EFFECTS OF CLASS PURITY OF TRAINING DATA ON CROP CLASSIFICATION USING 2D-CNN

Soyeon Park, No-Wook Park

Inha University, 100 Inha-ro, Michuhol-gu, Incheon 22212, Korea
Email: sypark531@inha.edu; nwpark@inha.ac.kr

KEY WORDS: deep learning, patch-based classification, training data

ABSTRACT: Proper collection of training data is an important factor in supervised classification for crop mapping. Each pixel in remote sensing imagery represents an area with various characteristics of surface objects and may have different spectral values for the same crop type. This mixed pixel effect in training data may greatly affect classification results. Although much effort has been made for the proper selection of training data in pixel-based classification, few studies have been conducted in patch-based classification with deep learning. In this study, we analyze the effect of class purity within the patch of training data on a patch-based 2D convolutional neural network (2D-CNN) model for crop classification. The classification performance of 2D-CNN was evaluated from two case study areas with different spatial characteristics of crops and input images with different spatial resolutions. In the area which consists of crop parcels with similar shapes and uniform patterns, the classification accuracy could be improved by collecting training samples with high class purity in the high spatial resolution imagery. On the contrary, using training samples with lower class purity in crop classification with Landsat images led to the improvement in the classification accuracy in the classification of areas where crop parcels had various shapes and sizes. These experimental results indicate that training data in the patch-based crop classification should be selected by taking into account the characteristics of the area to be classified.

1. INTRODUCTION

Crop type maps have been regarded as one of important factors in agricultural environment monitoring and crop yield prediction (Carranza-García et al., 2019; Kwak and Park, 2019). For the production of crop maps, remote sensing images have been widely used thanks to its ability to provide time-series information on various spatial scales (Kim et al., 2017). Deep learning algorithms such as convolutional neural networks (CNN), as well as conventional machine learning algorithms including support vector machine and random forest, have been applied to classification using remote sensing data (Mathur and Foody, 2008; Kussul et al., 2017). Unlike conventional pixel-based machine learning algorithms, CNN used patch-type inputs to extract optimal features for classification. Spatial information from neighboring pixels based on patches is known to be useful for classification of crop types with similar spatial features (Liang et al., 2016; Li et al., 2017; Kim et al., 2018).

The selection of training data is very important in supervised classification using remote sensing images. Since each pixel in the remote sensing image consists of several objects with different spectral characteristics (i.e., mixed pixels, Brown et al., 1999), we need the criteria for the proper selection of training data or pixels. In the pixel-based supervised classification, it is known that the acquisition of a large number of pure pixels is the goal of training, and that the class should be purely structured to properly train the model (Foody and Mathur, 2006). Conventional descriptive statistics including mean and standard deviation have been used to extract pure training pixels in the pixel-based classification (Ozdarici and Akyurek, 2012). Unlike the pixel-based classification, spatial correlation between neighboring pixels within patches greatly affects the classification performance. In particular, the composition of training data may change according to the class or spectral purity within patches. However, the effects of class purity of training data on patch-based supervised classification have not yet been fully investigated.

In this study, we quantitatively analyze how much the class purity within training data patches affects the performance of crop classification with patch-based 2D-CNN. Crop classification is carried out in two case study areas where crop types, the characteristics of crop parcels, and spatial resolutions of input images are quite different. Through this comparative case study, the criteria for the selection of training data in crop classification are discussed.

2. STUDY AREAS AND DATA

The two case study areas include Anbandegi in Korea and parts of Illinois state in USA. The Anbandegi area is a major highland Kimchi cabbage cultivation area in Korea. Three crop types including highland Kimchi cabbage, cabbage, and potato are cultivated in the area (Figure 1). Crop classification in Anbandegi was performed using a single-date unmanned aerial vehicle (UAV) image taken with an IXUS/ELPH drone and a Cannon S110 camera (Table 1). The image taken on August 25, 2017 which shows high vegetation vitality of the main crops was selected as input for classification. Training and test sets were extracted from ground truth data provided by the National Institute of Agricultural Sciences.

