

# Fuzzy Image Classification Using Multiresolution Neural Networks

## with Applications to Remote Sensing

Yannis S. Avrithis and Stefanos D. Kollias

Electrical Engineering Department, National Technical University of Athens,

Zographou 15773, Greece

e-mail: iavr@image.ntua.gr

### Abstract

*Recent progress in supervised image classification research, has demonstrated the potential usefulness of incorporating fuzziness in the training, allocation and testing stages of several classification techniques. In this paper a multiresolution neural network approach to supervised classification is presented, exploiting the inherent fuzziness of such techniques in order to perform classification at different resolution levels and gain in computational complexity. In particular, multiresolution image analysis is carried out and hierarchical neural networks are used as an efficient architecture for classification of the derived multiresolution image representations. A new scheme is then introduced for transferring classification results to higher resolutions based on the fuzziness of the results of lower resolutions, resulting in faster implementation. Experimental results on land cover mapping applications from remotely sensed data illustrate significant improvements in classification speed without deterioration of representation accuracy.*

## 1. Introduction

Recent advances in supervised image classification have shown that conventional ‘hard’ classification techniques, which allocate each pixel to a specific class, are often inappropriate for applications where mixed pixels are abundant in the image [1]. A typical application where the classes may not be considered to be discrete and mutually exclusive, is land cover mapping from remotely sensed data. For instance, in a remotely sensed image, there may be a continuous and gradual transition between land cover classes resulting in regions of mixed pixel composition. In addition, due to limited image resolution, pixels often represent ground areas which comprise two or more discrete land cover classes [2].

For this reason it has been proposed that fuzziness should be accommodated in the classification procedure so that pixels may have multiple or partial class membership [3,4]. In this case a measure of the strength of membership for each class is output by the classifier, resulting in a ‘soft’ classification technique. However, since most classification techniques comprise three common stages, namely, the training, the allocation and the testing stage, fuzziness should be incorporated in all three stages [5], as it will be explained in the next section.

Traditionally, image classification is performed by a maximum likelihood, or Bayesian classifier, which assigns the most likely class to the observed data, and is known to be optimal if the assumptions about the probability density functions are correct. Neural networks, on the other hand, with their ability of learning, have no need of assumptions about the probabilistic model. Moreover, recent results on classification of multisource

remote sensing data [6] and on multispectral Landsat images [7,8], have shown that neural networks are often able to perform better than maximum likelihood classifiers. For the above reason they are adopted in this paper.

The use of neural network classifiers in conjunction with multiresolution image analysis has led to efficient algorithms for image processing and analysis [9] as well as to the derivation of representations of images appropriate for invariant recognition [10]. This work has motivated us to exploit the fuzziness of the image classification in each resolution level and to derive an efficient scheme for fuzzy image classification using multiresolution neural network classifiers.

## 2. Fuzzy Image Classification

In this section, the process of incorporating fuzziness in the training, allocation and testing stages of a feed-forward neural network classifier will be described. Conventionally, in supervised classification (statistical or with neural networks) each pixel of an image is assigned a unique class according to a specified criterion, i.e., the class with which it displays the highest *a posteriori* probability of membership, or the one that corresponds to the maximum output of a neural network.

However, by taking the activation level of a neural network output node as a surrogate for the proportional coverage of a pixel associated with that node, mixed pixels may be relatively easily accommodated in training a neural network. This is achieved by defining the class composition of each training pixel, as the desired output from the network for that pixel [5]. The network may then be trained as usual, with a learning algorithm

that attempts to minimise the error  $E$  between the desired and actual network outputs  $d_i$  and  $o_i$ , typically expressed as

$$E = \frac{1}{2} \sum_{p=1}^P \sum_{i=1}^C (d_i(x_p) - o_i(x_p))^2 \quad (1)$$

where  $x_p$  is the  $p$ th training sample,  $P$  is the number of training patterns and  $C$  is the number classes.

In the *training stage* of the classification, where pixels of known class membership are identified and used to characterise the classes, the key is to include in the training set an adequate number of mixed pixels with their associated class compositions.

In the *allocation stage*, where class characterisations are used to allocate pixels of unknown class membership, fuzziness is accommodated by using the activation levels of the output nodes (which always lie in the area [0,1] for the sigmoid activation function) as measures of the strength of class membership that reflect the class composition.

