

Wavelet based approach for facial expression recognition

Zaenal Abidin^{a,b,1,*}, Alamsyah^{a,2}

^a Semarang State University, Building D2, Level 1, Kampus Sekaran, Gunungpati, Semarang, 50229, Indonesia

^b School of Engineering and Advanced Technology, Massey University, Albany, New Zealand

¹zaenalabidin@staff.unnes.ac.id*; ²alamsyah@mail.unnes.ac.id

ARTICLE INFO

Article history:

Received March 14, 2015

Revised March 29, 2015

Accepted March 30, 2015

Keywords:

Wavelet transforms

Backpropagation neural network

Facial expression

Pattern recognition

ABSTRACT

Facial expression recognition is one of the most active fields of research. Many facial expression recognition methods have been developed and implemented. Neural networks have capability to undertake such pattern recognition tasks. The key factor of the use of neural network is based on its characteristics. It is capable in conducting learning and generalizing, non-linear mapping, and parallel computation. Backpropagation neural networks (BPNN) are the approach methods that mostly used. In this study, BPNN was used as classifier to categorize facial expression images into seven-class of expressions which are anger, disgust, fear, happiness, sadness, neutral and surprise. For the purpose of feature extraction tasks, three discrete wavelet transforms were used to decompose images, namely Haar wavelet, Daubechies (4) wavelet and Coiflet (1) wavelet. To analyze the proposed method, a facial expression recognition system was built. The proposed method was tested on static images from JAFFE database.

Copyright © 2015 International Journal of Advances in Intelligent Informatics.
All rights reserved.

1. Introduction

In everyday life, people get in touch and communicate with others using verbal and non-verbal ways. One of non-verbal communications is emotions which are usually expressed through facial expressions. A facial expression is result of one or many movements or positions of facial muscles. By recognizing facial expressions, the emotion of people can be predicted. For instance, a smile expresses hospitality and affection, a lift of eyebrows shows confusion, a wince of forehead portrays fear and anxiety. According to Ekman [1], humans have universal facial expressions and the expressions categorized into six classes these are happiness, sadness, disgust, anger, surprise, and fear.

In recent years, researchers have been interested in studying facial expression recognition and over time, biometrics is widely used to recognize facial expression. It is applied using distinguishing traits. In general, these traits are divided into two, namely physical and behavioral characteristics. A facial expression is one of behavioural characteristics. The use of these characteristics allows us to recognize emotions of people.

Several methods for facial expression recognition have been developed and implemented. Long et al. [1] examined facial expression recognition using Pseudo Zernike Moment Invariant (PZMI) as a feature extraction from the global information of images and the Radial Basis Function (RBF) network was employed to be a classifier. In addition, Bashyal and Venayagamoorthy [2] applied Gabor wavelet and learning vector quantization (LVQ) in their study. Kulkarni [3] developed an applied smart system to recognize emotions using committee neural network. Meanwhile, Deng et al. [4] performed a research in comparing local Gabor filter bank using the approach of principal component analysis (PCA) and linear discriminant analysis (LDA). Furthermore, Ma and Khorasani [5] carried out facial expression recognition study using 2-D discrete cosine transform (DCT) and feedforward neural networks. Moreover, Abidin and Harjoko [6] analyzed facial expression recognition system using fisherface and BPNN approach.

Most of facial expression recognition studies applied neural networks that were combined by different kind of methods. The key factors of neural network are able to conduct learning and

generalization, non-linear mapping as well as parallel computation. However, neural networks sometimes trapped into local minima that lead to slow convergence level. According to Abiyev and Kaynak [7], to cope with neural network weaknesses, wavelet function applied in neural networks structure becomes an appropriate approach.

BPNN has been applied in many applications until such as in system modelling [8]–[10], remote sensing [11]–[13], in time series predication [14]–[16], recognition system [17]–[19], network security [20].

In this study, we developed a facial expression system using BPNN and wavelet based feature extraction which was expected to achieve similar or even better performance of facial expression recognition. In particular, the aims of this study are investigating the appropriate wavelet transform that can be used as feature extraction and analyzing the performance of the systems.

