Incoperating Texture in Remote Sensing Image Classification Using a MLP Deep Neural Network

Apiwat Lekfuangfu, Teerasit Kasetkasem
Department of Electrical Engineering
Kasetsart University
Bangkok, Thailand
Emails: apiwatlekfuangfu@gmail.com
fengtsk@ku.ac.th

Itsuo Kumazawa
Imaging Science and Engineering Laboratory
Tokyo Institute of Technology
R2-59, 4266 Nagatsuta-cho, Midori-ku
Yokohama, Japan
Email: kumazawa@isl.titech.ac.jp

Preesan Rakwatin
GISTDA
The Government Complex
Bangkok, Thailand
Email: preesan@gistda.or.th

Thitiporn Chanwimaluang
NECTEC
Pathumthani, Thailand
Email: thitiporn.chanwimaluang@nectec.or.th

Abstract—*In this paper, we investigated the use of texture information for image classification using deep-learning algorithm since intensity values alone cannot be used to separate one class from another. By incorporating texture information, we are able to significantly increase the classification performance.*

Keywords—Image classification; MLP; Texture extraction

I. INTRODUCTION

Image classification is very important tool in remote sensing image analysis. It enables the study of a large area on the ground relatively economically. As a result, there is a need for more accurate image classification techniques since remote sensing images often suffer from various factors. For instance, satellite images are very sensitive to the weather conditions such as cloud and haze.

Examples of image classification techniques developed specifically for remote sensing images can be seen in [1-2]. Jian Guo et al. [1] proposed to use MODIS data for the land cover classification data. They selected the area, which is located in Heilongjiang Province in China. They used Maximum Likelihood Classification (MLC), Self-Organized neural network (SONN) and Support Vector Machine (SVM) for classified the MODIS data. The result of image classification of MLC, SONN and SVM have accuracy 79.15%, 70.73% and 86.3% respectively. However, the data from MODIS is very sensitive with weather conditions. In the rainy season, the accuracy can greatly reduce.

The study by Jiming Li et al. [2] proposed to use Hyperspectral data for land-cover classification image. They used deep feature representation method to classifier for pixel-wise hyperspectral data classification. They proposed method has two main steps: principle components of the hyperspectral image cube are first filtered by three-dimensional Gabor wavelets. Second, they used stacked autoencoders are trained on the outputs of the previous step through unsupervised pretraining. Then, they used deep neural network is trained on those stacked autoencoders. However, the primary disadvantages of hyperspectral data are cost and complexity. Fast computers, sensitive detectors, and large data storage capacities are needed for analyzing hyperspectral data. Significant data storage capacity is necessary since hyperspectral cubes are large multi-dimensional datasets, potentially exceeding hundreds of megabytes. All of these factors increase the cost of acquiring and processing hyperspectral data.

The image classification algorithms described above used on the color information. However, there are many methods to extract features from image for classifier such as color feature, shape feature, texture feature, etc. The feature extraction techniques are selected to classification is deep learning. Hence, in this paper, we propose the use of texture information to train the deep learner to further increase the classification accuracy.

This paper is divided in four main sections. In section I, information of background study and satellite data. In section II, the methodology to classification land cover mapping by deep learning MLP neural network. In section IV, the result of this method compared with the ground data are shown.

II. BACKGROUND AND DATA

Currently, a subfield of machine learning, deep learning has attracted wide attention. In this paper we proposed algorithm for image classification based on multi-layer perceptron deep learning method, we begin the discussion by introducing the concept of multi-layer perceptron.

A. Multi-Layer Perceptron (MLP)

Multi-layer perceptron neural networks are feed forward networks with more layer of nodes between its input and output layer. Due to nonlinear activation function with each nodes, the MLP can accomplish complex decision making in the pattern space [3].

In this study, the most widely used supervised neural classifier, the MLP network, is used. MLP networks due to their generality, flexibility, and ability of handle nonlinear models. These networks have of a number of processing elements arranged into multiple layers, which is mathematical model to mimic the first-order behavior of the human nervous system. The general architecture consists of an input layer for the data, a hidden layer to apply some mathematical transformation, and an output layer to produce a label.