Finally, in the *testing stage*, an efficient method should be used for evaluating the classification accuracy. Conventional approaches based on confusion matrices, such as the percentage of correct allocation are applicable only to ‘hard’ classification. For this reason several approaches are suggested in [3] and [5] of which the most common is the Euclidean distance

$$\bar{D} = \frac{1}{N} \sum_{n=1}^N D_n, \quad D_n^2 = \sum_{i=1}^C (d_i(x_n) - o_i(x_n))^2 \quad (2)$$

where  $d_i(x_n)$  and  $o_i(x_n)$  are the desired and actual network outputs, respectively, i.e., the two classification measures to be compared,  $x_n$  is the  $n$ th evaluation sample and  $N$  is the number of evaluation patterns, i.e., the number of pixels in the evaluation image.

It should be noted that fuzziness can be equally well accommodated in the Bayesian classifier using similar arguments. Other techniques have also been proposed for obtaining directly a fuzzy classification such as the fuzzy  $c$ -means clustering algorithm [3].

### 3. Multiresolution Decomposition

In order to train a neural network to classify remotely sensed data, the land cover composition of training pixels should be known [1], and operationally this composition could be derived from a finer spatial resolution image co-registered to the image to be classified. It may also be possible in the near future to make use of data acquired from satellite sensors at two or more spatial resolutions simultaneously.

It is evident therefore that starting from an image with known ‘hard’ classification (spatial distribution of the classes represented by fraction images depicting the proportional coverage of a class in each pixel) to apply multiresolution decomposition to both the original image and its classification representation in order to de-

rive the fuzzy classification data that will be used as a training set for the neural network classifier.

Representation of signals at many resolution levels has gained much popularity especially with the introduction of the discrete wavelet transform, implemented in a straightforward manner by filter banks using quadrature mirror filters (QMF) [11]. Given an  $N \times N$  image  $x_0$  and using appropriate FIR perfect reconstruction filters  $h_L(n)$  (low pass) and  $h_H(n)$  (high pass), we can split the image into four  $(N/2) \times (N/2)$  images. In this way the *approximation* image at the lower resolution  $j=-1$  is derived by application of the low pass filter  $h_L(n)$  in the horizontal and vertical direction of the image:

$$x_{-1}^{LL}(m,n) = \sum_{k=1}^{N/2} \sum_{l=1}^{N/2} h_L(2m-k)h_L(2n-l)x_0(k,l) \quad (3)$$

Similarly, three *detail* images denoted as  $x_{-1}^{LH}$ ,  $x_{-1}^{HL}$  and  $x_{-1}^{HH}$  are obtained by applying all other possible combinations of the low and high pass filters in the horizontal and vertical directions.

Perfect reconstruction of the original image  $x_0$  can be achieved through synthesis of all subband components. It is, however, possible to obtain an approximate reconstruction of the original image by using only the low resolution component  $x_{-1}^{LL}$ :

$$\hat{x}_0(m,n) = \sum_{k=1}^{N/2} \sum_{l=1}^{N/2} h_L(m-2k)h_L(n-2l)x_{-1}^{LL}(k,l) \quad (4)$$

If the image decomposition procedure described in Eq. (3) is successively applied to the approximation images, a *multiresolution approximation* of the original image is obtained, consisting of images of continuously decreasing size. In this case, there exist optimal design techniques for analysis and synthesis filters, in terms of minimisation of the mean square error between the original image and its low resolution representation. Such a technique is used in our experiments, and efficiently implemented using auto-associative linear networks [10].

