A New Method for Image Classification Based on Multi-level Neural Networks

Samy Sadek, Ayoub Al-Hamadi, Bernd Michaelis, Usama Sayed

Abstract— In this paper, we propose a supervised method for color image classification based on a multilevel sigmoidal neural network (MSNN) model. In this method, images are classified into five categories, i.e., "Car", "Building", "Mountain", "Farm" and "Coast". This classification is performed without any segmentation processes. To verify the learning capabilities of the proposed method, we compare our MSNN model with the traditional Sigmoidal Neural Network (SNN) model. Results of comparison have shown that the MSNN model performs better than the traditional SNN model in the context of training run time and classification rate. Both color moments and multi-level wavelets decomposition technique are used to extract features from images. The proposed method has been tested on a variety of real and synthetic images.

Keywords— Image classification, multi-level neural networks, feature extraction, wavelets decomposition.

I. INTRODUCTION

IN recent years automatic image classification and retrieval have been increasingly investigated [1,2, 3]. However, content-based image classification and retrieval techniques are still immature for numerous reasons. The majority of approaches of image classification have addressed classification task using low-level image features such as color, texture, etc. Most of the work in this area [4,5,6] concentrates on semantic classes such as indoor/outdoor, people/non-people, city, landscape, sunset, forest, etc. These are more general, consumer-level, semantic classes compared to the business-use oriented semantic classes that are used by other researchers. Another pertinent work concerns the separation of computer-generated graphic images, such as presentation slides, from photographic images [7,8,9].

The greater part of the approaches addressed image classification deals with pure color images and presumes homogeneity in the color content of the image scene, either explicitly or implicitly. However, real images bring out a wide range of heterogeneity in the color content. This variation of color information induces varying degrees of uncertainty in the information content. The ambiguity in image information emerging from the admixtures of the color components has often been dealt with the soft computing paradigm.

S. Sadek, A. Al-Hamadi and B. Michaelis are with the Institute for Electronics, Signal Processing and Communications (IESK), Otto-von-Guericke University Magdeburg, Magdeburg, CO 39106 Germany (email: {Ayoub.Al-hamdi, Samy.Bakheet}@ovgu.de).
U. Sayed is with the Electrical Engineering Department, Assuit University, CO 71516 Egypt.

Artificial neural network (ANN) architectures have been increasingly employed to deal with many tasks of image processing especially image classification and retrieval. The neural classifier has the advantage of being fast (highly parallel), easily trainable and capable of creating arbitrary partitions of feature space [10]. However a neural network model, in the standard form, is unable to correctly classify images into more than two classes [11].This is due to the fact that each of the component single neuron employs the standard bi-level activation function. Since the bi-level activation function produces only binary responses, the neurons can generate only binary outputs. So, in order to produce multiple color responses either an architectural or a functional extension to the existing neural model is required.

In this work, a new approach for image classification based on multilevel neural network model is presented. Multilevel neural model makes use of a new activation function, named multilevel sigmoidal activation function.

The remainder of the paper is organized as follows. Section 2 shows multilevel activation functions used in this paper. Color moments and multilevel wavelet decomposition are briefly discussed in Section 3. In Section 4 the proposed image classification approach is introduced. Section 5 presents the simulation results of the proposed image classification method and Section 6 closes with a conclusion and discussion on possible enhancements.

II. MULTILEVEL ACTIVATION FUNCTIONS

A multilevel activation function is a functional extension of the standard activation functions in existence. Several multilevel forms pertaining to several standard activation functions can be designed. This section discusses the basic design mechanism of the multilevel versions of the standard activation function. The standard sigmoidal activation function is given by (see Fig.1):

$$f(x) = \frac{1}{1 + e^{-\beta x}} \tag{1}$$

where, β is the steepness factor of the function. The multilevel form of the sigmoidal functions is derived from the previous standard form as follows:

$$\phi(x) \leftarrow f(x) + (\lambda - 1)f(c) \tag{2}$$

where $(\lambda - 1)c \leq x \leq \lambda c$ and $1 \leq \lambda \leq n$.