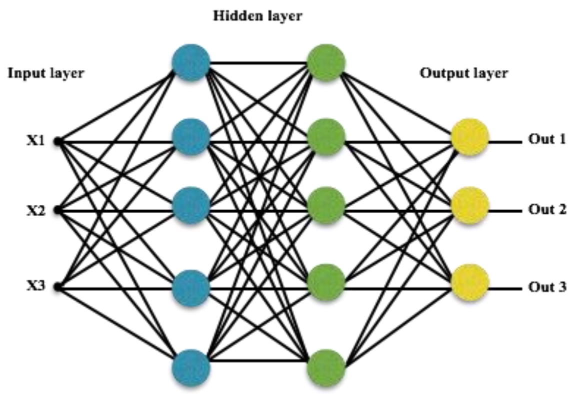

Figure 1: Multi-layer perceptron

In a multi-layer perceptron topology, neurons are grouped into distinct layers as depicted in Figure 1. Outputs of some neural are connected to inputs of other neural. Output of each layer is connected to input of nodes in the following layer. Inputs of the first layer are the inputs data to the network, while the outputs of the last layer, form the output of the network.

$$f(x) = G\left\{b^{(2)} + W^{(2)}\left(s\left(b^{(1)} + W^{(1)}x\right)\right)\right\} \quad (1)$$

In hidden layer of multilayer perceptron is a function $f: R^D \rightarrow R^L$, where D is the size of input x and L is size of the output $f(x)$. The parameter G and s are the softmax function and activity function in neural network respectively.

$$h(x) = \Phi(x) = s\left(b^{(1)} + W^{(1)}x\right) \quad (2)$$

where the parameter $h(x)$ is vector constituting the layer and $W^{(1)}$ is the weight matrix connecting the input vector to the hidden layer.

All in neural network connections have relationship of weights, which are usually randomized at the beginning of training data. When the data pass into a node through interconnections, they are multiplied by the weight associated with interconnections and calculation.

All in neural network connections have relationship of weights, which are usually randomized at the beginning of training process. When the data passes into a node through interconnections, they are multiplied by the weight associated with interconnections. After that, the activation function is used to determine the output value of this node. The activation function having a nonlinear character is very important since the nonlinearity enables the discrimination of the complex relationships that exist in the feature space. While linear functions are particularly used in input and output layers, non-linear activation functions can be used for hidden and output layers. The choice of activation functions may have mightily influence the complexity and performance of neural networks. The most commonly used nonlinear functional forms of activation functions satisfying the approximation conditions of activation functions are the sigmoid and hyperbolic tangent functions [4].

With the activation functions, the sigmoid and hyperbolic tangent functions, are the most widely used ones and they are defined as

$$\tanh(a) = \frac{(e^a - e^{-a})}{(e^a + e^{-a})} \quad (3)$$

and

$$\text{sigmoid}(a) = \frac{1}{(1 + e^{-a})} \quad (4)$$

From the above equation activation functions, the sigmoid function produces the output range of signal over the 0 to 1. The values 0 and 1 are obtained for only minus and plus infinities, respectively. As the output values come to close these limits, the derivations of this function decreases. On the other hand, the hyperbolic tangent function is a bipolar version of the sigmoid function. The hyperbolic tangent function produces the scaled output over the -1 to 1 closed range. The output values are obtained for minus and plus infinity, respectively. Because the output space of the tangent hyperbolic function is broader, it may be more efficient for the classification performance of the MLP.

B. Backpropagation

In layer of neural networks, we can train a network using gradient descent. This involves defining an error function, and then evaluating the derivatives $\partial E/\partial w_{ij}$. The evaluation of these error derivatives proceeds using a version of the chain rule of differentiation, referred to as back-propagation of error. The back-propagation is the multi-layer of neural network model, which used to calculate error between output and target of ground truth data.

$$E = \frac{1}{2}\sum_k (t_k - a_k)^2 \quad (5)$$

Note that it is only the outputs out of the final layer that appear in the output error function *E*. The parameter a_k is denotes an activation value of network and t_k is denotes a target value of classification. This research has six targets for learning image classification.

Then, we can use the error value into network for update weight of hidden layer.

$$w_{ij} = w_{ij} + \varepsilon \frac{\partial E}{\partial w_{ij}} \quad (6)$$

The subscript *j* and *i* denote are the hidden and output layer respectively. These equations constitute the Back-Propagation Learning Algorithm for Regression.