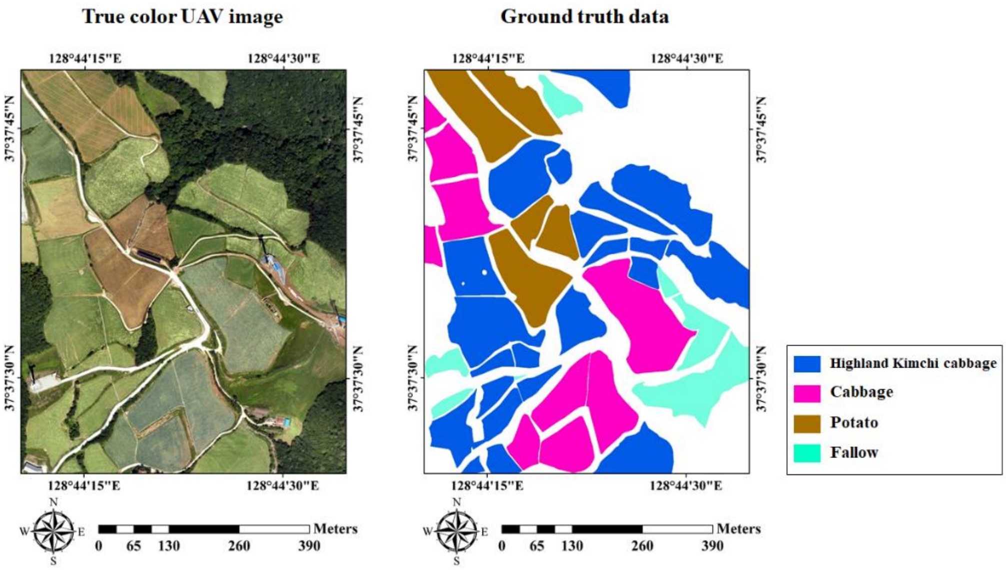

Figure 1. UAV imagery and ground truth data in Anbandegi.

Table 1. Summary of UAV and Landsat-8 OLI images used for crop classification

Category \ Region	Anbandegi	Illinois State
Image	UAV	Landsat 8 OLI
Spectral band	Blue, Green, Red	Red, NIR, SWIR
Spatial resolution	25cm	30m
Date	Aug. 25, 2017	Mar. 7, 2017
		Apr. 8, 2017
		May. 27, 2017
		Sept. 15, 2017
		Oct. 17, 2017

As the second study area, we chose a crop cultivation area near the downtown Waterloo in Illinois. Corn, soybean, and winter wheat are mainly cultivated in the study area (Figure 2). By considering the scales of the study area, time-series Landsat-8 OLI images were used to classify three crop types (Table 1). The cropland data layer provided by the USDA NASS was considered as ground truth and used for extracting training and test sets.