II. Research Method

A. The Need for System Analysis

Six basic facial expressions such as happiness, sadness, anger, surprise, fear, and disgust, plus one neutral expression will be investigated to measure performance of proposed method. This study was performed using JAFFE database [21]. It comprises 213 facial expression images from 10 Japanese women. Every subject posed three or four times to get six basic facial expressions and one neutral. The resolution of the image is 256×256 in *.tiff format with grayscale color mode. Images used for this study were split into two groups of training and testing data sets with 140 images and 73 images, respectively. The splitting image data follow the previous research performed by Abidin and Harjoko [6]. The training data set was used for learning process meanwhile the testing data set was used for analyzing system performance.

Each image is represented in a matrix in which each element of the matrix corresponds to the intensity level of a pixel. The representation matrix of image usually has a big dimension which can lead to the process is computationally expensive. Therefore, reduction dimension is necessary. In this study, wavelet functions will be used to reduce the dimension of images.

The following steps describe the process of analyzing facial expression system using the proposed method:

1) Image pre-processing and feature extraction.

The step of image pre-processing is started by automatically face detection. There are several methods available for performing face detection. In this study, the integral projection method proposed by Abidin and Harjoko [6] is used. Once a face is detected, image pre-processing then can be employed. Firstly, the size of images will be normalized into 130×114 pixel.

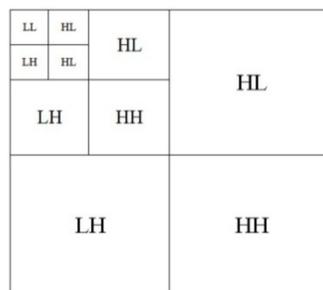


Fig. 1. Illustration of 3-Level 2D Decomposition

Secondly, image contrast is enlarged by applying histogram equalization and finally masking process is undertaken by covering each angle of image in order to decrease the variations appearing on them. After image pre-processing done, the next stage is feature extraction. In this study, three wavelet functions are used to extract the features, i.e. Haar wavelet, Daubechies (4) wavelet, and Coiflet (1) wavelet. The images are performed 3-level 2D decomposition and the dimension of images is changed into 17×15 pixel.

2) Learning process of neural networks for images in training data set.

In this step, all images in training data set were trained using a classifier that is BPNN. This neural network architectural design comprises fully weighted interconnected rows of nodes as processing units. Some nodes are organized into layers [6], [22]. A typical neural network consist of an input layer, an output layer, and one or more hidden layers [3]. The most popular learning algorithm for multi-layer perceptron is backpropagation. The backpropagation algorithm defines a systematic way to update the synaptic weights of multi-layer perceptron (MLP) networks.

The learning of backpropagation algorithm is performed in two phases: forward phase and backward phase [23]. During the first phase, the free parameters of the network are determined, and then the input signal is propagated through the network layer by layer. In this phase, an error signal is calculated by (1):

$$e_i = d_i - y_i \quad (1)$$

Where d_i is the desired response and y_i is the actual output produced by the network in response to the input x_i .

During this backward phase, the error signal e_i is propagated through the network in the backward direction. It is during this phase that adjustments are applied to the free parameters of the network in order to minimize the error e_i in a statistical sense.

In this study, backpropagation learning is implemented using pattern mode in which “weight updating is performed after the presentation of each training pattern” [24, p. 46]. The outcome of learning process is stored to database and is used for testing analysis of the facial expression recognition system.

3) Testing the facial expression recognition system both images in training and testing data sets.

The last stage of evaluating the facial expression recognition system is analyzing the performance of the system by measuring the recognition rate.

B. System Design

In this sub section describes design of the facial expression recognition system. To investigate the performance of proposed method, an interface of facial expression recognition system has been developed using MATLAB. The interface consists of three subsystems, which are:

1) Subsystem image registration

This subsystem comprises some features, i.e. the storing tool, face detector, and feature extractor. Storing tool is used to record keeping images into the system. When an image is registered into the system, the face detector will capture the face region, and do image pre-processing. Integral projection method was employed to accomplish face detection. An integral projection is obtained through the sum of given set of pixels along a given direction (for details, see [6]).

