JARQ 49 (4), 377 - 381 (2015) http://www.jircas.affrc.go.jp

Extreme Learning Machine-based Crop Classification using ALOS/PALSAR Images

Rei SONOBE[1]*, Hiroshi TANI[2], Xiufeng WANG[2], Yasuhito KOJIMA[2] and Nobuyuki KOBAYASHI[3]

[1] Faculty of Agriculture, Shizuoka University (Shizuoka, Shizuoka 422-8529, Japan)
[2] Research Faculty of Agriculture, Hokkaido University (Sapporo, Hokkaido 060-8589, Japan)
[3] Smart Link Hokkaido (Iwamizawa, Hokkaido 068-0034, Japan)

Abstract

Classification maps are required for agricultural management and the estimation of agricultural disaster compensation. The extreme learning machine (ELM), a newly developed single hidden layer neural network is used as a supervised classifier for remote sensing classifications. In this study, the ELM was evaluated to examine its potential for multi-temporal ALOS/PALSAR images for the classification of crop type. In addition, the k-nearest neighbor algorithm (k-NN), one of the traditional classification methods, was also applied for comparison with the ELM. In the study area, beans, beets, grasses, maize, potato, and winter wheat were cultivated; and these crop types in each field were identified using a data set acquired in 2010. The result of ELM classification was superior to that of k-NN; and overall accuracy was 79.3%. This study highlights the advantages of ALOS/PALSAR images for agricultural field monitoring and indicates the usefulness of regular monitoring using the ALOS-2/PALSAR-2 system.

Discipline: Agricultural engineering
Additional key words: Hokkaido, machine learning, sigma naught

Introduction

Land-cover classification is one of the most common applications of remote sensing. Crop-type classification maps are useful for yield estimation and agricultural disaster compensation, in addition to the management of agricultural fields. Optical remote sensing remains one of the most attractive options for the accumulation of biomass information and forest monitoring (Samreth *et al.* 2012, Sarker and Nichol 2011). In addition, while optical satellites such as ALOS/AVNIR-2 (Sonobe *et al.* 2014a), Landsat (Hartfield *et al.* 2013), MODIS (Sakamoto *et al.* 2009), and NOAA (Hirano and Batbileg 2013) have been employed in the identification of species and conditions of vegetation, cloud cover significantly limits the number of available optical images, radar is unaffected by cloud cover or low solar zenith angles (Bindlish and Barros 2001). And significant information about soil and vegetation parameters can also be obtained through microwave remote sensing (Sonobe *et al.* 2014b). These techniques are employed increasingly to manage land and water resources for agricultural applications (Sonobe *et al.* 2014c). The number of studies on rice and wheat monitoring and mapping using SAR data has increased, and some studies utilizing multi-temporal SAR data have reported high correlations between backscattering coefficients, and plant height and age (Chakraborty *et al.* 2005, Sonobe *et al.* 2014d, Waisurasingha *et al.* 2008). These examples highlight possible uses in the area of agricultural management, specifically for the identification of paddy fields. They indicate the potential use of SAR data for the discrimination of crop types. Furthermore, a number of examples have shown the practical usefulness of supervised learning for classification (Sonobe *et al.* 2014b). This paper reports a comparison of crop classification using PALSAR data performed by extreme learning machine (ELM) and k-nearest neighbor algorithm (k-NN).

*Corresponding author: reysnb@gmail.com
Received 12 May 2014; accepted 4 February 2015.

Materials and Methods

The experimental area of this study was agricultural fields in the western Tokachi plain, Hokkaido, Japan (142°55′12″ to 143°05′51″ E, 42°52′48″ to 43°02′42″ N). The mean size of fields was 2.16 ha (the largest and smallest field areas were 18.0 and 0.01 ha, respectively).

