Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 12-17, 2007

Breast Tissue Classification Based on Unbiased Linear Fusion of Neural Networks with Normalized Weighted Average Algorithm

Yunfeng Wu and S. C. Ng

Abstract— **The diagnosis of breast cancer is performed based on informed interpretation of representative histological tissue sections. Tissue distribution detected from cytologic examinations is useful for tumor staging and appropriate treatment. In this paper, we propose a normalized weighted average (Normwave) algorithm for the unbiased linear fusion, and also construct the multiple classifier system that includes a group of Radial Basis Function (RBF) neural classifiers for the classification of breast tissue samples. The empirical results show that the proposed Normwave algorithm may improve the performance of the RBF-based multiple classifier system, and also reliably outperforms some widely used fusion methods, in particular the simple average and adaptive mixture of experts.**

I. Introduction

Cancers of the breast, lung and bronchus, and colon and rectum, are top three most commonly diagnosed types of cancer among women in North America [1], [2]. Recent statistics [1] indicate that breast cancer is estimated to account for 26% of all new cancer cases among women in the United States in 2007. Early detection of primary tumors by taking one or more breast exams such as palpation, mammography, ultrasonography, and cytologic examinations, is important for an improved prognosis in the treatment that can help reduce the mortality rates [3], [4]. Screening mammography has been shown to be effective in detection of nonpalpable breast tumors at an early stage [4], [5], which helps a more accurate diagnostic decision by radiologists or medical specialists. Although mammography is commonly used in the diagnosis of malignant carcinoma, it is not conclusive in equivocal cases, and the result of modality study requires confirmation with the histological examination of relevant tissues [6]. The distribution of tissues detected from cytologic examinations can provide important information about the elastic characteristics of breast which would be incorporated in building the breast deformation model [7].

Recently, more and more computational systems and machine learning algorithms have been utilized in computer-aided diagnosis [8]. Artificial neural networks, with the advantages of experience-based learning and generalization ability [9], are also popular in a number of biomedical applications, such as the characterization of breast abnormalities [10], classification of microcalcifications [11] and masses [12], and early diagnosis of breast cancer [13].

Yunfeng Wu is with the School of Information Engineering, Beijing University of Posts and Telecommunications, No. 10 Xi Tu Cheng Road, Haidian District, Beijing 100876, China (e-mail: y.wu@ieee.org).

Dr. S. C. Ng is with the School of Science and Technology, The Open University of Hong Kong, 30 Good Shepherd Street, Homantin, Kowloon, Hong Kong (e-mail: scng@ouhk.edu.hk).

However, when given a noisy data set from the real world, artificial neural networks might produce multifarious generalizations by determining different decision boundaries [14]. In order to ameliorate accuracy of an individual neural classifier, the ensemble methods were proposed [15]–[19]. The scheme of an ensemble is to make a finite number of component neural networks work collectively, and then combine their knowledge to produce a consensus decision. The applications of neural network ensembles in biomedicine are rich in the literature, e.g., the spectroscopic detection of cervical pre-cancer [20], diagnosis of breast cancer [11], [21], [22], and prediction of protein subcellular locations [23]–[25].

According to Valentini and Masulli [26], the ensembles of neural networks can be categorized into two main styles: nongenerative and generative methods. The nongenerative ensembles confine themselves to combining a set of well-devised component learners by means of fusion strategies, e.g., the Simple Average (SA) [27], Adaptive Mixture of Experts (AME) [28], Majority Vote (MV) [29], Weighted Average (WA) [30], Perceptron Average (PA) [22], and Linear Least-Squares Fusion (LLSF) [23]. The MV and SA rules set the fixed fusion coefficients, whereas the WA, PA, and LLSF require a training process to initialize and update their fusion coefficients for an adaptation of component learners. On the other hand, the generative methods will produce a series of component neural learners by resampling the original data under a certain distribution. The representative algorithms are AdaBoost [31] and Bagging [32]. AdaBoost was proposed by Freund *et al.* [31] with the aim to find a final error-free mapping function according to the given probability distribution over the training data. Bagging introduces the bootstrap resampling procedure [33] into ensemble systems so that the final bias converges by average while the variance gets much smaller than that of each component learner. In this paper, we propose a novel ensemble algorithm for the unbiased linear fusion of neural networks, and only focus our study on the nongenerative ensembles.