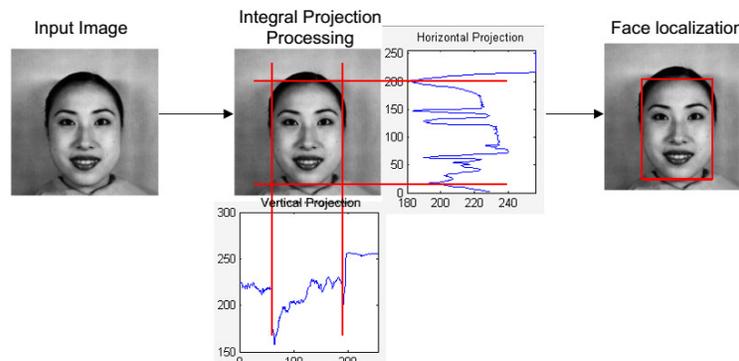


Fig. 2. Illustration of Integral Projection Usage for Face Detection Proposed by Abidin & Harjoko [6]

After image pre-processing, the second important task is feature extraction. The features are obtained by decomposing the images with wavelet transforms. In this study, Haar wavelet,

- Daubechies (4) wavelet, and Coiflet (1) wavelet are employed. Then the features are stored in a database.
- 2) Subsystem of neural network learning
Learning algorithm is built in this subsystem. BPNN's structure consists of three layers, namely input layer, hidden layer and output layer. The activation function for hidden layer is sigmoid bipolar while the activation function for output layer is sigmoid, hence the outcome of output layer is expected to be in the range [0, 1]. The weights of BPNN learning result then stored to database and will be used to recognition processes.
 - 3) Subsystem of facial expression recognition
This subsystem further classifies the facial expression into different categories based on the extracted data fed to it from a facial image. This subsystem can be seen in Fig. 3.

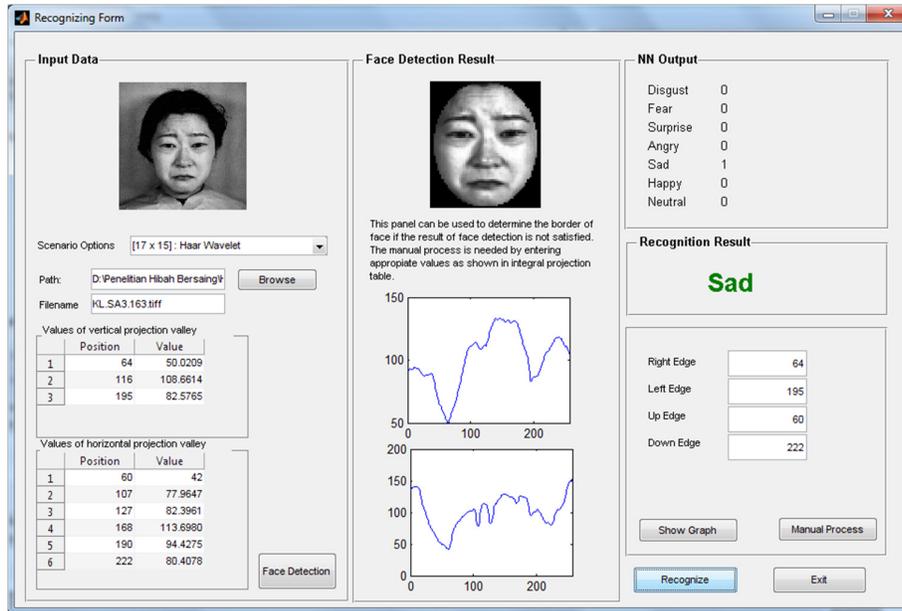


Fig. 3. Recognizing form

III. Experimental Results

In this study three experiments were conducted as trying to verify the effect of wavelet transforms on the performance of facial expression recognition system. The learning process was conducted by varying on the neural network parameters. The number of neurons in hidden layer was set by 5 while the learning rate was set by 0.25. Moreover, the minimum error was set by 0.001 and the maximum number of epoch was 1000.