The characteristics of the satellite images used in this study are shown in Table 1. PALSAR operates in the L band (1.27 GHz) and has five operational modes: FBS, a single-polarization high-resolution strip with a 70-km-wide swath; FBD, a dual-polarization receiver (HH+HV, with HV for transmission with horizontal polarization and reception with vertical polarization) enabling deforestation monitoring; SCANSAR, which has a 350-km-wide swath for quick deforestation and sea-ice monitoring; PLR, for the clarification of scattering mechanism; and DSN, which is a reduced-resolution strip-mode. Generally, only the FBS and FBD modes were adopted, and HH polarization data were regularly obtained. Furthermore, PALSAR backscattering coefficients were a powerful tool for biomass estimation (Kiyono *et al.* 2011). The PALSAR data (level 1.5 products) were converted from digital number images to sigma naught images, which are the radiometrically calibrated power images referenced to the ground, using the PALSAR Level 1.1/1.5 product format description procedures (JAXA 2008).

Reference data was provided by Tokachi Nosai as a polygon shape file in which the position of the fields and attribute data, such as crop type, were included. Any field below 0.5 ha was deemed too small for analysis and removed from the vector maps. All remaining fields were buffered inward 25 m, taking into account field shape. The buffer was used to avoid the selection of training pixels from the edge of a field, which would create a mixed signal and affect assessment accuracy. Then, mean sigma naught values were calculated (Bargiel and Herrmann 2011, Sonobe *et al.* 2014b).

ELM was applied using MATLAB and Statistics Toolbox Release 2014b (The MathWorks, Inc., Natick, Massachusetts, United States), and we used the source code published by Nanyan Technological University (http://www.ntu.edu.sg/home/egbhuang/index.html). Furthermore, we generated the code for K-fold cross validation. k-NN was applied using R version 3.0.0 (R Core Team 2013) and IBk function included in the RWeka package (Hornik *et al.* 2009). We used a stratified random sampling approach to select fields used for training (Foody 2009), and 20% of crop fields were selected at random as training samples (Hartfield *et al.* 2013). The stratified random sampling is a method of sampling that involves the division of a population into smaller groups known as strata. In stratified random sampling, the formation of strata is based on shared members attributes or characteristics. The remaining 80% of the fields were used for accuracy assessment. The classification maps was evaluated in terms of overall accuracy (OA), producer's accuracy (PA), user's accuracy (UA), and the kappa index of agreement (κ). PA is the result obtained from dividing the number of correctly classified fields in each crop type by the number of reference fields. This value represents how well reference fields of the crop cover type are classified. UA is computed by dividing the number of correctly classified fields in each crop type by the total number of fields that were classified in that crop type. It represents the probability that a field classified into a given crop type actually represents that crop type.

In this study, the ELM was selected as a classification algorithm, and the six crop types were identified using 10 PALSAR images. The training data comprised 555 fields (71 bean, 86 beet, 60 maize, 66 potato and 200 winter wheat, and 72 grasses) and the test data comprised 2217 fields (281 bean, 346 beet, 239 maize, 263 potato and 801

Table 1. Characteristics of satellite data

Date	ALOS/PALSAR mode	Off-nadir angle (°)	Pixel spacing (m)	Orbit	Path	Frame
03 May, 2010	FBS	34.3	6.25	Descending	55	2750
04 May, 2010	FBD	34.3	12.5	Ascending	396	850
21 May, 2010	FBD	34.3	12.5	Ascending	397	850
18 June, 2010	FBS	34.3	6.25	Descending	55	2750
19 June, 2010	FBD	34.3	12.5	Ascending	396	850
06 July, 2010	FBD	34.3	12.5	Ascending	397	850
03 August, 2010	FBS	34.3	6.25	Descending	55	2750
18 September, 2010	FBS	34.3	6.25	Descending	55	2750
19 September, 2010	FBD	34.3	12.5	Ascending	396	850
06 October, 2010	FBD	34.3	12.5	Ascending	397	850

winter wheat fields, and 287 grasses). The ELM-based classification approach is based on the use of a single hidden layer neural network (Huang *et al.* 2006). For the training data having K number of samples, represented by $\{x_i, y_i\}$, where $x_i = [x_{i1}, x_{i2}\ldots, x_{ip}]^T \in R^p$ is the input vector and $y_i = [y_{i1}, y_{i2}, \ldots, y_{iq}]^T \in R^q$ is the target vector and R^p and R^q are p- and q-dimensional vector spaces over R. A single hidden layer neural network having H hidden neurons and an activation function $f(x)$ can be represented as:

