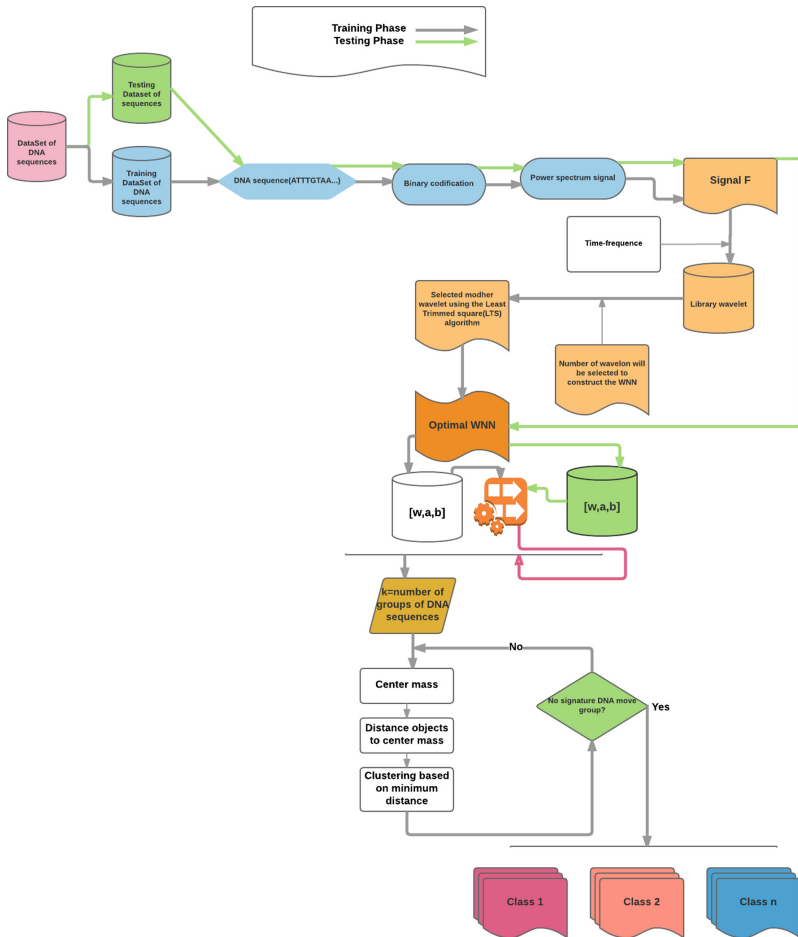


Fig. 2. Proposed approach

3 The Bêta Wavelet Family

The function beta is defined by $\beta(x) = \beta_{x_0, x_1, p, q}(x)$ [21, 22, 32, 33], x_0 and x_1 are real parameters.

$$\beta(x, p, q, x_0, x_1) = \begin{cases} \left(\frac{x - x_0}{x_c - x_0}\right)^p \left(\frac{x - x_1}{x_1 - x_c}\right)^q & \text{if } x \in [x_0, x_1] \\ 0 & \text{otherwise} \end{cases} \quad (17)$$

The derivatives of this function $\in L2(\mathfrak{R})$ are of class C^∞ . The general form of the nth derivative of Beta function is:

$$\psi_n(x) = \frac{d^n \beta(x)}{dx^n} = \left[(-1)^n \frac{n!p}{(x-x_0)^{n+1}} + \frac{n!q}{(x_1-x)^{n+1}} \right] \beta(x) + P_n(x) P_1(x) \beta(x) + \sum_{i=1}^n C_n^i \left[(-1)^n \frac{(n-i)!p}{(x-x_0)^{n+1-i}} + \frac{(n-i)!q}{(x_1-x)^{n+1-i}} \right] \times P_1(x) \beta(x) \quad (18)$$

where:

$$P_1(x) = \frac{P}{x-x_0} - \frac{q}{x_1-x} \quad (19)$$

$$P_n(x) = (-1)^n \frac{n!p}{(x-x_0)^{n+1}} - \frac{n!q}{(x_1-x)^{n+1}} \quad (20)$$

$$x_c = \frac{(px_1 + qx_0)}{(p+q)} \quad (21)$$

4 Results and Discussion

This paper used three datasets HOG100, HOG200, and HOG300 selected from microbial organisms [23]. In this study, different experiments are used to evaluate the performance of our approach. The data set of DNA sequences are divided into test and train data. The published empirical and synthetic datasets are selected to perform the clustering comparative analysis [23] (Table 1).