Facial expression recognition system has been tested by using 213 images from JAFFE database. The number of 140 images has been trained and the remaining 73 images were used for testing. Since the number of images in the JAFFE database is limited, we had performed the trial over 5 times to get the average classification rate.

Table 1. Recognition Rate for Each Wavelet Transform

| Wavelet | Recognition Rate (%) |
|----------------|----------------------|
| Haar | 94.12 |
| Daubechies (4) | 94.22 |
| Coiflet (1) | 93.08 |

Table 1 shows the average of recognition rate of wavelet-based facial expression with the trial over 5 times. From Table 1 it can be seen that the recognition rate of data set with Coiflet (1) wavelet

transform is lower than two others, however there is no significant difference in terms of recognition rate among wavelet transforms.

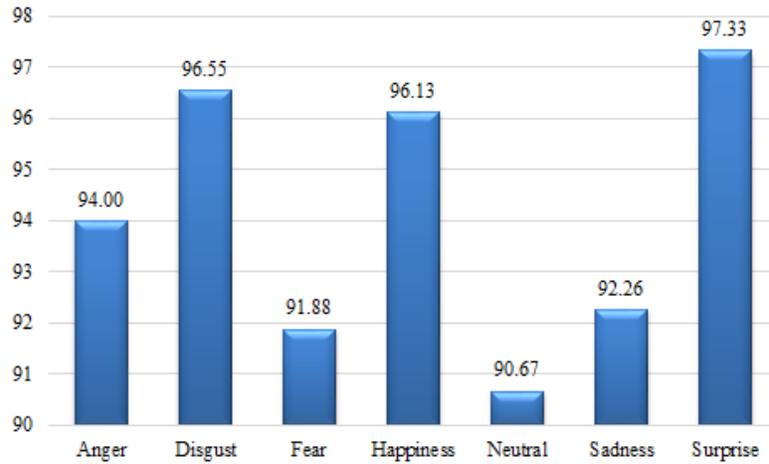


Fig. 4. Recognition Rate of Each Facial Expression for Data Set with Haar Wavelet (%)

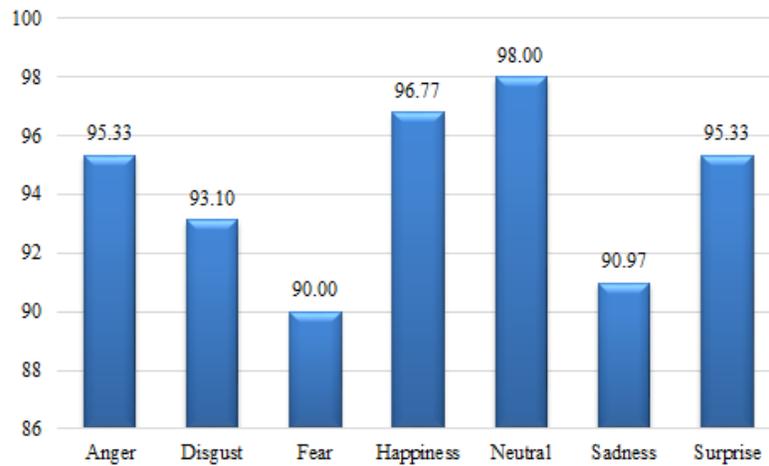


Fig. 5. Recognition Rate of Each Facial Expression for Data Set with Daubechies (4) Wavelet (%)