### 4. Hierarchical Neural Network Classification

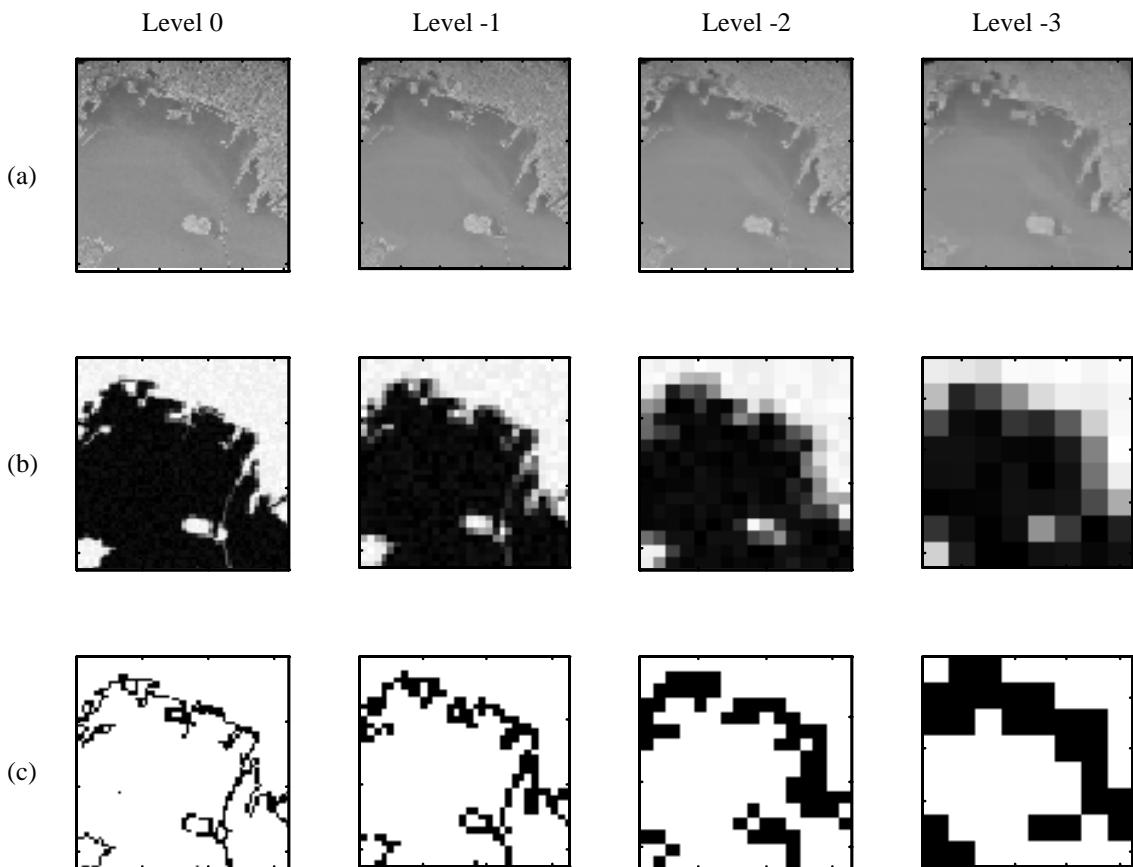
Hierarchical neural network classifiers can then be used as an efficient scheme for classifying the resulting multiresolution representations. A feedforward multilayer network is first used to classify an approximation image of quite low resolution, trained by some backpropagation variant. The hierarchical network is then recursively constructed to handle the image at higher resolution levels.

In the *training stage*, training starts at resolution level  $j \leq -1$  to classify approximation images  $x_j^{LL}$  and the performance is tested using a validation set of approximation images of the same resolution. Training then continues at resolution level  $j+1$  using as much as possible from the ‘knowledge’ acquired by the former problem

at level  $j$ . Transfer of a generally large number of (already computed) weights between the input and the first hidden layer of the low resolution network in corresponding positions of the high resolution network is permitted, so that these weights can remain fixed during training of the high resolution network, as described in detail in [10].

Moreover, in the *allocation stage*, since each pixel of a resolution level  $j$  approximation image substantially corresponds to four pixels of level  $j+1$  approximation image, and since a ‘hard’ or ‘crisp’ classification result for a pixel of the level  $j$  image will definitely lead to an analogous classification for all four corresponding pix-

els of the level  $j+1$  image, the fuzziness of the classification results of level  $j$  can be used as a measure for selectively continuing or not continuing classification at level  $j+1$ . In particular, classification can stop at level  $j$  when classification results near ‘0’ or ‘1’ are derived and all corresponding pixels of level  $j+1$  are assigned the same class compositions results. If, however, classification results are fuzzy, i.e., not near ‘0’ and ‘1’, classification for the corresponding pixels should be carried out at the next level  $j+1$ . Due to the recursive nature of the image classification process, a significant improvement in terms of computational complexity is to be gained by this selective classification algorithm.



**Figure 1:** Classification results for resolution levels 0 through -3: (a) input images to the neural network classifier, (b) activation level at the output of the classifier (sea in black and land in white), and (c) images illustrating the regions where the classification procedure should terminate (in white), or continue at the next resolution level (in black).

## 5. Experimental Results

The proposed technique was tested on an application of land cover mapping from aerial images for the simple case of classification into two classes: land and sea. Each image was segmented in blocks of 8x8 pixels and the 64 DCT coefficients of each block were used for classification of the block. Consequently, classification was performed in the DCT frequency domain and not in the more correlated 2-D image space. A land cover map was manually derived through visual analysis of the images resulting in a ‘hard’ classification (corresponding class membership ‘0’ for sea and ‘1’ for land) at resolu-

tion level 0. The fuzzy class membership of the 8x8 blocks was then derived as the approximation of the original ‘classification’ images at resolution level -3. Multiresolution approximation representations were derived for the original as well as for the ‘classification’ images using the optimal technique described in [10] for resolution levels  $j=0$  through -3.