III. Methodology

As discussed earlier, feature method in the image classification can be achieved by MLP neural network. Next, we define method for improve efficiency for image classification by normalization and texture classification before take data into neural network.

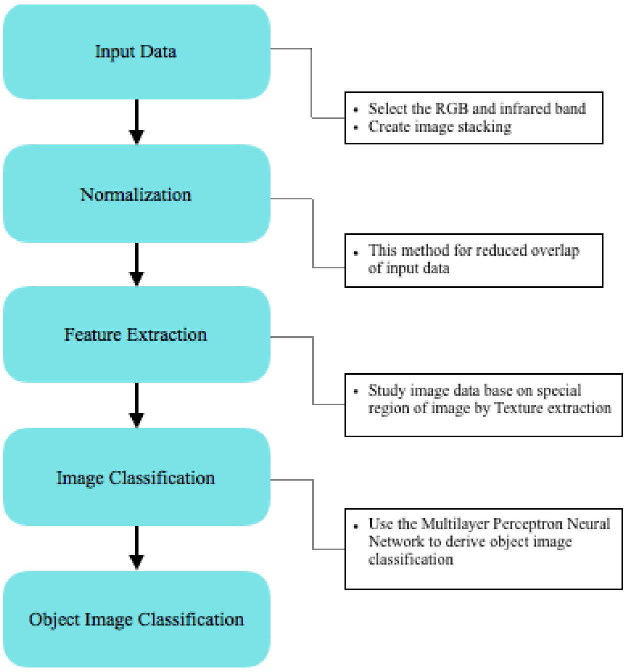

Figure 2: The diagram of the proposed algorithm

The steps of our algorithm are summarized in Figure 2. First, we have to select and input training samples. In order to assure the precision, we select training sample of each category through visual interpretation and field survey. At the beginning we use texture of image to interpret each ground object. In paper we use remote sensing image for classify the object. The examples of training samples can be categorized as shown in Figure 3.

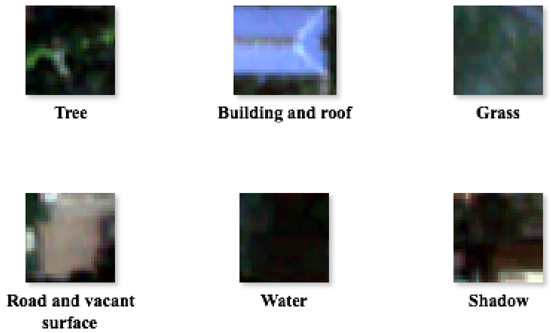

Figure 3: Training sample of each category

Because of the input images have many classes with different color, these high variations can decrease the performance of the MLP algorithm. As a result, the input data must be normalized first. The principles of normalization used to reduce duplication and possibility of error with data. The normalization for training samples and expected output vectors. Before we use MLP Neuron Network to train and classify the remote sensing image, in order to avoid the saturation in the process of network broadcasting, input vectors must be normalized. In this approach, the data is scaled to a fixed range usually 0 to 1. The cost of having this bounded range in contrast to standardization is that we will end up with smaller standard deviations, which can suppress the effect of outliers. Normalized formula [5]:

$$X_{norm} = \frac{(X - X_{min})}{(X_{max} - X_{min})} \quad (7)$$

In this formula, *X* means the original value of the input vector, *Xnorm* means normalized input value, *Xmin* and *Xmax* mean minimum and maximum value of the original entire training sample sets. Through this formula, value range of the input vectors is [0, 1].

The number of output layer neuron is determined by the classified number of the image to be classified. We use unit matrix to express expected output vector of training samples. One such example of expected output vector can be seen in Table I.

Input vectors	Expect vectors	Category
X_{norm_1}	{1,0,0,0,0,0}	Road and vacant surface
X_{norm_2}	{0,1,0,0,0,0}	Water
X_{norm_3}	{0,0,1,0,0,0}	Shadow
X_{norm_4}	{0,0,0,1,0,0}	Tree
X_{norm_5}	{0,0,0,0,1,0}	Grass
X_{norm_6}	{0,0,0,0,0,1}	Building and roof

Table I. Input vectors and expected output vectors.

