Wetland Classification Using Deep Convolutional Neural Network

Conference Paper · November 2018

DOI: 10.1109/IGARSS.2018.8517919

Masoud Mahdianpari
C-CORE and Memorial University of Newfoundland

Yun Zhang

Bahram Salehi
State University of New York (SUNY) College of Environmental Science and Forestry (…

Wetland Classification Using Deep Convolutional Neural Network

Masoud Mahdianpari[1], Mohammad Rezaee[2], Yun Zhang[2], Bahram Salehi[1]

1 C-CORE and Department of Electrical Engineering, Memorial University of Newfoundland,
St. John's, Newfoundland A1B 3X5, Canada. {bsalehi, m.mahdianpari}@mun.ca
2 CRC-Laboratory in Advanced Geomatics Image Processing, Department of Geodesy and Geomatics Engineering,
University of New Brunswick, Fredericton, New Brunswick, CANADA E3B 5A3. {mrezaee, yunzhang}@unb.ca

ABSTRACT

The synergistic use of spatial features with spectral properties of satellite images enhances thematic land cover information. This study aims to address the lack of high-level features by proposing a classification framework based on convolutional neural network (CNN) to learn deep spatial features for wetland. In particular, a CNN model was used for classification of remote sensing imagery with limited number of training data by fine-tuning of a pre-existing CNN (AlexNet). The classification results obtained by the deep CNN were compared with those based on well-known ensemble classifiers, namely Random Forest (RF), to evaluate the efficiency of CNN. Experimental results demonstrated that CNN was superior to RF for complex wetland mapping even by incorporating the small number of input features (i.e., 3 features) for CNN compared to RF. The proposed classification scheme serves as a baseline framework to facilitate further scientific research using the latest state-of-art machine learning tools for processing remote sensing data.

Index Terms— *Convolutional Neural Network, high-level features, AlexNet, Random Forest, Machine Learning, wetland mapping.*

1. INTRODUCTION

Wetlands are transitional zone between a water body and dry land, which may experience wet condition permanently or at least periodically during high water season [1]. Despite their high contributions to the ecosystem, they have been threatening by the anthropogenic and natural process during past decades [1]. Remote sensing tools have significantly contributed to the wetland mapping and monitoring for the restoration and preservation in a variety of aspects, including classification [2], change detection [3], and water level monitoring [4].

Despite significant improvements in remote sensing tools in both satellite image and applied techniques, classification of complex heterogeneous land cover such as wetland is challenging. Several approaches have been proposed in order to evaluate the efficiency of integrating spectral-spatial features for classification, including kernel methods [5] and Markov Random Field (MRF) [6]. However, these methods applied low-level features such as spectral information within neighboring pixels or morphological properties. Thus, the main disadvantage associated with these techniques is setting the proper parameters in order to produce suitable features for the different image objects [7].

Convolutional Neural Network (CNN) is one of the most efficient approaches among all deep learning based frameworks that does not require a prior feature extraction and thereby, has a greater generalization capability [8]. This is because a multi-layer based classifier has a high capacity to exploit abstract and invariable features. In particular, a deep CNN extracts the varying level of abstraction for the data in different layers. For example, low-level (e.g., edges), intermediate level (e.g., object fragment), and high-level information (e.g., full object) can be obtained in the initial, intermediate, and last layers, respectively [9].

Although the application of CNN was employed in a number of remote sensing studies for classification of different land cover types using hyperspectral imagery [10], its efficiency was not examined for complex land cover classification (e.g., wetland and sea ice). This study aims to remove the tedious process of feature generation and accordingly, feature selection; and determine the suitability of CNN for complex wetland classification.

2. METHODOLOGY

2.1. Dataset

The study area is located in Newfoundland and Labrador, Canada, covering an area of approximately 700 km2. Different wetland classes specified by Canadian Wetland Classification System (CWCS), including bog, fen, marsh, swamp, and shallow-water, are found within the study region [1]. Other land cover types are upland, urban, and deep-water in this ecoregion.

Table 1: Testing and training pixel counts for the Avalon reference data.

