Abstract
The use of neural networks to classify land-cover from remote sensing imagery relies on the ability to determine a winner from the candidate land-cover types based on the imagery information available. In the case of a “winner-takes-all” scenario, this does not allow us a measure of how much the prediction of each pixel’s land-cover can be trusted. We present a three-stage method where only winning candidates which are given a clear lead over the other land-cover types are accepted, with a neighborhood relationship and the application of mixed pixels being used to provide full classification. This method allows us to place more faith in the resulting map than simply taking the winner, and results in a higher accuracy of classification. The method is applied to Landsat imagery of an area of the Philippines where natural, urban, and cultivated land-cover types exist.

Introduction
Outline
Neural networks (NNs) have been proven capable of assisting land-cover mapping in cases where the number of land-cover classes described is not very great (i.e., <20 classes). Aitkenhead and Wright (2004) demonstrated a rapid and effective land-cover mapping method using the backpropagation neural network (a common neural network training approach which applies an error-minimization technique in order to adjust the connection weightings) and training areas of known land-cover type, for a total of 12 human and natural classes. Mendoza et al. (2004) used a remote sensing/neural network-based method to detect land-cover change through deforestation and agriculture in Amazonia, while Curran et al. (2000) demonstrated a link between disease vectors and land-cover, and showed how remote sensing information could be used to develop disease risk maps.

The use of neural networks is not restricted solely to providing a winner-take-all classification of basic remotely-sensed imagery. Linderman et al. (2004), in mapping bamboo distribution, successfully achieved the detection and recognition of understory vegetation through an overlying canopy, demonstrating that neural networks can operate successfully on noisy and disrupted datasets where the question being asked is not one of straightforward classification. Houet et al. (2003) used a neural network to monitor the ratio of vegetation cover to bare ground at different scales over a period of time, taking advantage of the utility of NNs in providing continuous variable prediction rather than all-or-nothing classification. This aspect of the NN paradigm is utilized here for a different reason, that of measuring the reliability of the pixel-by-pixel classification.

Laporte et al. (2003) used a neural network approach to integrate different data types for land-cover classification. Neural networks are very effective when applied to datasets containing more than one type of data, for example, spectral reflectance values and terrain physical parameters. All that is required is that the values given to the network are adjusted to lie on roughly the same scale, usually between 0 and 1. Similarly, Zhu et al. (2000) used a Bayesian approach to provide a priori information for a classifier using Landsat TM data.

The degree of accuracy of land-cover maps has generally in the past only been measured after the fact, when the mapping process is complete. However, recent work has focused more on determining the accuracy of applied methodologies as they are being carried out, and using this determination to adjust or compare methods. Brown (2003) examined the uncertainty within land-cover prediction results derived from remote sensing imagery, while Liu et al. (2004) presented a method which not only provided a land-cover classification from remote sensing data, but also provided a measure of how uncertain the classification was based on disagreement between different methods.

Kasetkasem et al. (2003) approached the problem of uncertain classification through a sub-pixel classification approach, and attempted to produce super-resolution land-cover maps using Markov methods. The method relied on the assumption that correlations existed between neighboring pixels in terms of class type. The same assumption is used here for dealing with problem pixels where there is no clear winner based solely on spectral information. Alimohammadi et al. (2004) discuss the importance of uncertainty as an indicator of how well remotely-sensed data has been interpreted, and of how much it can be trusted. The introduction of ancillary data, such as neighborhood information, was again shown to improve the classification accuracy. The problem however, lies with situations where the remote sensing data is all that is available. Remotely-sensed imagery is nowadays easy and inexpensive to obtain, whereas other types of spatial information which could be applied as ancillary data may be less so.

Civco et al. (2002), in a review of several different methods of land-cover classification from remote sensing, concluded that no single method exists which can accurately
work in all situations. A synthesis approach where several methods are applied in turn is more likely to resolve the situation, as each method has particular strengths which allow a subset of the problem to be dealt with. Steele and Patterson (2001) used an ensemble method, which relies on resampling of the remotely sensed data and development of new rules based on additional data.

Avrithis and Kollias (1997) used neural networks to provide fuzzy classification for remote sensing data, which is one way of dealing with the uncertainty caused by mixed classifications. In cases where a specific classification is desired, another method is to accept only those pixel classifications where there is a clear winner.