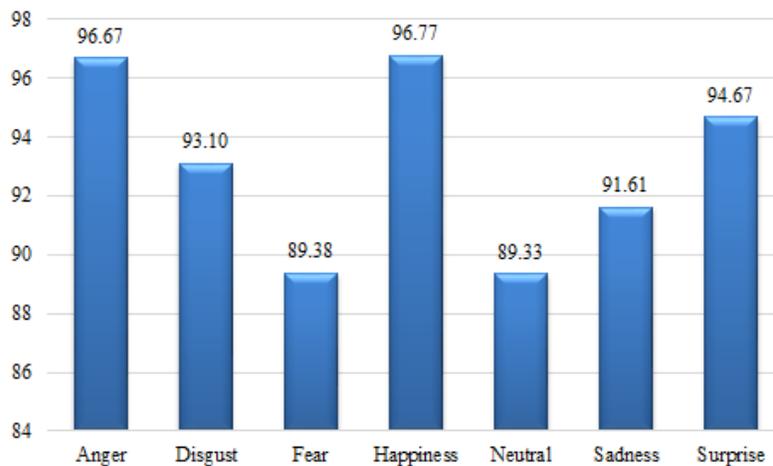


Fig. 6. Recognition Rate of Each Facial Expression for Data Set with Coiflet (1) Wavelet (%)

In detail, the recognition rate of seven-class facial expressions for each Haar wavelet, Daubechies (4) wavelet and Coiflet (1) wavelet is depicted in Fig. 4, Fig. 5 and Fig. 6 respectively. The recognition rates for each facial expression from data set with Haar wavelet are over 90%. Similarly, the recognition rates from data set with Daubechies (4) wavelet and Coiflet (1) wavelet show good performances.

From three figures (Fig. 4, Fig. 5 and Fig. 6), it can be seen that fear expression is always in the lowest rate. Meanwhile, in Table 2, Table 3 and Table 4 show that some fear expression data are dominantly recognized as surprise expression. Similarly, sadness expression is dominantly identified as disgust and anger expression. Since facial expressions are combination of facial muscles movements, these may lead ambiguity [25].

Table 2. Confusion Matrix of All Data Sets with Haar Wavelet (%)

| IO | Anger | Disgust | Fear | Happiness | Neutral | Sadness | Surprise |
|-----------|-------|---------|-------|-----------|---------|---------|----------|
| Anger | 94.00 | 2.67 | 0.67 | 0.00 | 2.67 | 0.00 | 0.00 |
| Disgust | 0.00 | 96.55 | 2.76 | 0.00 | 0.00 | 0.69 | 0.00 |
| Fear | 0.63 | 1.88 | 91.88 | 0.00 | 0.00 | 2.50 | 3.13 |
| Happiness | 0.00 | 0.65 | 0.00 | 96.13 | 1.94 | 1.29 | 0.00 |
| Neutral | 0.67 | 0.00 | 0.67 | 2.67 | 90.67 | 5.33 | 0.00 |
| Sadness | 0.00 | 3.23 | 0.65 | 2.58 | 1.29 | 92.26 | 0.00 |
| Surprise | 0.00 | 0.00 | 0.67 | 0.00 | 1.33 | 0.67 | 97.33 |

Table 3. Confusion Matrix of All Data Sets with Daubechies (4) Wavelet (%)

| IO | Anger | Disgust | Fear | Happiness | Neutral | Sadness | Surprise |
|-----------|-------|---------|-------|-----------|---------|---------|----------|
| Anger | 95.33 | 0.67 | 0.00 | 0.00 | 4.00 | 0.00 | 0.00 |
| Disgust | 0.69 | 93.10 | 4.83 | 0.00 | 0.00 | 1.38 | 0.00 |
| Fear | 1.25 | 0.63 | 90.00 | 0.63 | 0.63 | 1.88 | 5.00 |
| Happiness | 0.65 | 0.65 | 0.00 | 96.77 | 1.94 | 0.00 | 0.00 |
| Neutral | 0.67 | 0.00 | 0.00 | 0.00 | 98.00 | 0.00 | 1.33 |
| Sadness | 0.65 | 3.23 | 0.65 | 3.23 | 1.29 | 90.97 | 0.00 |
| Surprise | 2.00 | 0.00 | 0.67 | 0.67 | 0.00 | 1.33 | 95.33 |