The rest of this paper is organized as follows. Section II presents the detailed information of breast tissue data used in our experiments. Section III provides a detailed description of the multiple classifier system for breast tissue classification. Section IV presents the empirical results of three fusion methods: the SA, AME, and the proposed normalized weighted average fusion. The classification performance of the multiple classifier system studied is also compared with that of its component neural classifiers. Section V concludes our study as well as the directions for future work.

II. Data Description and Notation

The data set used in our experiments was provided by da Silva *et al.* [34], and it contains the electrical impedance measurements performed on 106 freshly excised breast tissue samples. The impedance measurements were taken at seven frequencies and plotted in the Argand plane. A total of nine features computed from the impedance spectrum are: impedance at zero frequency (I0); phase angle at 500 kHz; high-frequency slope of the phase angle; impedance distance between spectral ends (DA); area under the spectrum; area normalized by DA; maximum amplitude of the spectrum; distance between I0 and real part of the maximum frequency point; length of the spectral curve. The task is to categorize the breast tissues present in the data set into six classes, i.e., adipose (ADI), connective (CON), glandular (GLA), fibro-adenoma (FAD), mastopathy (MAS), and carcinoma (CAR). The detailed information about each tissue class is shown in Table I. For the purpose of notation description, let $S = \{(\mathbf{x}_i, t_i)\}_{i=1}^{N}$ $(N = 106)$ be the set of the sample-class pairs of breast tissues studied. The ith tissue sample that is characterized by the vector of the impedance-measurement features $\mathbf{x}_i$ $(\mathbf{x}_i \in \Re^d,\ d = 9)$, is expected to be categorized into one of the M possible classes, i.e., $t_i \in \{t^1, \cdots, t^M\}$ $(M = 6)$. For the purpose of numerical analysis, we converted the class names into a group of M-dimensional vectors, in which one of the M coordinates is one, and the rest are zero.

TABLE I
Classes of Breast Tissues

Class Names	Abbreviations	Number of Breast Tissue Samples
Carcinoma	CAR	21
Fibro-adenoma	FAD	15
Mastopathy	MAS	18
Glandular	GLA	16
Connective	CON	14
Adipose	ADI	22

III. The Scheme of the Multiple Classifier System for Breast Tissue Classification

A. Component RBF classifiers

An illustration of the multiple classifier system used for the classification of breast tissues is shown in Fig. 1. In our experiments, a total of four three-layer feedforward neural networks based on Radial Basis Function (RBF) were employed as component classifiers. The first layer of a component RBF classifier contains nine sensory neurons in accordance with tissue features. The hidden layer is comprised of nonlinear neurons with the Gaussian kernel function defined by

$$\phi(\cdot) = \exp\left(-\ln 2 \frac{\|\cdot\|^2}{\sigma^2}\right),$$

where σ is the spread parameter that determines the width of the area in the input space to which each hidden neuron responds. For example, a hidden neuron with a spread of 0.1 provides the output of 0.5 for any Euclidean distance of 0.1. To achieve ensemble diversity, the spread parameter was set to be incremental from 1.0 to 4.0 for each RBF classifier. The output layer is linear, and the response of the kth component RBF classifier to the ith tissue sample is denoted as $\hat{f}^k(\mathbf{x}_i)$, which is later sent to the succeeding linear fusion module. The learning of all four component RBF classifier was carried out with the orthognal least-squares method [35]. To select the optimal structure of each component classifier, the number of hidden neurons was varied over the range $[1, 25]$. From the mean-squared error (MSE) results tested with the leave-one-out method (listed in Table II), it can be inferred that the RBF classifiers with 14, 4, 10, and 10 hidden neurons could produce the minimum MSEs, respectively.