In the above equation, λ represents the color index, n is the

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:3, No:9, 2009

number of categories, and *c* represents the color scale contribution. Multilevel sigmoidal activation functions for three and five classes are depicted in Fig. 2and Fig. 3.
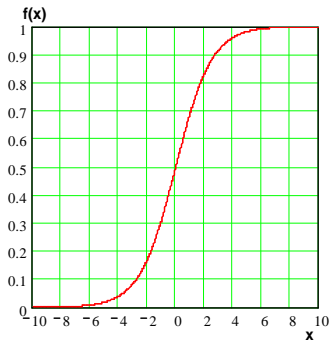

Fig. 1. Standard sigmoidal activation function

Fig. 2. Multilevel sigmoidal function for n=3.
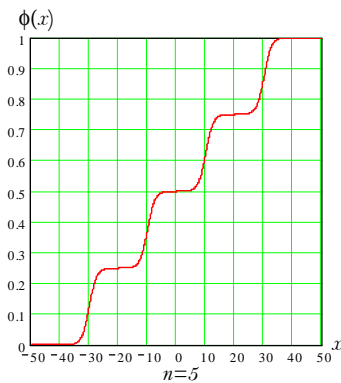

Fig. 3. Multilevel sigmoidal function for n=5.

III. COLOR AND TEXTURE FEATURES

Image classification is usually based on some image features that characterize the image. In the existing content-based image classification and retrieval systems the most common features are color, shape, and texture. Color histograms are regularly used in image classification and retrieval. In this paper, we extract features via two ways: color moments as a descriptor for color and approximation coefficients of multilevel wavelet decomposition as a descriptor for texture.

A. Color Moments

The basis of color moments lays in the assumption that the distribution of color in an image can be interpreted as a probability distribution [12]. Probability distributions are characterized by a number of unique moments (e.g. normal distributions are differentiated by their mean and variance). It therefore follows that if the color in an image follows a certain probability distribution, the moments of that distribution can then be used as features to identify that image based on color. The three central moments (Mean, Standard deviation, and Skewness) of an image's color distribution can be defined as:

$$\mu_k = \frac{1}{n} \sum_{i=1}^{n} p_i^k \tag{3.a}$$

$$\sigma_k = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (p_i^k - \mu_k)^2} \tag{3.b}$$

$$s_k = \sqrt[3]{\frac{1}{n} \sum_{i=1}^{n} (p_i^k - \mu_k)^3} \tag{3.c}$$

where p_i^k is the value of the *k-th* color channel for the *i-th* pixel, and *n* is the size of the image.

B. Wavelet decomposition

Discrete Wavelet Transform (DWT) captures image features and localizes them in both time and frequency content accurately. DWT employs two sets of functions called scaling functions and wavelet functions, which are related to low-pass and high-pass filters, respectively. The decomposition of the two into the different frequency bands is merely obtained by consecutive high-pass and low-pass filtering of the time domain signal. The procedure of multi-resolution decomposition of a signal x[n] is schematically shown in Fig. 3. Each stage of this scheme consists of two digital filters and two down-samplers by 2. The first filter H0 is the discrete mother wavelet; high pass in nature, and the second, H1 is its mirror version, low-pass in nature. The down-sampled outputs of first high-pass and low-pass filters provide the detail, D1 and the approximation, A1, respectively. The first approximation, A1 is further decomposed and this process is continued as shown in Fig. 4.

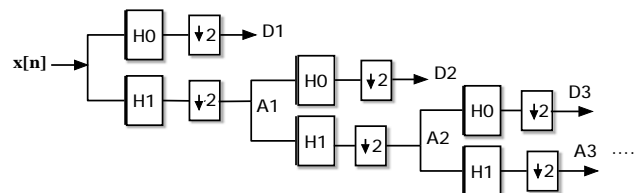Fig. 4. Multilevel wavelets decomposition.

IV. PROPOSED APPROACH

The main difficulty with any image classification process is

World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:3, No:9, 2009

that the unit of information in image is the pixel and each pixel has properties of position and color value; however, by itself, the knowledge of the position and value of a particular pixel should generally convey all information related to the image contents [13, 14]. To avoid this difficulty, features are extracted using two-way. The extracted features consist of two folds: color moments and approximate coefficients of multilevel wavelet decomposition. This allows us to extract from an image a set of numerical features, expressed as coded characteristics of the selected object, and used to differentiate one class of objects from another. The main steps of the proposed approach are shown in the following block diagram (see Fig. 5).
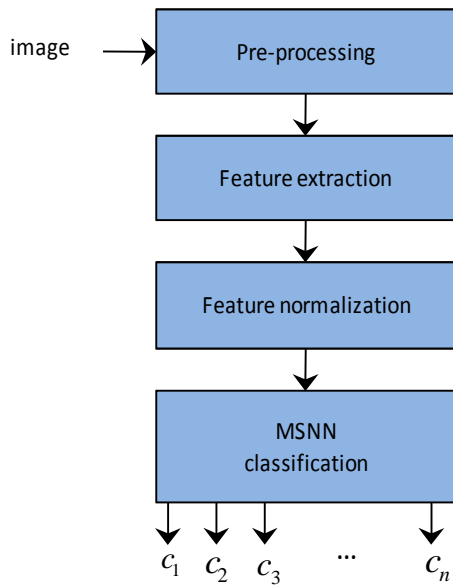

