第31回日本情報地質学会講演会 講演要旨集 19-20頁 2020年

Land Use/Land Cover Classification using Light Convolution Neural Network: a Case Study in Lao Cai, Vietnam

Hang T. DO*, Venkatesh RAGHAVAN** and Go YONEZAWA***

*Graduate School for Creative Cities, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558-5858, Japan
, * Graduate School of Engineering, Osaka City University, 3-3-138 Sugimoto, Sumiyoshi-ku, Osaka 558-5858, Japan
Email: *dohanghumg@gmail.com, ** raghavan@osaka-cu.jp , *** yonezawa@osaka-cu.jp

Key words: LULC, LCNN, Remote sensing, RapidEye, Object-based

1. Introduction

Convolution Neural Network (CNN) is currently the state of the art for many remote sensing image classification tasks. By learning with multiple hidden layers, CNN models have shown outstanding accuracy in various applications (Masi et al., 2017, Rezaee et al., 2018). Traditional CNN models which include numerous convolutional layers are very time-consuming and require a huge training data and computational resources. In comparison, Light CNN (LCNN) with only few convolutional layers can achieve high accuracy in a short processing time by using a small number of training samples (Song et al., 2019). In this study, pixel-based and object-based LCNN is applied to establish Land Use/ Land Cover (LULC) map of Lao Cai area in Vietnam.

2. Methodology

2.1. Data and study area

The study area, a part of Lao Cai province, located in the North of Vietnam covers an area of approximately 525 km^2. The study area is the settlement area for ethnic minorities. The main LULC classes are water, built-up, mining/bare land, rice terrace, paddy field, non-forest vegetation and forest.

RapidEye's sensors produce imagery in five spectral bands at 5m resolution. In this study, 5 bands of RapidEye image acquired on 9th September 2014 covering the study area are used for LULC extraction. Reference polygon samples of the 7 LULC classes were collected based on visual interpretation of RapidEye image with verification using Google Map (Figure 1). The reference samples cover 82.29 km^2, which equals 14.9% study area.

2.2. Methodology

In CNN model, convolutional layer is the first layer to extract feature maps from an input image. In convolution, the filter moves through the entire input image with the moving step decided by setting stride hyperparameter. If the stride is 1, the filter is moved to 1 pixel at a time. Sometimes the filter does not perfectly fit the input image size. Zero-padding is to pad the input image with zeros so that the filter will fit the image.

Pooling layers are used to progressively reduce the number of parameters when the images are too large. In fully connected layer, the feature maps extracted from previous layer are flattened into one-dimensional vector and fed to a fully connected network. While the convolutional layers learn the spatial features, the fully connected layers learn the classification rule to extract feature vectors by using an activation function. The output of full connected layer is determined by using an activation function.

2.2.1. PBIA and OBIA convolutional layer

Basically, in first convolutional layer, large filter sizes could obtain better classification result than smaller filter sizes considering the relation of a number of neighbor pixels of input image. In this case, it is considered as patch-based approach, or OBIA. However, when large filter sizes are used, such as 5×5 and 7×7, the number of parameters and computation cost increase drastically. To resolve the problem, a 1×1 filter which generates only a single parameter or weight for each channel of the input image, and small filter size 3x3 were employed (Lin, et al., 2013). The 1×1 filter does not involve any neighbor pixels in the input but works with the individual pixel itself. In this case, the convolutional layer could be considered as PBIA. However, the layer needed to be combined a nonlinearity with other convolutional layers, allowing the projection to perform non-trivial computation on the input feature maps. In this study, CNN model with PBIA convolutional layer is considered as PBIA-CNN, while the model with OBIA convolutional layer is named OBIA-CNN.

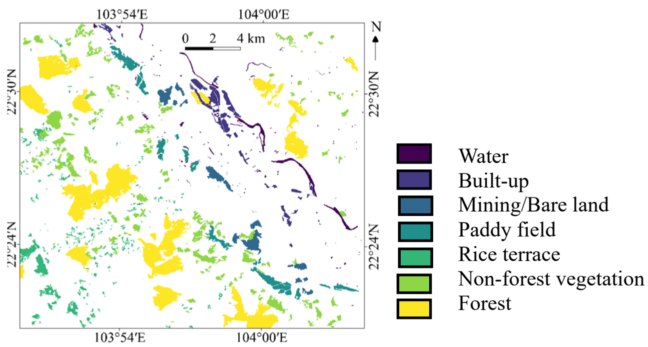

Fig 1. Reference map of this study

2.2.2. LCNN and OBIA-LCNN

In CNN model, when the number of network layers increase, the information in the neurons of the network is continuously combined. Eventually, the network extracts deep concepts and expresses abstract semantic features (Song et al., 2019). However,a deep network can lead to time-consuming and overfitting. To ignore these drawbacks, LCNN is applied for LULC classification in this study. The workflow of PBIA-LCNN and OBIA-LCNN are displayed in Figure 2. The two models have 3 convolutional layers. The first convolutional layer of PBIA-LCNN filters the 3D input with 20 filters of size 1×1×5. In case of OBIA-LCNN, the first convolutional layer includes 20 filters of size 3×3×5. The second and third convolutional layers of PBIA-LCNN and OBIA-LCNN have 20 filters of size 2×2×20. Zero padding and stride equals 1 are employed. The last Softmax layer which provides a probability distribution over 7 LULC classes. Fully connected layers and pooling layers are not employed. Instead, ReLU activation function and Adam optimizer with a learning rate of 10^{-5} are used. The number of epochs equals to 100 and early stopping technique is applied, the training processing stops when the different between two consecutive loss (r) is lower or equals to 10^{-6}. Google Colaboratory framework is selected to implement the models in this study.