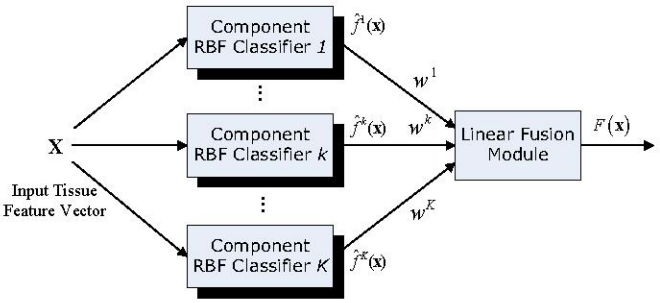

Fig. 1. Illustration of a multiple classifier system with the linear fusion of component RBF classifiers.

B. Normalized Weighted Average Algorithm for Unbiased Linear Fusion

To solve a multi-class pattern recognition problem, the number of fusion modules in the multiple classifier system is commonly set to be the same as the dimension of the class vector, i.e., each fusion module contributes a coordinate in the M-dimensional vector, and the interpretation of the predicted class should provide the maximum *a posteriori* probability among the coordinates.

For the sake of convenience of analytical study, we just clarify the training process of a single linear fusion module with the Normwave algorithm, and the working functions of all M linear fusion modules are similar in the breast tissue classification system studied. Consider the multiple classifier system that consists of a total of K $(K = 4)$ component RBF classifiers, the output produced by a unbiased linear fusion module functioned by the normalized weighted average can be expressed as

$$F(\mathbf{x}_i) = \sum_{k=1}^{K} w^k \hat{f}^k(\mathbf{x}_i), \tag{1}$$

and

$$\sum_{k=1}^{K} w^k = 1, \tag{2}$$

TABLE II
MEAN-SQUARED ERRORS PRODUCED BY A SINGLE RBF CLASSIFIER WITH DIFFERENT PARAMETERS

Number of Hidden Nodes	MSE associated with different spread parameters			
	σ = 1.0	σ = 2.0	σ = 3.0	σ = 4.0
1	0.1304	0.1184	0.1218	0.1182
2	0.1136	0.1064	0.1004	0.1006
3	0.1104	0.0909	0.0900	0.0900
4	0.1051	**0.0800**	0.0846	0.0855
5	0.1060	0.0802	0.0816	0.0835
6	0.1073	0.0805	0.0779	0.0859
7	0.1041	0.0810	0.0802	0.0805
8	0.0890	0.0824	0.0770	0.0803
9	0.0888	0.0819	0.0776	0.0794
10	0.0885	0.0801	**0.0769**	**0.0786**
11	0.0860	0.0816	0.0803	0.0790
12	0.0869	0.0819	0.0806	0.0802
13	0.0831	0.0826	0.0819	0.0805
14	**0.0823**	0.0841	0.0819	0.0838
15	0.0847	0.0842	0.0824	0.0856
16	0.0838	0.0844	0.0817	0.0836
17	0.0853	0.0834	0.0809	0.0842
18	0.0844	0.0830	0.0826	0.0853
19	0.0845	0.0855	0.0844	0.0880
20	0.0846	0.0847	0.0837	0.0888
21	0.0848	0.0865	0.0850	0.0904
22	0.0848	0.0873	0.0902	0.0926
23	0.0865	0.0863	0.0937	0.0929
24	0.0860	0.0857	0.0954	0.0960
25	0.0832	0.0871	0.0971	0.0950

where w^k is the normalized weighted coefficient assigned to the kth component RBF classifier. The MSE between the linear fusion output and desired classes on the whole data set of size N is written as

$$E_{fusion} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{2}\left(t_i - \sum_{k=1}^{K} w^k \hat{f}^k(\mathbf{x}_i)\right)^2\right]. \quad (3)$$