Methods and Results

Location and Imagery

The Landsat MSS image obtained is 1,151 \times 1,198 pixels and covers part of the province of Laguna, in the southeastern corner of Luzon, the largest, most populous island of the Philippines. The study area lies between 14° 05’ 07” and 14° 23’ 25” N, and between 121° 23’ 24” and 121° 41’ 25” E and is characterized by a highly dissected mosaic of urban, semi-urban, rural, and disturbed natural forest land-use typical of large areas throughout the Philippines. Dominant natural features in the image include the southeastern portion of the 900 km² Laguna de Bay (LDB) in the north, and the dormant volcano, Mt. Banahao, which rises to 2,170 m in the south. The Balanac-Pagsanjan Rivers comprise the major drainage systems, traversing the upland-lowland transect from south/north to empty into LDB. Santa Cruz, the provincial capital on the southern shores of LDB, and the municipal towns of Luchan, Majayjay, Magdalena Luisiana, and Lilio are the most populated urban areas, although numerous small barangays (villages) and rural settlements are scattered throughout the study area, particularly adjacent to paved roads. Agriculture is dominated by smallholder cropping of irrigated rice and vegetables in the lowlands, while terraced rice, vegetable production, and small formal plantations, and widespread scattered plantings of coconut and other minor tree crops occupy increasing areas of the steep uplands.

Aims

In this work, a neural network is used instead of a Bayesian network as the inputs given to the classifier system (spectral reflectance and measures of land-cover type presence) are continuous rather than discrete, a situation which NNs lend themselves to more easily than Bayesian statistics. A method of developing ancillary neighborhood data is demonstrated which relies on the remote sensing information, with the proviso being that if other data were to be made available, the accuracy of the system would doubtless benefit.

The use of neighborhood data developed from the original imagery allows us to carry out an ancillary information-based approach by using the pixels which have been confidently classified as an additional data source to bootstrap a method of classification based on neighboring pixels. The method of confidently classifying particular pixels relies on a thresholding method where only pixels where there is a candidate land type which wins by more than a set margin is accepted.

Methods using a combination of spectral and spatial information have been demonstrated in the literature (Berberoglu et al., 2000; Chen et al., 1997; Kimes et al., 1999). However, these methods usually involve the measurement of textural information, and the application of this alongside spectral information in a standard winner-takes-all technique, rather than using the two-stage method outlined here of first identifying those pixels that can be confidently classified, and then using this information to develop neighborhood information.

Candidate Thresholding

The backpropagation neural network training method is one that can be readily applied to the identification of land-cover classes from multispectral imagery. Selected training areas with known land-cover were identified by in-field georeferenced ground-observations and expert knowledge and are used in the adjustment of the neural network’s weights. This is followed by identification of the unknown pixels by passing their reflectance information through the trained network.

Land-cover types for classification included urban, rice cultivation (cropped, bare, flooded), vegetable cultivation (cropped, bare), coconut, forest (trees present and forest cleared), surface water, cloud, and shadow. The six input nodes supplied the three pixel reflectance values (red, blue, and green) (a) in their absolute values scaled in the range \([0,1]\), and (b) as a proportion of the total reflectance values for each pixel. The use of proportional, as well as absolute values, was intended to solve problems caused by inconsistent illumination over the image, for example by having slopes of different aspect which would receive different amounts of sunlight from one another. Kolmogorov’s theorem for neural network modeling states that the number of nodes in the hidden layers should be twice the maximum in the input or output layers in order to guarantee the perfect fit of any continuous function (Bishop, 1995). Therefore, the neural network used in this case had six input nodes, two hidden layers of twenty-six nodes each and thirteen output nodes each corresponding to one of the land-cover classes used. An additional class of “no image” was used as where there were areas on the image that were blank, and for which the pixel reflectance values were all set to zero. This class is ignored in later sections as it is not useful to the statistical evaluation of the method.)

The neural network was used to provide an output weighting for each of the candidate land-cover types. A standard method of classifying land-cover using neural networks is to take the candidate output node with the highest weighting as the one that is classified on the image, but this does not take into account the fact that there is an unavoidable and unpredictable error in the activation of
each output node, and that a secondary candidate might have activation only just below that of the winner. In order to avoid errors caused by this assumption, a threshold difference value was assigned to the network which meant that only winning candidates which had a weighting more than that of the second-place candidate plus this threshold difference were accepted. For example, if the threshold difference was set at 0.2 and the winning output node had a value of 0.7, then the pixel would only be considered classified if the second-place candidate had a value of less than 0.5. From visual examination of sets of output activation values and the range that they could take, it was assumed that sensible threshold difference values would lie in the range 0.1 to 0.5, where the range in possible output node values lay between 0 and 1. Figure 1 shows the proportion of the image which remained unclassified using different values of hidden layer size and difference thresholds. In each case, the number of training steps used was 100,000. Two conclusions can be drawn: first, that there is a large variation in network performance between using five nodes per hidden layer and using ten or more nodes per hidden layer; secondly, winner threshold values below 0.1 or above 0.6 (approximately) are likely to be ineffective in producing classified images. Below a threshold of 0.1, the proportion of pixels remaining unclassified is lower than the expected error rate of an NN system, and so a single winner-takes-all method would be just as effective. Above a threshold of approximately 0.6, the large proportion of pixels unclassified will make the application of the secondary method detailed below impractical and will result in large unclassified patches.