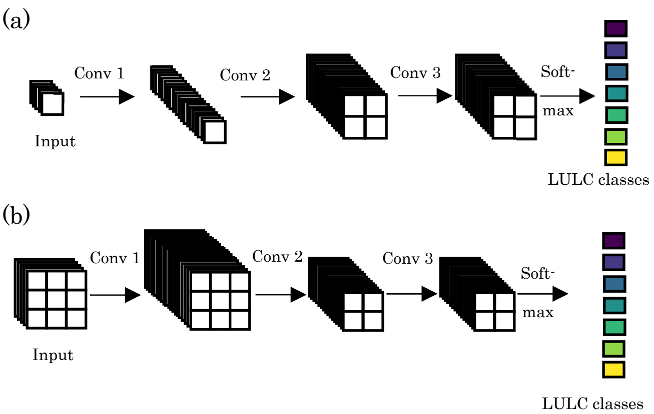

Fig 2. Architecture of (a) PBIA-LCNN, (b) OBIA-LCNN

Table 1: Classification accuracy (MB: Mining/bareland, NF: Non-forest vegetation)

LULC class	Accuracy (%)					
	PBIA			OBIA		
	PA	UA	OA	PA	UA	OA
Water	97	96	96	98	96	97
Built-up	94	91	92	94	92	93
MB	90	88	89	90	91	91
Paddy field	83	88	88	85	90	87
Rice terrace	95	94	95	97	94	95
NF	90	88	89	89	91	90
Forest	96	97	97	98	96	97
OA	**94**			**94**		

UA: User's Accuracy, PA: Producer's Accuracy, OA: Overall Accuracy

3. Result and Conclusions

Figure 3 shows the result LULC maps of PBIA-LCNN and OBIA-LCNN using RapidEye image. Table 4.1 gives producer's, user's and overall accuracy of the classification.

In PBIA, all LULC maps extracted from RapidEye image produce high overall accuracies (Table 1). Among them, water, rice terrace and forest achieve excellent accuracies, at 96%, 95% and 97%, respectively. Mining/bare land, built-up, paddy field and non-forest vegetation attain lower accuracies, at 89%, 92%, 88% and 89%, correspondingly. The difference between producer's accuracies and user's accuracies of all classes are small, less than or equal to 5%. Overall accuracy of the classification is 94%.

Classification accuracies of all LULC classes in OBIA are more than 87%. Similar to the PBIA, water, rice terrace and forest are the most accurate classes, at 97%, 95%, 97%, respectively, followed by mining/bare land, built-up, paddy field and non-forest vegetation, at 91%, 93%, 87% and 90%, respectively. In general, the difference between producer's and user's accuracies of individual LULC classes are small, no more than 5%. Overall classification accuracy is equal to PBIA, at 94%.

The result shows that both PBIA-LCNN and OBIA-LCNN are effective classification techniques for LULC mapping. Moreover, the LCNN models have capability of handling large RS datasets to aid in monitoring LULC change on a local as well as regional scale using RS images.

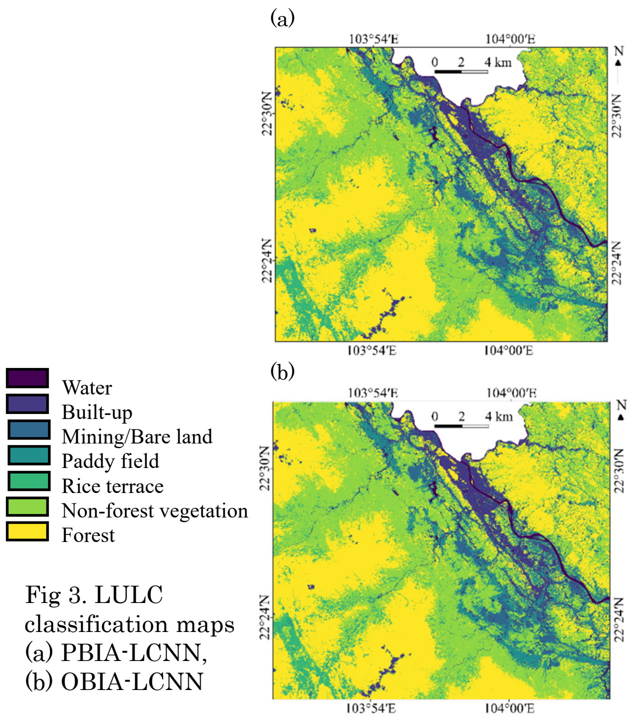

Fig 3. LULC classification maps (a) PBIA-LCNN, (b) OBIA-LCNN

References

Lin, Min, Chen, Qiang, and Yan, S. (2013) Network in network. *arXiv preprint* arXiv:1312.4400.

Masi, G., Cozzolino, D., Verdoliva, L. and Scarpa, G. (2017) CNN-based pansharpening of multi-resolution remote-sensing images, *2017 Joint Urban Remote Sensing Event (JURSE)*, pp. 1-4.

Rezaee, M., Mahdianpari, M., Zhang, Y. and Salehi, B. (2018) Deep convolutional neural network for complex wetland classification using optical remote sensing imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 9, pp. 3030-3039.

Song, H.; Kim, Y. and Kim, Y. A. (2019) Patch-Based Light Convolutional Neural Network for Land-Cover Mapping Using Landsat-8 Images. *Remote Sensing*. no. 11, 114.