$$\sum_{j=1}^{H} \alpha_i f(w_i x_j + c_i) = y_j \qquad (1)$$

where $w_i\, x_i$ is the inner product of w_i and x_i, w_i and α_i are the weight vectors connecting inputs, the i th hidden neuron, the i th hidden neuron and output neurons, respectively; and c_i is the complex bias of the i th hidden neuron. The ELM algorithm works for any infinitely differentiable activation function $f(x)$. These activation functions include the sigmoidal functions as well as the radial basis, sine, cosine and exponential functions. In this study, a sigmoidal activation function based ELM was applied. The parameter (H) was determined by K-fold cross validation technique using training data selected by stratified random sampling. This technique repeatedly generates training and test data sets from populations with a known land cover class membership. This data are separated into equally-sized K subsets (here, K = 10). A classifier is trained using nine subsets and is tested using the excluded one subset. Accuracy measures are obtained by interchanging test data and the resultant ten accuracy measures are averaged. The cross validation process including separation is then repeated 10 times.

For comparison with the result of ELM classification, we conducted k-NN, one of the traditional classification algorithms. This non-parametric procedure was introduced by Fix and Hodges (1951), and has since become well-known in the pattern recognition literature as the voting k-NN rule. The k-NN classifier assigns a class to unclassified data using its k-NNs in the training set. Cover and Hart (1967) have provided a statistical justification of this procedure by showing that, as N samples and k both tend toward infinity in such a manner that k/N→0, the error rate of the k-NN rule approaches the optimal Bayes error rate. Beyond this remarkable property, the k-NN rule owes much of its popularity in the pattern recognition community to its good performance in practical applications. However, in the finite sample case, there is no guarantee that the voting k-NN rule is the optimal way of using the information contained in the classified data in the neighborhood (Denoeux 1995). RWeka automatically finds the best value for k between 1 and 20; and the K value is determined by 10 iterations.

The significant differences among the results of the two classifications were determined at a 95% level of significance using the Z-test, which is performed for pairwise comparison of the proposed methods and takes into account the ratio between the difference values of two κ coefficients and the difference in their respective variances (Congalton and Green 2008).

Results and Discussion

For application of the ELM, the number of hidden nodes was tuned using training data and the K-fold cross validation technique. Fig. 1 represents the relationships between the number of hidden nodes and the averaged accuracy rate calculated using 10-fold cross validation. Higher accuracy was observed when the number of hidden nodes was 450, thus the value was adopted for ELM classification. For application of k-NN, The proper K value ranged from 5 to 10, and the average value was 8 by cross validation for RWeka (Table 2). In this study, 8 was selected as the k value

The corresponding confusion matrices of classifications using PALSAR data are given in Table 3. In terms of the agreement indices, those of the ELM are superior to those of k-NN. However, in terms of the PA of beans, beet and grasses, the results of k-NN are superior to those of the ELM. Furthermore, in the case of the UAs of maize, potato and wheat, those of the ELM are inferior to those of k-NN.

We used the Z-test to compare the accuracy of classification methods because the same test data were used for each classification. The Z score was 2.44, indicating that the difference is significant. The variances and standard deviations related to Kappa statistics are given in Table 4. Unlike passive systems, synthetic aperture radar (SAR) systems are not dependent on atmospheric influences or weather conditions; thus they are especially suitable for multi-temporal classification approaches, and the advantage leads to high accuracy (Bargiel and Herrmann 2011). However, computational cost of a classifier, such as training and test time, often represents a significant proportion of the total cost in classifications using multi-temporal remote sensing data. The ELM is a fast algorithm among machine learning methods because it is based on the use of a single hidden layer neural network, and input weights and biases are randomly assigned and need not be tuned. This approach should be applicable to the generation of land cover/use classification maps, in particular, in the area of agricultural.