Table 1. Distribution of available data into training and testing set of DNA sequence

Dataset	Total	Training	Test
HOG100	500	300	200
HOG200	600	400	200
HOG300	700	600	100

4.1 Selecting the Mother Wavelet

To evaluate the performance of our method, the NSRMSE and the training time serve. The LTS has a better performance to select the wavelet candidates from the Multi Library Wavelet Neural Networks (MLWNN). In the beginning, during the phase of approximation our approach tried to decompose the input signal for every DNA sequence and in the end, it tried to reconstruct the input signal of DNA. The estimation of the performance of this phase is measured by the NSRMSE. (Table 4) shows that

obtained NSRMSE are low (0.000869) and the run time increases relatively with size of the DNA sequence. The result shows that the size of DNA sequence increases the time of the training phase. This time depends in the length of a DNA sequence. The training time increases (12.213 s) when the size is equal to 700. The complexity of the WNN structure increases exponentially with the input dimension. To solve the approximation problem we applied the library wavelet which incorporates a family wavelet. This library is called the Multi Library Wavelet Neural Network Model (MLWNN) (Tables 2 and 3).

Table 2. Selected mother wavelets and normalized square root of mean square error using residual based regressor selection (RBRS)

Dataset	Size	Beta 1	Beta 2	Beta 3	Mexhat4	Slog5	NSRMSE	Training time
HOG100	500	0	1	3	2	2	3.68777	43.337
HOG200	600	1	2	2	3	1	3.7984	47.654
HOG300	700	2	1	3	1	1	3.9523	52.654

Table 3. Selected mother wavelets and normalized root mean square using stepwise regressor selection by orthogonalization (SRSO)

Dataset	Size	Beta 1	Beta 2	Beta 3	Mexhat4	Slog5	NSRMSE	Training time
HOG100	500	2	2	1	1	2	0.675238	13.584
HOG200	600	0	3	2	1	2	0.78235	16.213
HOG300	700	1	1	2	2	2	0.79856	17.654

Table 4. Selected mother wavelets and normalized root mean square using the least trimmed square (LTS).

Dataset	Size	Beta 1	Beta 2	Beta 3	Mexhat4	Slog5	NSRMSE	Training time
HOG100	500	2	2	1	3	1	0.000869	9.869
HOG200	600	3	1	2	1	2	0.00543	10.542
HOG300	700	2	1	1	1	3	0.00896	12.213

4.2 Classification Results

Experiment results were performed to prove the effectiveness of our proposed approach. Evaluation metrics namely Precision, Recall and F-measures are used to compare our approach with other competitive methods. The F-measure combines the precision and the recall metric. We then calculate the recall and the precision of that cluster for each given class. More specifically, for cluster j and class i . The F-measure of the class j and the group i is the given by

$$F(j, i) = (2 * Recall(i, j) * Precision(i, j)) / ((Precision(i, j) + Recall(i, j))) \quad (22)$$

Table 5. The classification results of WNN- LTS (Our Method) and other models (WFV, K-tuple, DMK) on different datasets of DNA sequences

Dataset	Our method		WFV		K-tuple		DMK	
	F-Score	# class	F-Score	#class	F-Score	# class	F-Score	# class
HOG100	0.68	330	0.58	854	0.52	451	0.55	658
HOG200	0.7	468	0.57	845	0.53	566	0.55	754
HOG300	0.78	232	0.58	185	0.52	654	0.54	256

Table 5 shows that WNN-LTS (our method) is better than other models(WFV, K-tuple and DMK) in terms of the classification results and optimal settings. The number of classes obtained by our approach is little less than in the other methods. The F-score (F-measure) proves the efficiency of our method. The F-measure is increased using WNN and LTS method. The LTS is applied to optimize the WNN structure.

4.3 Running Time

Tables 5 and 6 show that the WNN can produce very good the prediction accuracy. The results of our approach WNN-LTS tested on datasets show that accuracy outperforms the other techniques in terms of percentage of the correct species identification. Tables 5 and 6 show the distribution of the good classifications by class as well as the rate of global classification for all the DNA sequences of the validation phase. The WNN-LTS (our approach) is faster than the other methods. This speed is due to the use of the Least Trimmed Square (LTS) algorithm.

Table 6. Running time in seconds of each method on all datasets.