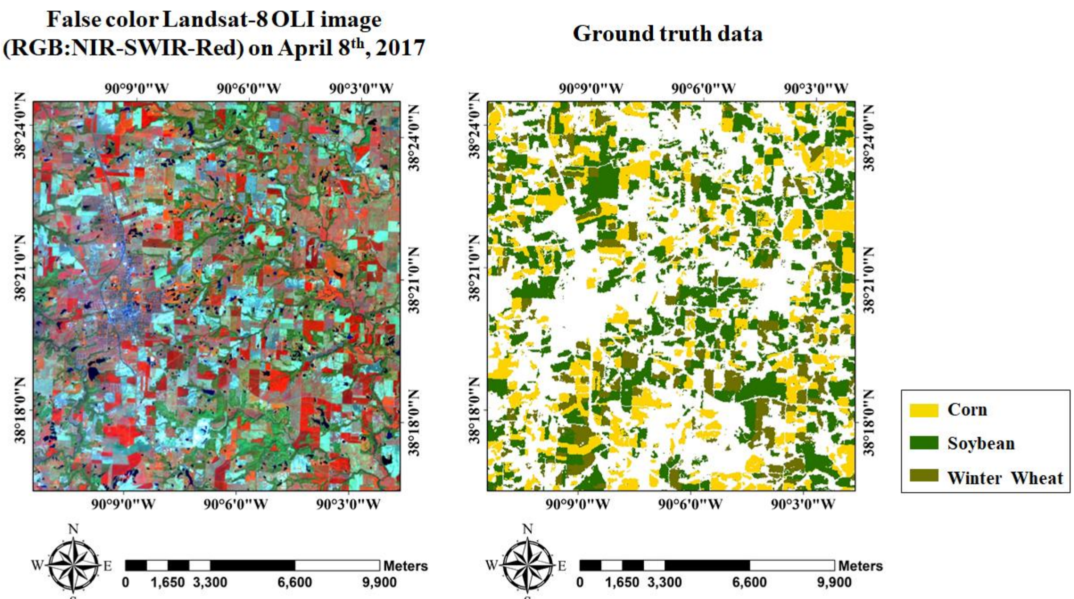

Figure 2. Landsat-8 OLI image on April 8, 2017 and CDL data in Illinois.

3. METHODOLOGY

In this study, the class purity (CP) is defined as the matching degree or ratio between pixels within the patch and the representative class of the patch. For example, if the representative class of a 5 by 5 patch is cabbage and fifteen cabbage pixels are included in the patch, CP is defined as 60 % (CP60). Remaining ten pixels, except for the fifteen cabbage pixels, may be other crops or non-crop such as roads, buildings, etc.

The selection of training data on patch-based classification starts with a single pixel. First, a patch is generated with a predefined size at all pixels of the entire image. Then, the number of pixels corresponding to a specific ratio (i.e., threshold value) is determined from the defined patch size. After extracting pixels corresponding to a certain ratio of the number of pixels allocated to the specific patch, the patch is selected as a candidate for training data if the class except for non-crop is the same as the class of the center of the patch. Final training samples are collected by random sampling of 0.1% of the training candidates considering the ratio of crops. In this study, two case study areas were classified with three different CP values including 60, 80, 100 % of class purity (CP60, CP80, CP100). Since the number of pixels and the class distribution corresponding to CP are different with respect to the patch size, we also compared and analyzed classification results according to different patch sizes for the two study areas (Table 2).

As for the path-based classifier, we applied 2D-CNN that extracts spatial features by applying a series of convolution filters to input data. Spectral information in patch units from the 2-dimensional input data is used for the extraction of spatial features and classification by 2D-CNN is then performed using the extracted spatial features (Kim et al., 2018). We used a common 2D-CNN model with fixed architecture and parameterization for both study areas to analyze the variations of classification accuracy according to class purity (Table 2). To avoid overfitting by using much deeper layers, we used shallow layers for the extraction of spatial features through a preliminary test. Since initial weights and kernel configuration are randomly chosen in 2D-CNN, the final classification accuracy was calculated by averaging four classification results.

Table 2. Summary of training parameters in each study area

Classifier	Parameter	Study area & values	
		Abandegi	Illinois
2D-CNN	Dropout rate	0.2	
	Patch size	5, 9, 13, 17, 21	5, 9, 15
	Kernel size	3	
	Number of filters	32	

Table 3. Summary of 2D-CNN architecture. P and F refer the input patch size and the number of filters, respectively

2D-CNN architecture		
Layer (type)	Output shape (dimension)	Number of parameters
Conv2D_1	(P, P, F)	896
Conv2D_2	(P, P, F)	9248
Max-pooling2D	(P/2, P/2, F)	0
Conv2D_3	(P/2, P/2, F*2)	18496
Dropout	256 neurons	0
Flatten	256 neurons	0
ReLu	64 neurons	16448
softmax	4 neurons	260