Fig. 6: Coefficients of wavelet decomposition

C. Feature normalization

To prevent singular features from dominating the others and to obtain comparable value ranges, we do feature normalization by transforming the feature components x to a random variable with zero mean and one variance as follows:

$$\bar{x} = \frac{x - \mu}{\sigma} \qquad (4)$$

where μ and σ are the sample mean and the sample standard deviation respectively.

If we assume that each feature is normally distributed, the probability of $\bar{x}$ being in the $[-1,1]$ is 0.68. An additional shift and rescaling as

$$\bar{x} = \frac{\frac{x - \mu}{3\sigma} + 1}{2} \qquad (5)$$

could guarantee 0.99 of $\bar{x}$ values to be in $[0,1]$ range.

D. MSNN image classification

In this step, the multilevel neural network is used. The hidden layer has 50 neurons which use the designed multilayer activation functions. In the training stage, the training algorithm used the back-propagation method: the first pattern is given to the input neurons and the network gives its output. If it is not equal to the desired output pattern, then the algorithm computes the difference (mean squared error) between these two values and changes the weights in order to minimize the error. These operations are repeated for each input pattern until the error is minimized. In the testing stage, the feature vector which holds the normalized features of an input image is fed into network, then the network directly outputs a value that refers to the category in which the input image belongs.

V. EXPERIMENTAL RESULTS

Firstly, in order to evaluate the learning capability of the neural model used by the proposed approach, a comparison with the standard sigmoidal neural network was carried out. The run consists in training multilevel sigmoidal neural network and traditional sigmoidal neural networks with the same numbers of network parameters. Fig. 7 shows the plots of the Mean Squared Error (denoted as MSE) during the learning phase.

Fig. 5. Block diagram of the proposed approach

In Fig. 5 above, the symbols $c_1, c_2 \ldots, c_n$ refer to the classes or categories that images are classified into. In the following subsections, the main steps of the proposed approach mentioned above are briefly explained.

A. Pre-processing

A preliminary step of creating a symbolic representation of the source images is required before applying any data classification methods. The images are thus normalized by bringing them to a common resolution, performing histogram equalization and applying the median filter to remove small distortions without reducing the sharpness of the image.

B. Feature extraction

In this step, we utilize 2D multilevel Haar wavelets transform to decompose an image. Each level of decomposition gives two categories of coefficients, i.e., approximate coefficients and details coefficients as shown in fig. 6. Approximate coefficients and color moments are both considered as the features for our classification problem.
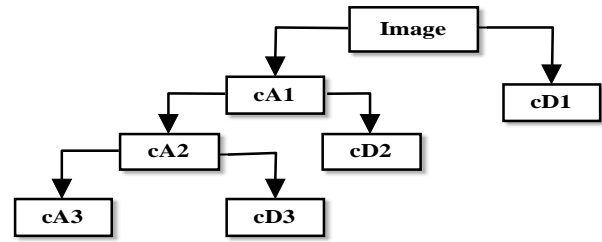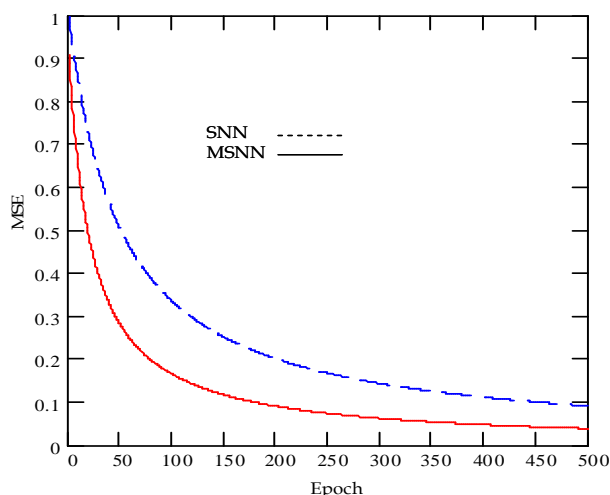
Fig.7: Averaged learning curve comparison between MSNN and SNN models.