Therefore, the gradient of overall MSE with respect to the weighted coefficients is computed as follows,

$$\begin{aligned}\nabla_{w^k} E_{fusion} &= \frac{\partial}{\partial w^k}\frac{1}{N}\sum_{i=1}^{N}\left[\frac{1}{2}\left(t_i - \sum_{k=1}^{K} w^k \hat{f}^k(\mathbf{x}_i)\right)^2\right] \\ &= -\frac{1}{N}\sum_{i=1}^{N}\hat{f}^k(\mathbf{x}_i)\left[t_i - \sum_{j=1}^{K} w^j \hat{f}^j(\mathbf{x}_i)\right] \quad (4)\end{aligned}$$

By substituting (2) and (4) into the steepest-descent method, we may derive the training rule of the fusion weighted coefficients as

$$\begin{aligned}\hat{w}^k &= w^k + \eta[-\nabla_{w^k} E_{fusion}] \\ &= w^k + \frac{\eta}{N}\sum_{i=1}^{N}\hat{f}^k(\mathbf{x}_i)\left[t_i - \sum_{j=1}^{K} w^j \hat{f}^j(\mathbf{x}_i)\right] \\ &= w^k + \frac{\eta}{N}\sum_{i=1}^{N}\hat{f}^k(\mathbf{x}_i)\left[\sum_{j=1}^{K} w^j\left(t_i - \hat{f}^j(\mathbf{x}_i)\right)\right] \\ &= w^k + \frac{\eta}{N}\sum_{i=1}^{N}\hat{f}^k(\mathbf{x}_i)\sum_{j=1}^{K} w^j e^j(\mathbf{x}_i), \quad (5)\end{aligned}$$

where η is the learning rate ($0 < \eta \leq 1$), and $e^j(\mathbf{x}_i)$ represents the difference between the desired class value and actual response of the jth component RBF classifier for the ith tissue sample, i.e.,

$$e^j(\mathbf{x}_i) = t_i - \hat{f}^j(\mathbf{x}_i). \quad (6)$$

Equation (5) can be rewritten in the matrix form, and then the normalized weighted average (Normwave) algorithm can be formulated to be

$$\hat{w}^k = w^k + \eta \mathbf{w}^T \mathbf{e}\hat{\mathbf{f}}^k, \quad (7)$$

where $\mathbf{w}$ is the $K \times 1$ weighted coefficient vector, the instantaneous estimated error $\mathbf{e}$ is a $K \times N$ matrix, and $\hat{\mathbf{f}}^k$ is the $N \times 1$ output vector of component RBF classifier in accordance with the kth tissue sample.

IV. EMPIRICAL RESULTS

We carried out the 10-fold Cross-Validation method [36] to evaluate the numerical experiments. The data set was randomly partitioned into 10 disjoint subsets of equal size. Each time with a single subset is retained as a validation set, and the remaining 9 subsets are used as training data. The 10 results from the folds were then combined to produce an overall classification for 106 tissue samples. For the purpose of comparison, we also implemented the most frequently used SA and AME ensemble methods.

Fig. 2 shows the overall classification accuracy performed by the three ensemble methods, together with the average results of the component RBF neural classifiers. It is noted that the fusion methods can ameliorate the classification performance in relation to that of a single RBF classifier at an average level, in particular that the SA, AME, and Normwave fusion improve 4.95%, 4.01%, and 7.78% accuracy, respectively.

The classification results on the specific tissue categories performed by the SA, AME, and Normwave fusions are tabulated in Table III. It can also be observed that the Normwave fusion is superior in the tasks of categorizing the specific breast tissues, whereas the SA and AME methods failed to ameliorate the classification performance on the CAR and ADI.

In addition, we also applied the measure of reliability index [24], [25] as an indicator of the certainty level of the ensemble results. A reliability index (RI) is defined to be

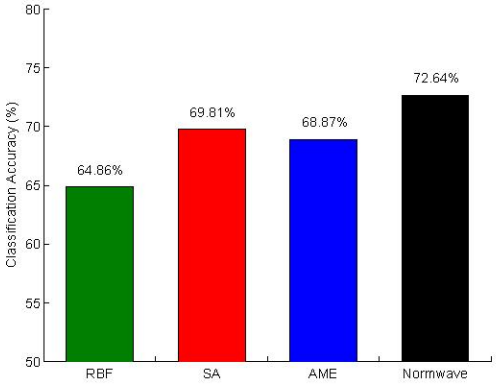

Fig. 2. Classification accuracy performed by different fusion methods.