4. RESULTS AND DISCUSSION

Figure 3 presents variations of overall accuracy in two case study areas with respect to different CP values and patch sizes. In Anbandegi, the classification accuracy increased as CP increased, regardless of patch sizes (Figure 3(a)). When pure training samples were used for classification (i.e., CP100), the best classification accuracy was achieved. As the patch size increases, the impact of the CP value on classification accuracy decreased. In particular, the difference in overall accuracy between CP60 and CP100 was about 4 %p when the 5 by 5 patch was used. In contrast to Anbandegi, the classification accuracy in Illinois decreased as CP values became larger (Figure 3(b)). When the patch size was 9 by 9, the difference in overall accuracy between CP60 and CP100 was about 6.5 %p. In the case of using the 5 by 5 patch, the significant difference in overall accuracy was not observed, but CP100 still showed the lowest accuracy.

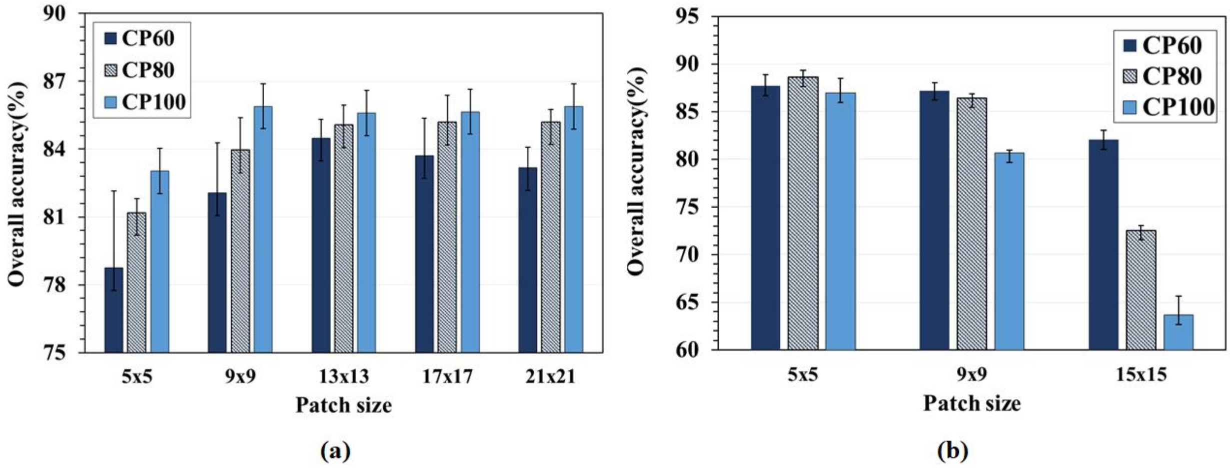

Figure 3. Overall accuracy by CP and patch size of training samples in study areas: (a) Anbandegi and (b) Illinois.

The classification result in Anbandegi is shown in Figure 4. In the classification result using CP 60, areas near the parcel boundary were well classified, but misclassification occurred within parcels. In contrast, misclassification of potato and fallow decreased for CP100 than CP60, but misclassification of highland Kimchi cabbage as cabbage was observed at the parcel boundaries. In all classification results, particularly, two specific cabbage parcels were misclassified as potato and highland Kimchi cabbage, mainly due to cabbage harvest. In the case of Anbandegi, the shape of parcels is not irregular and crop patterns within the parcels are uniform. Consequently, the classification performance got better when pure training samples were used for classification. As the CP values of training samples increased, the inside of each parcel which has a relatively high proportion of pixels with the same class, could be well classified, but the misclassification pattern appeared near the parcel boundary. The increase of the classification accuracy by using the larger patches may be explained by the fact that the larger patch includes more pixels with uniform spatial patterns and less diverse classes.