Table 4. Confusion Matrix of All Data Sets with Coiflet (1) Wavelet (%)

| IO | Anger | Disgust | Fear | Happiness | Neutral | Sadness | Surprise |
|-----------|-------|---------|-------|-----------|---------|---------|----------|
| Anger | 96.67 | 1.33 | 0.00 | 0.00 | 2.00 | 0.00 | 0.00 |
| Disgust | 0.00 | 93.10 | 4.14 | 0.00 | 0.00 | 2.76 | 0.00 |
| Fear | 0.00 | 1.25 | 89.38 | 0.63 | 1.88 | 3.13 | 3.75 |
| Happiness | 0.65 | 0.00 | 0.00 | 96.77 | 0.65 | 1.94 | 0.00 |
| Neutral | 0.67 | 0.00 | 1.33 | 3.33 | 89.33 | 4.00 | 1.33 |
| Sadness | 2.58 | 1.29 | 1.29 | 0.65 | 1.94 | 91.61 | 0.65 |
| Surprise | 0.00 | 0.00 | 5.33 | 0.00 | 0.00 | 0.00 | 94.67 |

Although some expressions may be ambiguous, six expressions, namely anger, disgust, happiness, neutral, sadness, and surprise were well classified. The recognition rate of those expressions was that over 90% meanwhile fear expression was around 89.38% to 91.88%. Convolution matrices of Table 2, Table 3 and Table 4 give details information about facial expression recognition rate for each wavelet transform.

Overall, there are no significantly differences among those outcomes. In other words, Haar wavelet, Daubechies (4) wavelet or Coiflet (1) wavelet give a similar influence in classifying facial expression into seven group classes.

We have compared our proposed with the existing method which is method that uses the same data for training and testing, face detection method and also the same pre-processing of images. From

Table 5 it can be seen that the proposed method shows better performance than method using fisherface + BPNN.

Table 5. Comparative Results of the Classification Rate of Different Approaches

| Method | Recognition rate |
|-----------------------|------------------|
| Fisherface + BPNN [6] | 86.85% |
| Proposed method | 94.22% |

IV. Conclusions and Future Works

In this paper, a wavelet based facial expression recognition system is proposed. Three wavelet transforms namely Haar wavelet, Daubechies (4) wavelet, and Coiflet (1) wavelet were selected to perform feature extractions. Moreover, BPNN was chosen as a classifier. The proposed method has shown good performances in recognizing facial expression images of JAFFE database. However, the task of facial expression recognition is still challenging since the appearance of expressions can vary by person. Ambiguity of expression may occur and become hard to classify.

In future work the ambiguity of expression can be considered to improve better performance of facial expression recognition rate. It is also important to select different wavelet function to be employed in developing neural networks based facial expression recognition system as comparative study.

Acknowledgment

This research was funded by the *Hibah Bersaing* Research Grant (No: 529/UN37.3.1/LT/2014, 07 May 2014) from the Directorate General of Higher Education, Ministry of National Education and Culture, Indonesia.

References

- [1] T. B. Long, L. H. Thai, and T. Hanh, "Facial expression classification method based on pseudo zernike moment and radial basis function network," in *Proceeding of IEEE 3rd Int. Conf. Machine Learning and Computing (ICMLC)*, Singapore, 2011, pp. 310–313.
- [2] S. Bashyal and G. K. Venayagamoorthy, "Recognition of facial expressions using Gabor wavelets and learning vector quantization," *Eng. Appl. Artif. Intell.*, vol. 21, no. 7, pp. 1056–1064, Oct. 2008.
- [3] S. S. Kulkarni, "Facial image based mood recognition using committee neural networks," University of Akron, 2006.
- [4] H. B. Deng, L. W. Jin, L. W. Zhen, and J. C. Huang, "A new facial expression recognition method based on local Gabor filter bank and PCA plus LDA," *Int. J. Inf. Technol.*, vol. 11, no. 11, pp. 86–96, 2005.
- [5] L. Ma and K. Khorasani, "Facial expression recognition using constructive feedforward neural networks," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 34, no. 3, pp. 1588–1595, Jun. 2004.
- [6] Z. Abidin and A. Harjoko, "A Neural Network based Facial Expression Recognition using Fisherface," *Int. J. Comput. Appl.*, vol. 59, no. 3, pp. 30–34, Dec. 2012.
- [7] R. H. Abiyev and O. Kaynak, "Fuzzy wavelet neural networks for identification and control of dynamic plants - A novel structure and a comparative study," *IEEE Trans. Ind. Electron.*, vol. 55, no. 8, pp. 3133–3140, Aug. 2008.
- [8] S. A. Billings, H. B. Jamaluddin, and S. Chen, "Properties of neural networks with applications to modelling non-linear dynamical systems," *Int. J. Control*, vol. 55, no. 1, pp. 193–224, Jan. 1992.
- [9] S. Chen and S. A. Billings, "Neural networks for nonlinear dynamic system modelling and identification," *Int. J. Control*, vol. 56, no. 2, pp. 319–346, Aug. 1992.
- [10] A. T. C. Goh, "Back-propagation neural networks for modeling complex systems," *Artif. Intell. Eng.*, vol. 9, no. 3, pp. 143–151, 1995.
- [11] J. F. Mas and J. J. Flores, "The application of artificial neural networks to the analysis of remotely sensed data," *Int. J. Remote Sens.*, vol. 29, no. 3, pp. 617–663, Feb. 2008.
- [12] M. Reichstein, P. Ciais, D. Papale, R. Valentini, S. Running, N. Viovy, W. Cramer, A. Granier, J. Ogee, V. Allard, M. Aubinet, C. Bernhofer, N. Buchmann, A. Carrara, T. Grünwald, M. Heimann, B. Heinesch, A. Knohl, W. Kutsch, D. Loustau, G. Manca, G. Matteucci, F. Miglietta, J. m. Ourcival, K. Pilegaard, J.