**Neighborhood Data and Secondary Network**

In order to classify the image pixels for which there had been no clear winner, it was necessary to develop a method which relied not only on the raw pixel values in question. A secondary neural network was used for which the outputs were the same (i.e., weightings for each of the land-cover classes), but which had additional inputs corresponding to the presence of specific land-cover types in the immediate neighborhood (the eight cells surrounding each pixel). The training data for this network was derived from the results of the first network, which had been confidently predicted by the thresholded winner-take-all network being selected at random from the image. The presence or absence of each land-cover class in any of the pixels neighboring the selected pixel was determined (The importance was not attached to which of the neighbors contained a particular land-cover class; only the presence in any neighboring pixel was taken into account.) and used as additional input data, with 100,000 training steps being used with a hidden layer size of 38 (using the same rule of having twice as many nodes in each hidden layer as there are input nodes, which in this case is six for the spectral information, and thirteen for the land-cover classes). The trained network was then used to classify those pixels which the first network had been unable to identify with sufficient confidence, using the same winner threshold value as in the first classification step. It was necessary to run the second classification step more than once in order to eliminate unclassified pixels for which there were originally no classified neighbors, and for which the secondary network would only make a difference once neighboring pixels had been classified. The process was repeated until no more pixels were identified as being clear winners by the network (the number of repetitions required to reach this point ranged from 0 for a threshold value of 0 to a maximum of 18 for a threshold value of 0.5).

**Mixed Pixels**

Following the second pixel classification effort for the image, there remained a small number of pixels which could still not be identified. The number of these varied between approximately 1.4 and 2.5 percent of the total image for the different winner threshold values used, with no apparent relation between winner threshold value and number of problem pixels. As the primary and secondary methods were both incapable of categorizing these problem pixels, a third method of classifying them was required. It was assumed that because no clear winner could be identified either using a thresholded spectral classification method, or using both spectral and neighborhood information, the problem pixels had a higher probability than those previously classified of being composed of mixed pixels. As the majority of these pixels were assumed to be a mixture of the two land-cover classes with the most weighting and were classified as such, using the secondary neighborhood-trained neural network. The number of mixed classes ranged between 29 and 70. Figure 2 demonstrates the three-step process of (a) classification using a thresholded winner-take-all neural network, (b) classification of the majority of the remaining pixels using an NN trained with a combination of spectral and neighborhood data, and (c) categorization of mixed pixels.

**Comparison**

Land-cover classification was carried out over the whole image using the above three-step method for a range of winner threshold values between 0 and 0.5. Figure 3 shows the results of comparison of the predictions with 423 field observations, showing that the highest levels of accuracy were achieved for a winner threshold value of 0.25, which required six iterative steps to identify each pixel. The additional amount of time taken by the system, following identification of the training pixels, to automatically select training data for the secondary neural network, train this network, and then classify the remaining pixels was relatively small compared to the amount of time required to classify the entire 1.4 megapixel image using a simple winner-takes-all NN method with the same 12 land-cover classes (61 minutes as compared to 54 minutes to complete the map) on a 2.8 GHz desktop PC running Microsoft® Windows® XP 6.0).

Figure 4 shows the results of the primary neural network classification for a winner threshold of 0.25, with unclassified pixels marked in black. This Figure gives us an indication of the number and distribution of pixels that cannot be confidently classified using a straightforward backpropagation neural network. Figure 5 shows the results of further
classification using the secondary, neighborhood-trained network, while Figure 6 gives the final classified image with mixed pixels and legend. Table 1 gives a confusion matrix for the 423 field observations, providing an indication of which classes the method tended to confuse with one another. This table contains only the classes that are obtained from observations, and which are not mixed, along with the two classes of cloud and shadow obtained from examination of the image. The overall accuracy of the system was 0.913 over all ground observations, with accuracy levels ranging from 0.85 to 0.96 for individual classes. The number of observations for each class was not identical, resulting in a slightly higher value for the overall accuracy than if the mean of individual classes was taken, which was 0.911.

Using only the winner-takes-all method with no thresholding (i.e., the traditional single NN method), the prediction accuracy is 0.844.