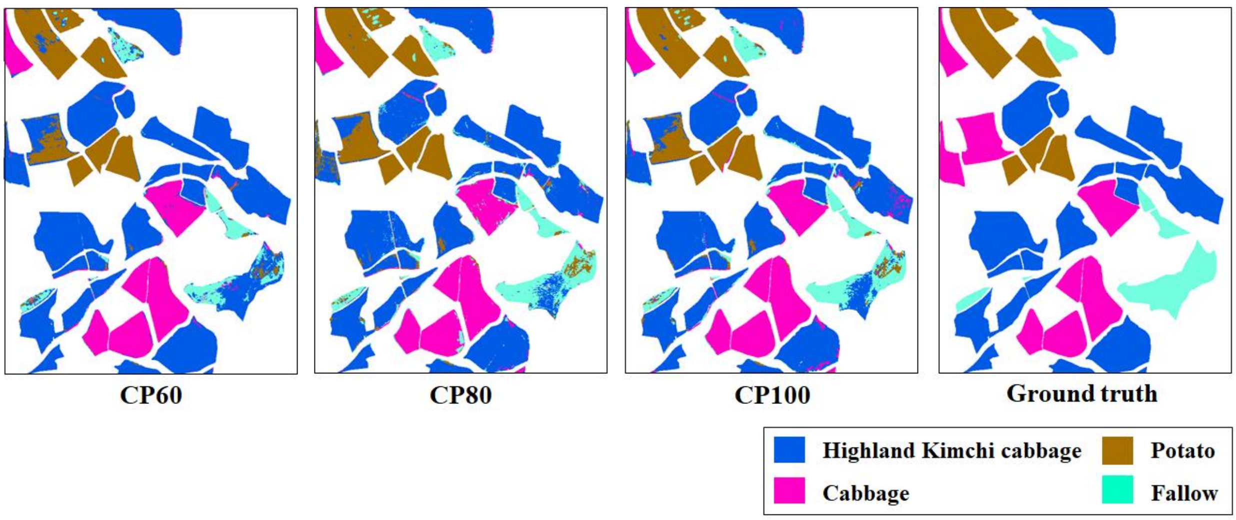

Figure 4. Classification results according to different CP when using the 13 by 13 patch in Anbandegi.

As shown in Figure 5, in the case of CP60 in Illinois, the distinction between corn and soybean was clear at the boundary, and the inside of each parcel was well classified. However, most of corn was misclassified as soybean for CP100. The Illinois area has diverse sizes of crop parcels, compared to the Anbandegi area. Thus, as the CP value become lower, more training samples were selected near the parcel boundary, leading to the clear distinction between corn and soybean at the boundary. The increase of the patch size also led to the decrease in classification accuracy, because many pixels with too diverse spatial patterns were included in the training patch.

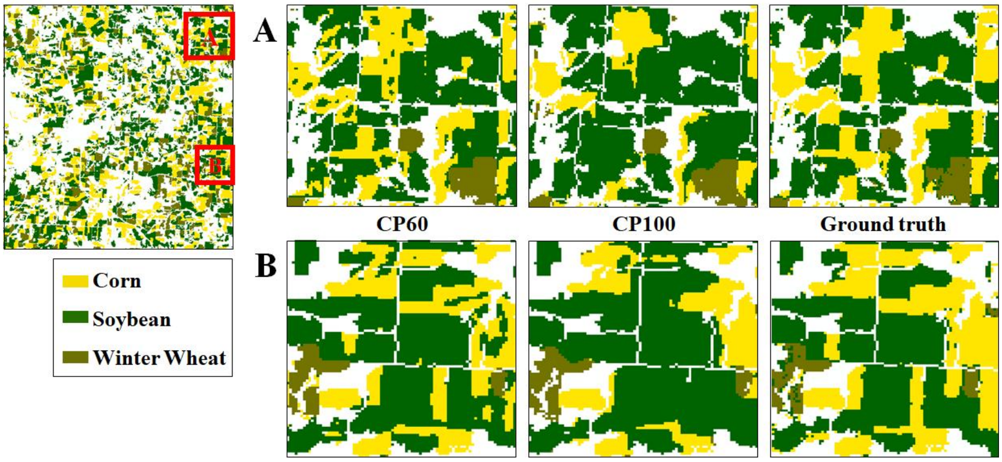

Figure 5. Classification results according to difference CP when using the 9 by 9 patch in sub-areas of Illinois.