To increase the performance of image classification even further, we incorporate textures into the MLP. To do this, instead of submitting only intensity vector of a pixels of interest into the image, we train our MLP with a block of neighboring pixels [6] as shown in Figure 4.

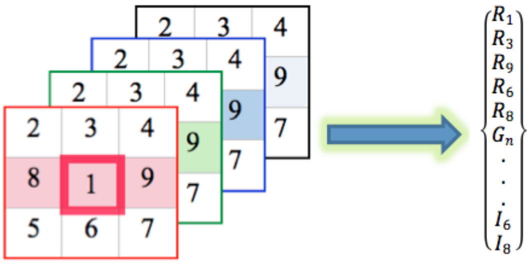

Figure 4: Set data based on texture before training

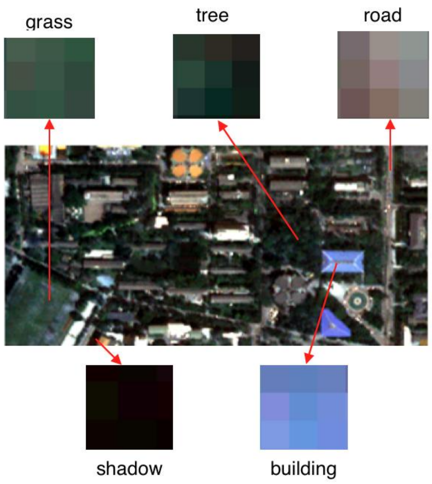

Figure 5: Block of neighbor from ground data

Next step, we used the Multi-layer perceptron neural network to classification image data. We must divide the data set for test and training in neural network. We defined the value of dense in neural network is a fully connected. In the first layer, we must specify the expected input data shape such as we have 20 input data that mean the first layer has 20 dimensional vectors.

Then, we used nonlinear logistic (sigmoid and hyperbolic tangent) to activation functions, which used effectively with MLP network for image classification. The activation function is one of the main characteristic elements of an artificial neural network.

Deep neural network with a large data of parameters are very powerful machine learning systems. However, over-fitting is a serious problem in neural network. Over-fitting has many effect for neural. Because of this problem, we used the dropout technique for addressing this problem. Dropout consists in randomly setting a fraction of input units to 0 at each update during training time, which are randomly drop units from the neural network during training. Dropout method used to prevents units from co-adapting too much. Then, we use fivelayers for training data until the final layer. Finally, we define the target path in final layer by ground truth data.

III. EXAMPLE

In this example, we examine the performance of our proposed algorithm on real dataset. Here, the MLP written in Python Language by the Theano and Keras deep learning library were employed. The input image is QuickBird data and day/night band (DNB) data provided by Geo-Informatics and Space Technology Development Agency (GISTDA), which is a fine resolution remotely sensed product available to the public through Digital Globe. The Quickbird image has been used to extract the information of density and height of ground, which has high-resolution (2.4 meter) for study area. The multispectral bands are red (630- 690 nm), green (520-600nm), blue (450-520nm) and near infrared (760-900nm). The study area covered of Kasetsart University [7].

When the Multi-layer perceptron neural network training success, we obtain MLP neural network in accordance with the ground truth of the image data. The land cover of image is large, which has range size is 300×150 pixel, which have 4 bands of input data as shown in Figure 6. The ground truth was manually labeled into six classes and it is given in Figure 7 where black, blue, dark green, bright green, white and red colors representing shadow, water, trees, grass, roof type 1 and root type 2 classes, respectively.

The day/night band (DNB) data is a new generation of night-time light data sensed by the Visible Infrared Imaging Radiometer Suite (VIIRS) at 750-meters resolution carried by the Suomi National Polar-Orbiting Partnership satellite. In this paper, we study area covered Bangkok, Thailand, was chosen because Bangkok is the capital of development of Thailand [8]. We classified urban and non-urban area from DNB data based on MLP neural network. The DNB image has range size 250×250 pixel, which have 7 bands of input data and ground truth (Figure 8.) have two classes are white and black colors representing urban area and non-urban area respectively.