The above mentioned approach for image classification has been tested using a test set containing 500 images which covers 5 categories, Coast, Mountain, Car, Farm, and Building as shown in Fig. 8. We used 20% of images from each category for training to get the category property, and the 80% of remaining images are used for testing. The classification results are shown in Table 1.

| Coast | Mountain | Car | Farm | Building |

Fig. 8. Categories of the image database used as a test set.

TABLE I
CLASSIFICATION RESULTS FOR TEST IMAGES COLLECTION

Categories	Precision	
	SNN	MSNN
Coast	0.92	0.98
Mountain	0.76	0.78
Car	0.79	0.84
Farm	0.95	0.99
Building	0.80	0.82
Average:	84.4	88.2

The figures in the above table conspicuously demonstrate that the MSNN model does better in the classification domain than SNN counterpart. The proposed approach was implemented in C/C++ codes. The implementation carries out the classification in few seconds using Desktop PC with Intel® Pentium Dual-Core processor and Windows Vista operating system.

VI. CONCLUSION

In this paper, we have proposed a method for image classification. The proposed method made use of a neural network model called Multilevel Sigmoidal Neural Network (MSNN). The main property of this neural model is the low computational complexity as well as the easiness of implementation. The simulation results on image classification shows that the MSNN model is very effective in terms of learning capabilities and classification accuracies. Although the proposed method has dealt with five-category database, it can be straightforwardly extended to deal with database of higher number of categories.

REFERENCES

[1] Goodrum, "Image Information Retrieval: An Overview of Current Research", *Special Issue on Information Science Research*, vol. 3, no. 2, 2000.
[2] N. O' Connor, E. Cooke, H. Le Borgne, M. Blighe, and T. Adamek," The aceToolbox: Lowe-Level Audiovisual Feature Extraction for Retrieval and Classification". *Proc. of EWIMT'05*, Nov. 2005.
[3] Deng, H. and D.A. Clausi,"Gaussian MRF Rotation-Invariant Features for SAR Sea Ice Classification," *IEEE PAMI*, 26(7): pp. 951-955, 2004.
[4] R. Zhao and W. I. Grosky, Bridging the Semantic Gap in Image Retrieval, Distributed Multimedia Databases: Techniques and Applications, T. K. Shih (Ed.), Idea Group Publishing, Hershey, Pennsylvania, pp. 14-36, 2001.
[5] J. Luo, and A. Savakis, "Indoor vs Outdoor Classification of Consumer Photographs using Low-level and Semantic Features," *Proc. of ICIP*, pp.745-748, 2001.
[6] A.K. Vailaya, Jain, and H.-J. Zhang, "On Image Classification: City Images vs. Landscapes," *Pattern Recognition Journal*, vol. 31, pp 1921-1936, December, 1998.
[7] J. Z. Wang, G. Li, and G. Wiederhold, "SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries," *In IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, pages 947-963, 2001.
[8] S. Prabhakar, H. Cheng, J.C. Handley, Z. Fan Y.W. Lin, "Picture-graphics Color Image Classification," *Proc. of ICIP*, pp. 785-788, 2002.
[9] Hartmann and R. Lienhart,"Automatic Classification of Images on the Web," *In Proc of SPIE Storage and Retrieval for Media Databases*, pp. 31-40, 2002.
[10] S. W. Kuffler and J. G. Nicholls, "From Neuron to Brain," (Sinauer Associates, Sunderland, 1976; Mir, Moscow, 1979).
[11] S. Bhattacharyya and P. Dutta, "Multiscale Object Extraction with MUSIG and MUBET with CONSENT: A Comparative Study," *Proceedings of KBCS 2004*, pp. 100-109, 2004.
[12] Yu, H., Li, M., Zhang, H.-J., Feng, J., Color texture moments for content-based image retrieval, In*: Internat. Conf. on Image Processing*, vol. 3, pp. 929-932, 2002.
[13] Der-Chiang Li and Yao-Hwei Fang, "An algorithm to cluster data for efficient classification of support vector machines," *Expert Systems with Applications*, vol. 34, pp. 2013-2018, 2008.
[14] R. Marmo *et al.* "Textural identification of carbonate rocks by image processing and neural network: Methodology proposal and examples," *Computers and Geosciences*, 31, pp. 649–659, 2005.