TABLE III
ACCURACY OF BREAST TISSUE CLASSIFICATION

Class Names	Classification Accuracy (%)			
	RBF	SA	AME	Normwave
CAR	85.00	83.81	83.81	91.43
FAD	32.67	44.00	44.00	57.33
MAS	14.44	26.67	26.67	25.56
GLA	61.88	63.75	63.75	65.00
CON	84.29	92.86	92.86	92.86
ADI	92.50	91.82	91.82	97.27

difference $(diff)$ between the highest and the second highest output value of a classification system, i.e.,

$$RI = \begin{cases} \text{INTEGER}(diff) + 1, & \text{if } 0 \leq diff < 9.0 \\ 10, & \text{if } diff \geq 9.0 \end{cases}$$

where INTEGER$(\cdot)$ denotes the round function towards zero.

The curves and values of normalized classification accuracy over the reliability index provided by the ensemble methods are presented in Fig. 3 and Table IV, respectively. It can be observed that the performance of Normwave fusion is consistently better than the SA or AME except for the RI value of 7 and 10.

TABLE IV
RELIABILITY INDEX RESULTS OF THE ENSEMBLE METHODS STUDIED

Reliability Index	Normalized Classification Accuracy		
	SA	AME	Normwave
1	0.39	0.34	0.40
2	0.63	0.65	0.76
3	0.80	0.75	0.80
4	0.78	0.60	1.00
5	0.50	0.83	1.00
6	1.00	0.88	1.00
7	0.83	1.00	**0.83**
8	1.00	1.00	1.00
9	1.00	1.00	1.00
10	1.00	1.00	**0.89**

In Fig. 4, the normalized classification accuracy was calculated cumulatively with respect to a range of reliability indices by starting with tissue samples with the highest RI, and then progressively including those with lower indices until the lowest RI. It can be observed that around 67% of all the tissue samples have $RI = 2$, and among these tissues about 0.89 accuracy are correctly classified by the Normwave fusion, which is notably higher than that of the SA (0.84) and AME (0.84) methods.

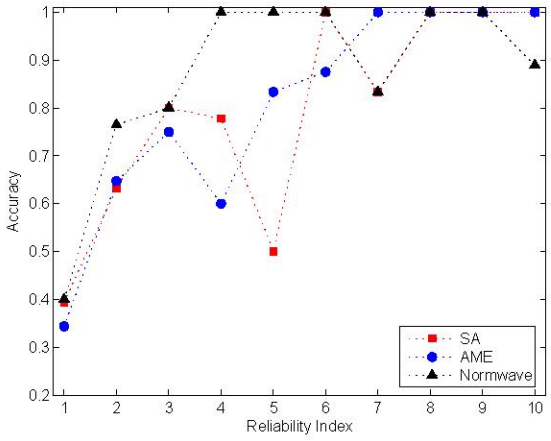

Fig. 3. Classification accuracy over a range of reliability indices.

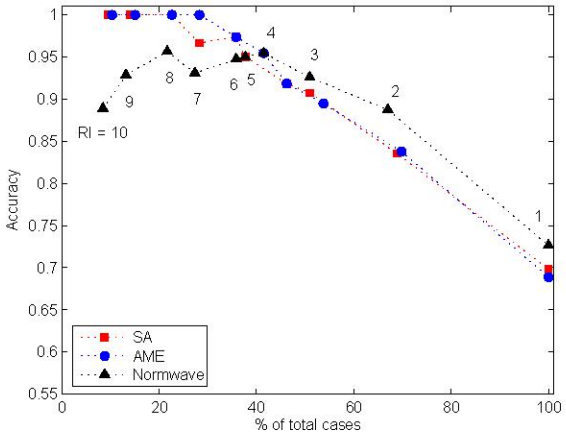

Fig. 4. Plots of cumulative accuracy and percentage of classified tissue samples with descending reliability indices..