Dataset	Model	Length of feature vector	Time of building feature vector	Time of k-means	Total running time
HOG100	Our method	128	5.3569	75.645	81.0019
	WFV	32	8.4857	102.2634	110.7491
	K-tuple	64	7.9544	763.5223	771.4767
HOG200	Our method	128	16.758	350.664	367.422
	WFV	32	20.5239	645.8376	666.3615
	K-tuple	64	19.6306	3010.5426	3030.1732

5 Conclusions

In this study, we have used the LTS method to select a subset of wavelet function from the Library Wavelet Neural Network Model. This subset wavelet is applied to build Wavelet Neural Network (WNN). The WNN is used to approximate function $f(x)$ of a DNA sequence signal. Firstly, the binary codification and Power Spectrum are used to process the DNA sequence signal. Secondly, the Library Wavelet is constructed. The LTS method is used to select the best wavelet from library. These wavelets are

applied to construct the WNN. Thirdly, the k-means classification is used to classify the similar DNA sequences according to some criteria. This clustering aims at distributing DNA sequences characterized by p variables X_1, X_2, \dots, X_p in a number m of subgroups which are homogeneous as much as possible while every group is well differentiated from the others. The proposed approach helps to classify DNA sequences of organisms into many classes. These clusters can be used to extract significant biological knowledge.

Acknowledgment. We would like to acknowledge the financial support, under the form of grant, from the General Direction of Scientific Research (DGRST), Tunisia.

References

1. Zhang, Q., Benveniste, A.: Wavelet networks. *IEEE Trans. Neural Netw.* **3**(6), 889–898 (1992)
2. Zhang, J., Walter, G., Miao, Y., et al.: Wavelet neural networks for function learning. *IEEE Trans. Signal Process.* **43**(6), 1485–1497 (1995)
3. Zhang, Q.: Using wavelet network in nonparametric estimation. *IEEE Trans. Signal Process.* **8**, 227–236 (1997)
4. Pati, Y.C., Krishnaprasad, P.S.: Analysis and synthesis of feed-forward neural networks using discrete affine wavelet transformations. *IEEE Trans. Neural Netw.* **4**, 73–85 (1993)
5. Billings, S.A., Wei, H.L.: A new class of wavelet networks for nonlinear system identification. *IEEE Trans. Neural Netw.* **16**, 862–874 (2005)
6. Xu, J.H., Ho, D.W.C.: A basis selection algorithm for wavelet neural networks. *Neurocomputing* **48**, 681–689 (2002)
7. Mallat, S.G., Zhifeng, Z.: Matching pursuits with time-frequency dictionaries. *IEEE Trans. Signal Process.* **41**, 3397–3415 (1993)
8. Chen, S., Wigger, J.: Fast orthogonal least squares algorithm for efficient subset model selection. *IEEE Trans. Signal Process.* **43**, 1713–1715 (1995)
9. Han, M., Yin, J.: The hidden neurons selection of the wavelet networks using support vector machines and ridge regression. *Neurocomputing* **72**, 471–479 (2008)
10. Wu, C.H.: Artificial neural networks for molecular sequence analysis. *Comput. Chem.* **21**(4), 231–256 (1997)
11. Jach, A.E., Marín, J.M.: Classification of genomic sequences via wavelet variance and a self-organizing map with an application to mitochondrial DNA. *Stat. Appl. Genet. Mol. Biol.* **9**, 1544–6115 (2010)
12. Zhao, J., Yang, X.W., Li, J.P., Tang, Y.Y.: DNA sequences classification based on wavelet packet analysis. In: Tang, Y.T., Wickershauser, V., Yuen, P.C., Li, C.-H. (eds.) WAA 2001. LNCS, vol. 2251, pp. 424–429. Springer, Heidelberg (2001)
13. Wu, C., Berry, M., Fung, Y.-S., McLarty, J.: Neural networks for molecular sequence classification. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **1**, 429–437 (1993)
14. Vinga, S., Almeida, J.: Alignment-free sequence comparison—a review. *Bioinformatics* **19**, 513–523 (2003)
15. Wei, D., Jiang, Q. A.: DNA sequence distance measure approach for phylogenetic tree construction. In: Proceedings of the IEEE Fifth International Conference on Bio-Inspired Computing: Theories and Applications (BIC-TA), pp. 204–212. IEEE (2010)
16. Shi, L., Huang, H.: DNA sequences analysis based on classifications of nucleotide bases. In: Luo, J. (ed.) *Affective Computing and Intelligent Interaction*. AISC, vol. 137, pp. 379–384. Springer, Heidelberg (2012)