Training of a feedforward neural network with one hidden layer was then performed at resolution levels 0 through -3 with the backpropagation algorithm, using as inputs the 64 DCT coefficients of each block and as output the corresponding fuzzy class memberships. Care

was taken so as to include in the training set enough representative blocks of each class, as well as mixed blocks which appear particularly near imprecise or fuzzy class boundaries. Training was carried out independently at each resolution level although this could be more efficiently done using the hierarchical approach outlined in the previous section and more thoroughly described in [10].

Figure 1(a) depicts an image of size 512x512 pixels used for evaluation of the classification mechanism, at resolution levels 0 through -3. In Figure 1(b) the corresponding classification results are illustrated. The classification accuracy was similar to the one derived in [5]. In particular, the Euclidean distance  $\bar{D}$  between the desired and actual classifications, defined in Eq. (2), was used as a measure of classification accuracy, resulting in a mean accuracy of  $\bar{D} = 0.62$  for the whole evaluation set at resolution level 0.

Finally, the recursive classification technique is illustrated in Figure 1(c). For each resolution level a new image is constructed, based on the fuzziness of the corresponding classification results. Specifically, white regions of the images of Figure 1(c) correspond to ‘hard’ classifications, i.e., to classifications for which the output activation level lies in the regions  $[0,t]$  or  $[1-t,1]$ , where  $t$  is in general a small threshold. For these regions, the classification procedure stops at the current resolution level and pixels in the same regions of higher resolution representations are assigned the same activation level as a classification result. Black regions, on the other hand, correspond to ‘soft’ classification and thus for those regions the classification procedure continues at the next resolution level. Using  $t=0.1$ , only 500-600 blocks out of  $64 \times 64 = 4096$  were required for the classification of a 512x512 image, resulting effectively in an 85% reduction in the computational time, while classification accuracy was only slightly decreased ( $\bar{D} = 0.67$ ).

## 6. Conclusions

A multiresolution hierarchical neural network approach to supervised classification has been presented, exploiting the fuzziness of the classification results in order to selectively perform classification at different resolution levels and gain in computational complexity. Multiresolution approximations have been obtained in an optimal way and hierarchical neural networks have been used for classification of the derived representations. It has been indeed demonstrated that transferring classification results to higher resolutions, based on the fuzziness of the results of lower resolutions, can lead to faster implementation. Promising experimental results have been presented, illustrating significant improvements in classification speed without deterioration of representation accuracy.

Extensive research is currently carried out on the potential application of the derived fuzzy classification technique on other problems, such as efficient indexing and

content-based retrieval of images and video sequences from large databases.

## References

- [1] G. M. Foody, “Land cover mapping from remotely sensed data with a neural network: accommodating fuzziness,” *Proc. 1<sup>st</sup> COMPARES Workshop*, York, UK, July 1996.
- [2] G. M. Foody and D. S. Boyd, “Observation of temporal variations in environmental gradients and boundaries in tropical West Africa with NOAA AVHRR data,” *Proceedings RSS’ 96: Science and Industry*, Remote Sensing Society, Nottingham, UK.
- [3] G. M. Foody, “Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data,” *Int. J. Remote Sensing*, vol. 17, no. 7, pp. 1317-1340, 1996.
- [4] G. M. Foody, R. M. Lucas, P. J. Curran, and M. Honzak, “Non-linear mixture modelling without end-members using an artificial neural network,” *Int. J. Remote Sensing*, vol. 18, no. 4, pp. 937-953, 1997.
- [5] G. M. Foody and M. K. Arora, “Incorporating mixed pixels in the training, allocation and testing stages of supervised classifications,” *Pattern Recognition Letters*, vol. 17, pp. 1389-1398, 1996.
- [6] J. A. Benediktsson, P. H. Swain and O. K. Ersoy, “Neural network approaches versus statistical methods in classification of multisource remote sensing data,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 28, no. 4, pp. 540-552, July 1990.
- [7] H. Bischof, W. Schneider and A. J. Pinz, “Multispectral classification of Landsat - images using neural networks,” *IEEE Trans. Geoscience and Remote Sensing*, vol. 30, no. 3, pp. 482-490, May 1992.
- [8] H. Bischof and A. Leonardis, “Optimising neural networks for land use classification,” *Proc. 1<sup>st</sup> COMPARES Workshop*, York, UK, July 1996.
- [9] S. D. Kollias, “Efficient processing and analysis of images using neural networks,” *Proc. 1<sup>st</sup> COMPARES Workshop*, York, UK, July 1996.
- [10] S. D. Kollias, “A multiresolution neural network approach to invariant image recognition,” *Neurocomputing*, vol. 12, pp. 35-37, 1995.
- [11] S. G. Mallat, “A theory for multiresolution signal decomposition: the wavelet representation,” *IEEE Trans. PAMI*, vol. 11, no. 7, pp. 674-693, 1989.