V. CONCLUSIONS

The empirical results of the classification of breast tissues demonstrate that the proposed Normwave algorithm for the unbiased linear fusion in the multiple classifier system works effectively in improving classification accuracy, and also achieves confident performance in terms of reliability index. The tangible superiority versus the SA and AME methods indicates the merits of the Normwave algorithm for the design of multiple classifier systems.

The theoretical analysis of the generative ensemble methods and more experimental comparison with other nongenerative ensemble rules, e.g., the Median, Min , Max rules [37], [38], would be the next step of our work.

ACKNOWLEDGMENT

The authors would like to thank Ms. Ye Wu and Ms. Yachao Zhou for their technical assistance and review. This work was supported in part by the National Science Foundation of China under grant No. 60575034; the Doctoral Program Foundation of Ministry of Education of China under grant No. 20060013007; and the 2005 Innovation Research Funds from the Graduate School, Beijing University of Posts and Telecommunications. Yunfeng Wu receives a grant of the 2007 "Visitorship for Scholar in Mainland of China" sponsored by the Croucher Foundation in Hong Kong.

REFERENCES

[1] A. Jemal, R. Siegel, E. Ward, T. Murray, J. Xu, and M. J. Thun, "Cancer statistics, 2007," *CA Cancer J. Clin.*, vol. 57, no. 1, pp. 43–66, 2007.

[2] *Canadian Cancer Statistics 2007*, National Cancer Institute of Canada, 2007.

[3] W. L. Donegan, J. S. Spratt, and J. A. Orsini, *Cancer of the Breast, 5th ed.*, Amsterdam, Netherlands: Elsevier Science, 2002.

[4] R. M. Rangayyan, *Biomedical Image Analysis*, Boca Raton, FL: CRC Press, 2005.

[5] L. Shen, R. M. Rangayyan, and J. E. L. Desautels, "Detection and classification of mammographic calcifications," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 6, pp. 1403–1416, 1993.

[6] X. Zhou and R. Gordon, "Detection of early breast cancer: An overview and future prospects," *Critical Reviews in Biomedical Engineering*, vol. 17, no. 3, pp. 203–255, 1989.

[7] P. Miller and S. Astley, "Classification of breast tissue by texture and analysis," *Image and Vision Computing*, vol. 10, no. 5, pp. 277–282, 1992.

[8] D. L. Hudson and M. E. Cohen, *Neural Networks and Artificial Intelligence for Biomedical Engineering*, New York, NY: IEEE Press, 1999.

[9] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification, 2nd ed.*, New York, NY, USA: Wiely, 2001.

[10] R. Panchal and B. Verma, "Characterization of breast abnormality patterns in digital mammograms using auto-associator neural network," in *Proc. 13th Int'l Conf. Neural Information Processing (ICONIP'06), LNCS 4234*, Hong Kong, 2006, pp. 127–136.

[11] ——, "A fusion of neural network based auto-associator and classifier for the classification of microcalcification patterns," in *Proc. 11th Int'l Conf. Neural Information Processing (ICONIP'04), LNCS 3316*, Calcutta, India, 2004, pp. 794–799.

[12] Y. F. Wu, J. J. He, Y. Man, and J. I. Arribas, "Neural network fusion strategies for identifying breast masses," in *Proc. 2004 IEEE Int'l Joint Conf. Neural Networks*, Budapest, Hungary, 2004, pp. 2437–2442.

[13] B. Verma and J. Zakos, "A computer-aided diagnosis system for digital mammograms based on fuzzy-neural and feature extraction techniques," *IEEE Trans. Information Technology in Biomedicine*, vol. 5, no. 1, pp. 46–54, 2001.

[14] S. Haykin, *Neural Networks: A Comprehensive Foundation, 2nd ed.*, Englewood Cliffs, NJ, USA: Prentice Hall PTR, 1998.

[15] L. K. Hansen and P. Salamon, "Neural network ensembles," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 10, pp. 993–1001, 1990.