Figure 6: True color composited of Quickbird image over the Kasetsart University area.

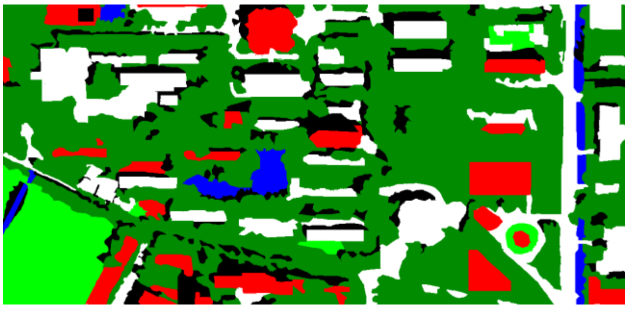

Figure 7: Ground truth image

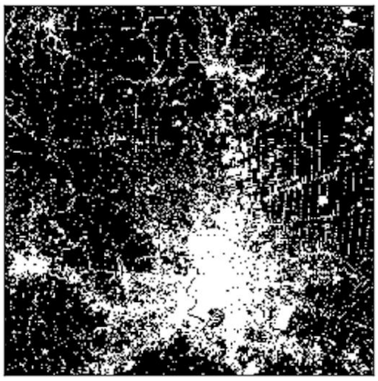

Figure 8: Ground truth of DNB image

The first step of our proposed algorithm is to normalized the remote sensing image using Eq. (7). The normalized data is shown in Figure 9.

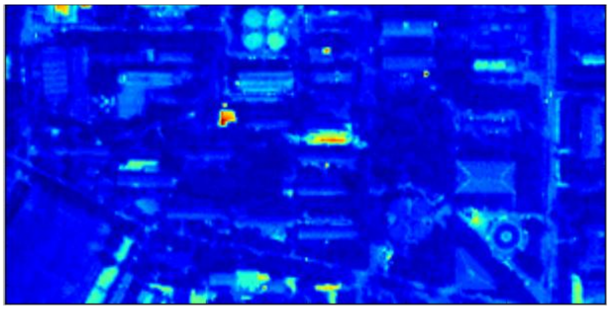

Figure 9: Image data of one band after normalization

The normalized image is submitted to the Multi-layer perceptron neural network using the training process and texture as described in the previous section and the classified map using our proposed method is shown in Figure 10. For comparison purpose, we also use the traditional MLP neural network without using any texture information, and the resulting map is shown in Figure 11. By visual comparison, we observed that your proposed method provides more connected map, especially for the Class of Grass and water.

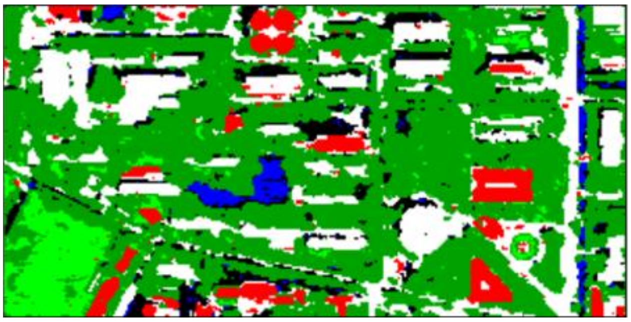

Figure 10: Classified Map using the proposed algorithm

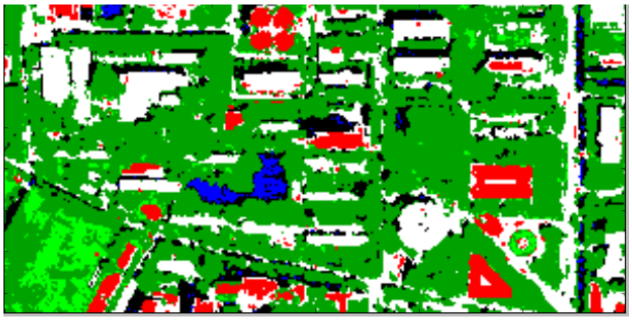

Figure 11: The classified map using the traditional MLP neural network

The result of day/night band data Figure 12 based on texture extraction and MLP neural network compare with Ground truth data (Figure 8), the proposed algorithm achieves the percentage of correctly classified pixels of 82.12%. On the other hand, the result without texture (Figure 13.) has accuracy 81.25%.