17. Bauer, M., Schuster, S.M., Sayood, K.: The average mutual information profile as a genomic signature. *BMC Bioinf.* **9**, 48 (2008)
18. Qi, J., Wang, B., Hao, B.I.: Whole proteome prokaryote phylogeny without sequence alignment: a K-string composition approach. *J. Mol. Evol.* **58**, 1–11 (2004)
19. Carter, O.B., et al.: Alignment-free genetic sequence comparisons: a review of recent approaches by word analysis. *Briefings Bioinf.* **15**(6), 890–905 (2013)
20. Bao, J.P., Yuan, R.Y.: A wavelet-based feature vector model for DNA clustering. *Genet. Mol. Res.* **14**, 19163–19172 (2015)
21. Ben Amar, C., Bellil, W., Adel Alimi, M.: Beta function and its derivatives: a new wavelet family. *Trans. Syst. Signals Devices* **1**, 275–293 (2006)
22. Bellil, W., Othmani, M., Ben Amar, C.: Initiazation by selection for multi-library wavelet neural network training. In: Conference: Artificial Neural Networks and Intelligent Information Processing (ANNIIP), Angers, France (2007)
23. <http://doua.prabi.fr/databases/hogenom/>
24. Mejdoub, M., Ben Amar, C.: Classification improvement of local feature vectors over the KNN algorithm. *Multimedia Tools Appl.* **64**(1), 197–218 (2013)
25. Zaied, M., Said, S., Jemai, O., Amar, C.: A novel approach for face recognition based on fast learning algorithm and wavelet network theory. *Int. J. Wavelets Multiresolut. Inf. Process.* World Sci. **19**, 923–945 (2011)
26. Said, S., Ben Amor, B., Zaied, M., Ben Amar, C., Daoudi, M.: Fast and efficient 3D face recognition using wavelet networks. In: 16th IEEE International Conference on Image Processing, Cairo, Egypt, pp. 4153–4156 (2009)
27. Jemai, O., Zaied, M., Ben Amar, C.: Fast learning algorithm of wavelet network based on fast wavelet transform. *Int. J. Pattern Recognit. Artif. Intell.* **25**(8), 1297–1319 (2011)
28. Jemai, O., Zaied, M., Ben Amar, C., Alimi, A.: Pyramidal hybrid approach: wavelet network with OLS algorithm- based image classification. *Int. J. Wavelets Multiresolut. Inf. Process.* **9**, 111–130 (2011). World Scientific Publishing Company
29. Ejbali, R., Benayed, Y., Zaied, M., Alimi, A.: Wavelet networks for phonemes recognition. In: International Conference on Systems and Information Processing (2009)
30. Ejbali, R., Zaied, M., Ben Amar, C.: Multi-input multi-output beta wavelet network modeling of acoustic units for speech recognition. In: International Journal of Advanced Computer Science and Applications (IJACSA). The Science and Information Organization (SAI), vol. 3 (2012)
31. Ejbali, R., Zaied, M., Ben Amar, C.: Wavelet network for recognition system of Arabic word. *Int. J. Speech Technol.* **13**, 163–174 (2010). Springer edition
32. Bouchrika, T., Zaied, M., Jemai, O., Ben Amar, C.: Ordering computers by hand gestures recognition based on wavelet networks. In: International Conference on Communications, Computing and Control Applications (CCCA), Marseilles, France, pp. 36–41 (2012)
33. Mejdoub, M., Fonteles, L., Ben Amar, C., Antonini, M.: Embedded lattices tree: an efficient indexing scheme for content based retrieval on image databases. *J. Vis. Commun. Image Represent.* **20**(2), 145–156 (2009)
34. Dammak, M., Mejdoub, M., Zaied, M., Amar, C.B.: Feature vector approximation based on wavelet network. In: ICAART 2012 - Proceedings of the 4th International Conference on Agents and Artificial Intelligence, vol. 1, pp. 394–399 (2012)
35. Ben Aoun, N., Mejdoub, M., Ben Amar, C.: Graph-based approach for human action recognition using spatio-temporal features. *J. Vis. Commun. Image Represent.* **25**(2), 329–338 (2014)