- Pumpanen, S. Rambal, S. Schaphoff, G. Seufert, J.-F. Soussana, M.-J. Sanz, T. Vesala, and M. Zhao, "Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis," *Glob. Change Biol.*, vol. 13, no. 3, pp. 634–651, Mar. 2007.
- [13] M. Schlerf and C. Atzberger, "Inversion of a forest reflectance model to estimate structural canopy variables from hyperspectral remote sensing data," *Remote Sens. Environ.*, vol. 100, no. 3, pp. 281–294, Feb. 2006.
- [14] K. Huarng and T. H.-K. Yu, "The application of neural networks to forecast fuzzy time series," *Phys. Stat. Mech. Its Appl.*, vol. 363, no. 2, pp. 481–491, May 2006.
- [15] D. Ömer Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Eng. Appl. Artif. Intell.*, vol. 23, no. 4, pp. 586–594, Jun. 2010.
- [16] G. P. Zhang, B. E. Patuwo, and M. Y. Hu, "A simulation study of artificial neural networks for nonlinear time-series forecasting," *Comput. Oper. Res.*, vol. 28, no. 4, pp. 381–396, Apr. 2001.
- [17] L. Liang and D. Wu, "An application of pattern recognition on scoring Chinese corporations financial conditions based on backpropagation neural network," *Comput. Oper. Res.*, vol. 32, no. 5, pp. 1115–1129, May 2005.
- [18] R. Prabhavalkar and E. Fosler-Lussier, "Backpropagation training for multilayer conditional random field based phone recognition," in *2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)*, 2010, pp. 5534–5537.
- [19] J. Sutha and N. Ramaraj, "Neural network based offline tamil handwritten character recognition system," in *Proceeding International Conference on Conference on Computational Intelligence and Multimedia Applications, 2007*, 2007, vol. 2, pp. 446–450.
- [20] K. Shihab, "A Backpropagation Neural Network for Computer Network Security," *J. Comput. Sci.*, vol. 2, no. 9, pp. 710–715, Sep. 2006.
- [21] M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba, "Coding facial expressions with Gabor wavelets," in *Proceeding the third IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 200–205.
- [22] R. J. Erb, "Introduction to Backpropagation Neural Network Computation," *Pharm. Res.*, vol. 10, no. 2, pp. 165–170, Feb. 1993.
- [23] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning Internal Representations by Error Propagation," Sep. 1985.
- [24] D. Svozil, V. Kvasnicka, and J. Pospichal, "Introduction to multi-layer feed-forward neural networks," *Chemom. Intell. Lab. Syst.*, vol. 39, no. 1, pp. 43–62, Nov. 1997.
- [25] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 699–714, May 2005.