5. CONCLUSIONS

This study investigated the effects of class purity of training data on crop classification based on patch-based 2D-CNN. The classification performance was evaluated through crop classification in two case study areas where the scale and shapes of crop parcels were quite different. When crop parcels with relatively simple shapes were classified with high resolution UAV imagery, the use of larger CP values increased the classification accuracy. In contrast, the classification accuracy decreased when the large CP values were used for the classification of crop parcels with diverse scales and shapes. These case study results indicate that the characteristics of target crops and the study area should be considered for the selection of optimal training samples. For the classification of areas where the spectral distinction of different crops is relatively clear and crop patterns are uniform, training data with larger CP should be selected. On the other hand, mixed pixels with lower CP should be selected as training samples when the study area consists of various crop parcels and the spectral variability of different crops is similar. To verify the major findings of this study, the effects of the spectral purity of crop types will be investigated as future work.

ACKNOWLEDGEMENT

This study was carried out with the support of "Cooperative Research Program for Agriculture Science & Technology Development (Project No. PJ01350004)" Rural Development Administration, Republic of Korea. The authors thanks National Institute of Agricultural Sciences for providing UAV imagery and ground truth data in the Anbandegi area.

REFERENCES

Brown, M., Gunn, S.R., and Lewis, H.G., 1999. Support vector machines for optimal classification and spectral unmixing. Ecological Modelling, 120 (2-3), pp. 167-179.

Carranza-García, M., García-Gutiérrez, J., and Riquelme, J. C., 2019. A framework for evaluating land use and land cover classification using convolutional neural networks. Remote Sensing, 11 (3), 274.

Foody, G.M. and Mathur, A., 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. Remote Sensing of Environment, 103, pp. 179-189.

Kim, Y., Kwak, G.-H., Lee, K.-D., Na, S.-I., Park, C.-W., and Park, N.-W., 2018. Performance evaluation of machinelearning and deep learning algorithms in crop classification: Impact of hyper-parameters and training sample size. Korean Journal of Remote Sensing, 34 (5), pp. 811-827.

Kim, Y., Park, N.-W., and Lee, K.-D., 2017. Self-learning based land-cover classification using sequential class patterns from past land-cover maps. Remote Sensing, 9 (9), 921.

Kussul, N., Lavreniuk, M., Skakun, S., and Shelestov, A., 2017. Deep learning classification of land cover and crop types using remote sensing data. IEEE Geoscience and Remote Sensing Letters, 14 (5), pp. 778-782.

Kwak, G. H. and Park, N. W., 2019. Impact of texture information on crop classification with machine learning and UAV images. Applied Sciences, 9 (4), 643.

Liang, H. and Li, Q., 2016. Hyperspectral imagery classification using sparse representations of convolutional neural network features. Remote Sensing, 8 (2), 99.

Li, Y., Zhang, H., and Shen, Q., 2017. Spectral–spatial classification of hyperspectral imagery with 3D convolutional neural network. Remote Sensing, 9 (1), 67.

Mathur, A. and Foody, G.M., 2008. Crop classification by support vector machine with intelligently selected training data for an operational application. International Journal of Remote Sensing, 29 (8), pp. 2227-2240.

Ok, A.O., and Akyurek, Z., 2011. Automatic training site selection for agricultural crop classification: a case study on Karacabey Plain, Turkey. ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 3819, pp. 221-225.