Conclusions

To generate classification maps, ten images from ALOS/PALSAR were used, and two algorithms, the ELM and k-NN, were applied. Although this study is simply a case in point and these results may not always be practical, the results of the ELM were superior to those of k-NN in

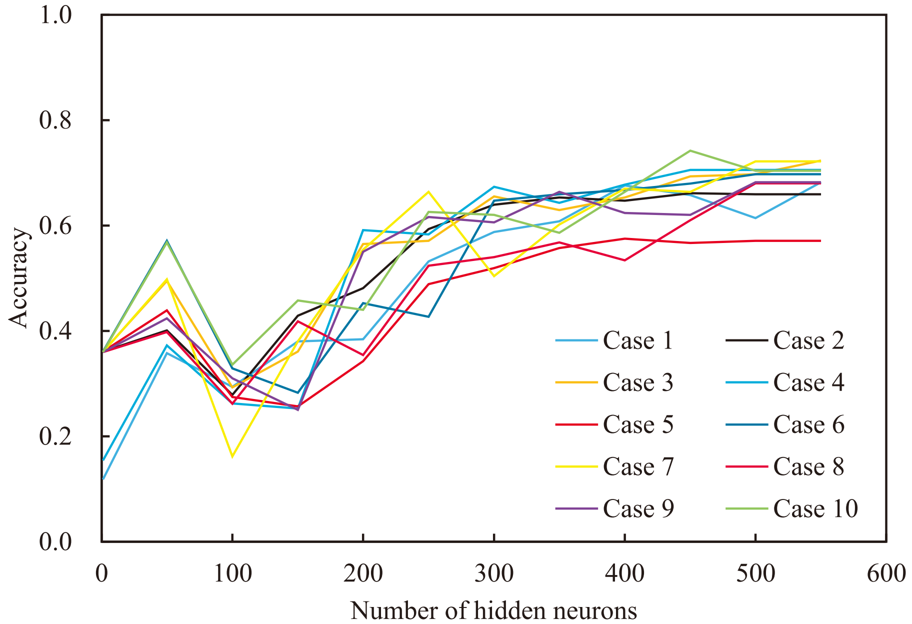

Fig. 1. Relationships between the number of hidden nodes and the averaged accuracy rate calculated using 10-fold cross validation

Table 2. Nearest neighbors and correctly classified instances

	Nearest neighbors	Correctly classified instances
Case 1	8	0.777
Case 2	5	0.766
Case 3	5	0.775
Case 4	10	0.760
Case 5	6	0.760
Case 6	10	0.791
Case 7	10	0.758
Case 8	8	0.762
Case 9	7	0.787
Case 10	6	0.760
Average	8	0.769

terms of overall accuracy and kappa. Furthermore, the ELM requires only one parameter whereas the support vector machine and neural network require a number of tuning parameters. In addition, the ELM algorithm is fast among machine learning methods because it is based on the use of a single hidden layer neural network, and input weights and biases are randomly assigned and need not be tuned. This approach should be applicable to the generation of land cover/use classification maps, in particular, in the area of agricultural.

This study also demonstrated the great potential of L-band SAR data for agricultural applications. The Advanced Land Observing Satellite-2 is a follow-up to the ALOS

Table 3. Classification accuracy

Class \ Algorithm	Extreme Learning Machine (ELM)	k-Nearest Neighbor (k-NN)
Producer's accuracy		
Beans	0.764	0.765
Beet	0.873	0.908
Grass	0.651	0.753
Maize	0.627	0.523
Potato	0.794	0.650
Wheat	0.874	0.833
User's accuracy		
Beans	0.726	0.655
Beet	0.916	0.781
Grass	0.721	0.643
Maize	0.640	0.683
Potato	0.703	0.810
Wheat	0.864	0.881
Overall acuracy	0.793	0.770
Kappa	0.737	0.709

Table 4. Variance and standard deviation related to the Kappa statistic

	Extreme Learning Machine (ELM)	k-Nearest Neighbor (k-NN)
Variance	0.00011	0.00012
Standard deviation	0.01040	0.01076

mission. For this new satellite, further studies testing the applicability of the method will be necessary. In addition, operational methodology for the mapping will be tested. Furthermore, cross-year classification may increase the ability to generate crop-type classification maps without concurrent training data and may prove useful in reducing labor costs for management in the area of agricultural and early information gathering. In future studies, the potential of PALSAR data for cross-year classification will be tested.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 26·5253. The authors also express their gratitude to the Japan Aerospace Exploration Agency (JAXA) for providing ALOS/PALSAR data.