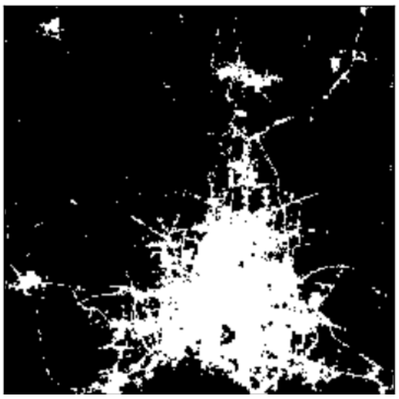

Figure 12: The classified DNB urban area based on MLP and Texture

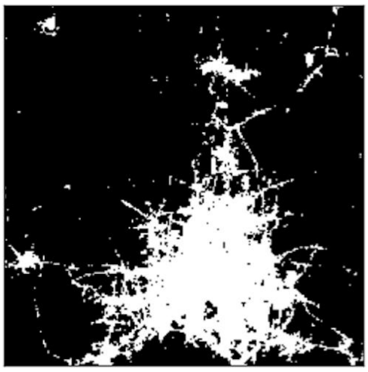

Figure 13: The classified DNB urban area without Texture

When compared the result of Quickbird image classification with Ground truth data (Figure 7), the proposed algorithm achieves the percentage of correctly classified pixels of 83% whereas the traditional approach can only achieve 79%. This result shows that the texture information can improve the classification accuracy of the land cover mapping. However, the results of DNB between using texture extraction (Figure 12) and without texture extraction (Figure 13) have a little change because the data of day/night band is very low texture in data. Therefore, texture extraction is very helpful for image classification in the data have many texture in image.

IV. CONCLUSION

We have proposed an approach to classify a land cover map. This article focuses on the incorporation of texture information in the MLP neural network. We found that the classification accuracy can be improved by incorporating the texture data into the MLP neural network.

ACKNOWLEDGMENT

This research is financially supported by Thailand Advanced Institute of Science and Technology (TAIST), Kasetsart University (KU) under the research grant from KURDI, and computer for running program by Geo-Informatics and Space Technology Development Agency (GISTDA).

REFERENCES

[1] Joo, J., Zhang, J., Zhang, Y., Cao, Y., *Study on the Comparison of the Land Cover Classification for Multitemporal MODIS Images, The International Workshop on Earth Observation and Remote Sensing Applications, June 2008, Pages 1 - 6.*

[2] L. Jiming, L. Bruzzone, S. Liu, *Deep Feature Representation For Hyperspectral Image Classification, IGRASS 2015 International Conference, Pages 4951-4953*

[3] D.R. Guha, S.K. Patra, *Cochannel Interference Minimization Using Wilcoxon Multilayer Perceptron Neural Network, International Conference on Recent Trends in Information, Telecommunication and Computing 2010, Pages 145 – 150.*

[4] C. Özkan, F.S. Erbek,, *The Comparison of Activation Functions for Multispectral Landsat TM Image Classification, Photogrammetric Engineering & Remote Sensing, November 2003, Pages 1225–1234.*

[5] J. Jiang, J. Zhang, G. Yang, D. Zhang, L. Zhang, *Application of Back Propagation Neural Network in the Classification of High Resolution Remote Sensing Image, Geoinformatics, 18-20 June 2010, Pages 1 - 6.*

[6] S. Pharsook, P. Larmsrichan, T. Kasetkasem, S. Siddhichai, *"The texture classification using the fusion of decisions from different texture classifiers", ECTI-CON, 17-19 May 2011, Page 1003 - 1006.*

[7] J.L. Wang, X.H. Wang, *Information Extraction of Building Height and Density Based on Quick Bird Image in Kunming, China, Urban Remote Sensing Event, 2009, Pages 1 - 8.*

[8] R. Jaturapitpornchai, T. Kasetkasem, I. Kumazawa, P. Rakwatin, T. Chanwimaluang, *"A level-based method for urban mapping using NPP-VIIRS nighttime light data", Information and Communication Technology for Embedded Systems (IC-ICTES), 2015, Pages 1 – 6.*