[16] S. Hashem and B. Schmeiser, "Improving model accuracy using optimal linear combinations of trained neural networks," *IEEE Trans. Neural Networks*, vol. 6, no. 3, pp. 792–794, 1995.

[17] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matus, "On combining classifiers," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 20, no. 3, pp. 226–239, 1998.

[18] N. Ueda, "Optimal linear combination of neural networks for improving classification performance," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 2, pp. 207–215, 2000.

[19] L. I. Kuncheva, "A theoretical study on six classifier fusion strategies," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 281–286, 2002.

[20] K. Tumer and J. Ghosh, "Analysis of decision boundaries in linearly combined neural classifiers," *Pattern Recognition*, vol. 29, no. 2, pp. 341–348, 1996.

[21] Y. F. Wu, J. M. Zhang, C. Wang, and S. C. Ng, "Linear decision fusions in multilayer perceptrons for breast cancer diagnosis," in *Proc. 17th IEEE Int'l Conf. Tools with Artificial Intelligence (ICTAI'05)*, Hong Kong, 2005, pp. 699–700.

[22] Y. F. Wu, C. Wang, S. C. Ng, A. Madabhushi, and Y. X. Zhong, "Breast cancer diagnosis using neural-based linear fusion strategies," in *Proc. 13th Int'l Conf. Neural Information Processing (ICONIP'06), LNCS 4234*, Hong Kong, 2006, pp. 165–175.

[23] Y. F. Wu and C. Wang, "Linear least-squares fusion of multilayer perceptrons for protein localization sites prediction," in *Proc. 32nd IEEE Annu. Northeast Bioeng. Conf. (NEBC'06)*, Easton, PA, USA, 2006, pp. 157–158.

[24] S. Hua and Z. Sun, "Support vector machine approach for protein subcellular localization prediction," *Bioinformatics*, vol. 17, no. 8, pp. 721–728, 2001.

[25] A. Reinhardt and T. Hubbard, "Using neural networks for prediction of the subcellular location of proteins," *Nucleic Acids Research*, vol. 26, no. 9, pp. 2230–2236, 1998.

[26] G. Valentini and F. Masulli, "Ensembles of learning machines," in *Proc. 13th Italian Workshop on Neural Nets (WIRN'02), LNCS 2486*, Vietri sul Mare, Italy, 2002, pp. 3–19.

[27] Y. F. Wu and J. I. Arribas, "Fusing output information in neural networks: Ensemble performs better," in *Proc. 25th IEEE EMBS Annu. Int'l Conf. (EMBC'03)*, Cancun, Mexico, 2003, pp. 2265–2268.

[28] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixtures of local experts," *Neural Computation*, vol. 3, no. 11, pp. 79–87, 1991.

[29] N. M. Wanas and M. S. Kamel, "Decision fusion in neural network ensembles," in *Proc. 2001 Int'l Joint Conf. Neural Networks (IJCNN'01)*, Washington, DC, USA, 2001, pp. 2952–2957.

[30] G. Fumera and F. Roli, "A theoretical and experimental analysis of linear combiners for multiple classifier systems," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 27, no. 6, pp. 942–956, 2005.

[31] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, 1997.

[32] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.

[33] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap*. New York, NY: Chapman and Hall, 1993.

[34] J. da Silva, J. Marques de Sa, and J. Jossinet, "Classification of breast tissue by electrical impedance spectroscopy," *Med. Biol. Eng. Comput.*, vol. 38, no. 1, pp. 26–30, 2000.

[35] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Networks*, vol. 2, no. 2, pp. 302–309, 1991.

[36] T. Y. Kwok and D. Y. Yeung, "Efficient cross-validation for feedforward neural networks," in *Proc. 1995 Int'l Joint Conf. Neural Networks (IJCNN'95)*, Perth, WA, Australia, 1995, pp. 2789–2794.

[37] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. Systems, Man, and Cybernetics*, vol. 22, no. 3, pp. 418–435, 1992.

[38] L. I. Kuncheva, "Switching between selection and fusion in combing classifiers: An experiment," *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 32, no. 2, pp. 146–156, 2002.