References

Bargiel, D. and Herrmann, S. (2011) Multi-temporal land-cover classification of agricultural areas in two European regions with high resolution spotlight Terrasar-X data. *Remote Sens*., **3**, 859-877.

Bindlish, R. and Barros, A.P. (2001) Parameterization of vegetation backscatter in radar-based, soil moisture estimation. *Remote Sens. Environ*., **76**, 130-137.

Chakraborty, M. *et al.* (2005) Rice crop parameter retrieval using multi-temporal, multi-incidence angle Radarsat SAR data. *ISPRS J. Photogramm. Remote Sens*., **59**, 310-322.

Congalton, R. G. and Green, K. (2008) *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC Press, Florida, USA, pp.183.

Cover, T.M. and Hart, P.E. (1967) Nearest neighbor pattern classification. *IEEE Trans. Inform. Theory*, **13**, 21-27.

Denoeux, F. (1995) A k-nearest neighbor classification rule-based on Dempster-Shafer theory. *IEEE Trans. Syst., Man, Cybern*., Syst., **25**, 804-813.

Fix, E. and Hodges, J.L. (1951) Discriminatory analysis, nonparametric discrimination: Consistency properties. Technical Report 4, USAF School of Aviation Medicine, Randolph fields, TX.

Foody, G. M. (2009) Sample Size Determination for Image Classification Accuracy Assessment and Comparison. *Int. J. Remote Sens*., **30**, 5273-5291.

Hartfield, K. *et al.* (2013) Contemporary and historical classification of crop types in Arizona. *Int. J. Remote Sens*., **34**, 6024-6036.

Hirano, A. and Batbileg, B. (2013) Identifying Trends in the Distribution of Vegetation in Mongolia in the Decade after its Transition to a Market Economy. *JARQ*, **47**, 203-208.

Hornik, K *et al.* (2009) Open-Source Machine Learning: R Meets Weka. *Computational Statistics*, **24**, 225-232.

Huang, G. *et al.* (2006) Extreme learning machine: theory and applications. *Neurocomputing*, **70**, 489-501.

JAXA: ALOS PALSAR ALOS/PALSAR Level 1.1/1.5 product Format description. http://www.eorc.jaxa.jp/ALOS/en/doc/fdata/PALSAR_x_Format_EL.pdf.

Kiyono, Y. *et al.* (2011) Practicalities of non-destructive methodologies in monitoring anthropogenic greenhouse gas emissions from tropical forests under the influence of human intervention. *JARQ*, **45**, 233-242.

R Core Team, R: A language and environment for statistical computing. http://www.R-project.org/.

Sakamoto *et al.* (2009) Detection of Yearly Change in Farming Systems in the Vietnamese Mekong Delta from MODIS Time-Series Imagery. *JARQ*, **43**, 173-185.

Samreth *et al.* (2012) Tree Biomass Carbon Stock Estimation using Permanent Sampling Plot Data in Different Types of Seasonal Forests in Cambodia. *JARQ*, **46**, 187-192.

Sarker, L. R. and Nichol, J. E. (2011) Improved forest biomass estimates using ALOS AVNIR-2 texture indices. *Remote Sens. Environ*., **115**, 968-977.

Sonobe, R. *et al.* (2014a) Detection of ambrosia beetles using a pan-sharpened image generated from ALOS/AVNIR-2 and ALOS/PRISM imagery. *For. Syst*., **23**, 178-182.

Sonobe, R. *et al.* (2014b) Parameter tuning in the support vector machine and random forest and their performances in cross- and same-year crop classification using TerraSAR-X. *Int. J. Remote Sens*., **35**, 7898-7909.

Sonobe, R. *et al.* (2014c) Winter wheat growth monitoring using multi-temporal TerraSAR-X dual-polarimetric data. *JARQ*. **48**, 465-470.

Sonobe, R. *et al.* (2014d) Application of multi-temporal TerraSAR-X data to map winter wheat planted areas in Hokkaido, Japan. *JARQ*. **48**, 471-476.

Waisurasingha *et al.* (2008) Application of Remote Sensing and GIS for Improving Rice Production in Flood-prone Areas: A Case Study in Lower Chi-River Basin, Thailand. *JARQ*. **42**, 193-201.