Further statistical analyses of the system accuracy can be obtained through the use of Kappa Index of Agreement (KIA) measurements, which provides a superior indication of the ability of the system to categorize pixels than straightforward prediction accuracy. The KIA value (K) can be given by Equation 1:

\[
K = \frac{A - E}{N - E}
\]

where \(A\) is the number of samples of that class predicted accurately, \(E\) is the number that would be predicted accurately by random class selection, and \(N\) is the total number of samples of that class. The value of \(K\), which lies between
0 and 1, can be calculated both for individual classes and for all classes. Table 2 gives the value of $K$ for each main class, and for all ground observations. As can be seen, the difference here between predicted accuracy and $\kappa$ is small for each class.

Examination of the land-cover classification obtained using winner threshold values higher than 0.25 showed large clusters of unclassified pixels following classification with the primary neural network. These clusters were eliminated by the secondary network, but tended to result in large homogeneous patches of classified land-cover. This would suggest that the secondary network, trained using both pixel reflectance values and neighboring land-cover type information, is effective at eliminating small areas of unclassified pixels but should not be trusted for larger areas. Visual examination of the heterogeneity of the classified images for different winner thresholds appears to provide evidence for this, although it has not been quantified.
the system’s ability to classify land-cover types against existing boundary definitions.

Discussion
The aim of this paper was to demonstrate the ability of a combination of two methods in improving the utility of neural networks for land-cover classification from remote sensing. These two methods, the introduction of a threshold requirement for the winning candidate land-cover class and the use of neighborhood information to bootstrap an additional dataset for identifying problem pixels, are not difficult to implement. We believe that the level to which they improve the accuracy of the system (the increase in accuracy from 0.844 to 0.913 almost halves the error rate from 0.156 to 0.087) more than balances any additional complexity of the method. This conclusion is reached by comparing the reduction in error rates (44 percent) with the increase in time taken (approximately 13 percent).

The methods outlined here also provide an indicator of land-cover map reliability, and in comparison with a

Figures 7 and 8 show (a) an expanded area of the original image alongside, and (b) the same area as classified according to land-cover. This allows visual comparison of
standard winner-takes-all neural network classification, provide more reliable land-cover mapping. The use of a threshold to define a clear winner, and the subsequent use of a second network using neighboring classified cells, is almost as rapid as a single network method while adding to the reliability of the system. As applied here to a landscape containing a mixture of obscured, managed, and natural pixels, it has been shown capable of identifying regions of problem pixels; thus, not only of improving the accuracy of the map but also of providing us with a way of revising the training data used in order to avoid mixed or problem pixels.

An additional advantage to using neural networks for land-cover classification is their versatility with multiple data sets. The work carried out here used only remote sensing imagery with three spectral bands, but if additional datasets (e.g., DEM, population density, temperature) had been available, it would have required little or no effort to integrate this information with the imagery and produce a training data set that could be used to develop a more sophisticated neural network land-cover mapping system. Any factors that can be meaningfully converted to values, whether as measures of some variable or as presence/absence indicators, could be included using this method. The use of winner thresholds could also be applied to additional datasets, allowing for (a) a succession of procedures which could eliminate more and more problem pixels with reliable predictive ability, and (b) an investigation into which factors provide the strongest indicators of specific land-cover types.

The addition of more information is also related to the issue of network scale. As the network node population increases, the system's performance tends to improve, but the computational cost also increases. Generally, a balance must be struck between accuracy and the time taken to produce a land-cover map, taking into account available computer power, pixel count, number of classes, and required accuracy. Neural networks are relatively computationally expensive for pixel-by-pixel land-cover classification, particularly as the number of classes increases (see below), and so they are possibly more useful as a method of providing broad categorization prior to the application of more rapid, class-specific, and expert knowledge-based methods (e.g., decision trees and Bayesian statistics). Different training algorithms and node activation functions will also play a part in determining how well the system performs in a specific situation, but difficulties remain in knowing which method to apply in a particular situation.

Potential problems with this method include the transferability of the winner threshold from one image to another. It may be necessary to calibrate the threshold for different images, although we feel that the range within which the winner threshold will be most effective is likely to lie between 0.1 and 0.3 for most cases. There is no direct proof for this, however, and varying classification systems and degrees of image quality will result in a range of values for the threshold variable. In addition, the optimal value of this threshold may vary with the number of inputs or outputs of the neural network. However, in comparison
to the normal winner-takes-all approach, it is felt that due to its nature of only accepting clear winners, this method will provide accurate results even when the number of classes becomes as high as 20 or 30, provided the imagery is of sufficient quality (the availability of multispectral imagery is extremely important, although textural information may also be obtained from sufficiently high-resolution imagery). Beyond this number of classes, confusion between spectrally similar land-cover classes and the increased computational requirements of the system (which get larger as the square of the number of hidden pixels in a two-hidden-layer NN) will, it is felt, restrict the applicability of the system.

References


(Received 04 August 2005; accepted 08 November 2005; revised 07 December 2005)