Class	Class Description	#Training Pixels	#Testing Pixels	Total
Bog	Peatland dominated by Sphagnum species	20650	19565	40215
Fen	Peatland dominated by graminoid species	11183	8794	19977
Swamp	Mineral wetlands dominated by woody vegetation	3197	9491	12688
Marsh	Mineral wetlands dominated by emergent graminoid species	10869	5238	16107
Shallow Water	Mineral wetlands dominated by submerged and floating vegetation	6205	5679	11884
Urban	Human-made structures	66339	67125	133464
Deep Water	Deep water areas	62927	89194	152121
Upland	Dry forested upland	73458	89878	163336
Total		254828	294964	549792

Field data were acquired for 191 sample sites in summers and falls of 2015, 2016 (Table 1). For classification, two RapidEye optical images in Level 3A (radiometrically and geometrically corrected) that were acquired on June 18 and October 22, 2015, were used.

2.2. Patch-Based Image Labeling (PBIL)

The patch-based image labeling method was introduced to convert the categorization problem in deep learning into the classification in order to make CNN compatible for remote sensing applications. Since the number of training data is not high, the AlexNet, which just has five convolution layer is utilized in this study [11].

2.3. Training Procedure

The main challenges associated with the network training are the limited number of training samples and determining an optimum batch size to be utilized in PBIL. Instead of full training a network from scratch, a pre-trained network can be utilized. In this approach, the parameters of the last layers are mostly updated. Furthermore, the update's values are small since the updating is carried out on a pre-trained network. Due to the limited number of *in-situ* data in this study, the last four layers of the network were updated to ensure the sufficient amount of sampling data for both training and testing the network. The patch size of 30 pixels (i.e., 150 *m* on the ground) found to be an optimum value given the spatial resolution of 5 m for the image and the smallest object size for this study. This patch size was obtained by a trial and error procedure.

3. RESULTS AND DISCUSSION

The CNN network was trained using Caffe library [12]. Specifically, the training was carried out on a computer with an Intel® Xenon 2.80GHz CPU (16GB memory) and a Nvidia Quadro k2200 GPU (4GB memory). By setting 30,000 iterations, the training step was completed in approximately 12 hours. Fig. 1 illustrates the loss and accuracy curve for the validation set.

As seen in Fig. 1, the speed of convergence is high in the first epochs, since the training is actually a fine-tuning of a pre-trained network. Fig. 2 depicts the first convolution layer, its corresponding kernels and features in AlexNet.

To evaluate the efficiency of CNN for wetland mapping, the classification results of CNN were compared with Random Forest (RF) classifier. RF is an ensembles classifier and has shown good results for several lands cover mappings, such as wetland [2]. For classification based on the RF classifier, all original spectral bands of RapidEye image were used. However, for CNN only three original spectral bands of RapidEye image, including red, green, and NIR2 were applied. The classification maps obtained by RF and CNN are depicted in Fig. 3.

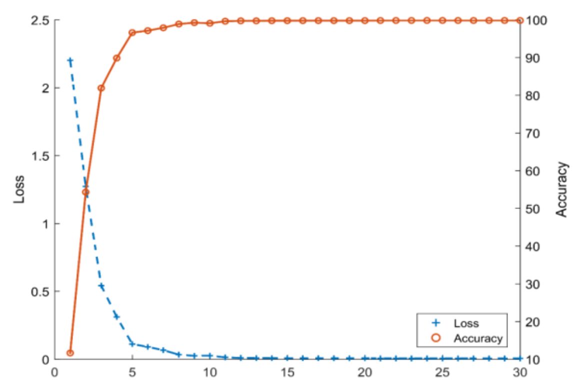

Figure 1: The value of validation accuracy and loss as a function of epochs.

As can be seen, there is a significant degree of disagreement between two classified maps for all wetland classes. For example, the dominant wetland classes obtained by RF are swamp and marsh wetlands. Whereas, the dominant wetland classes for CNN are bog and fen. As reported by field biologists participating during field data collection, bog and fen wetlands are dominant classes. This is attributed to the

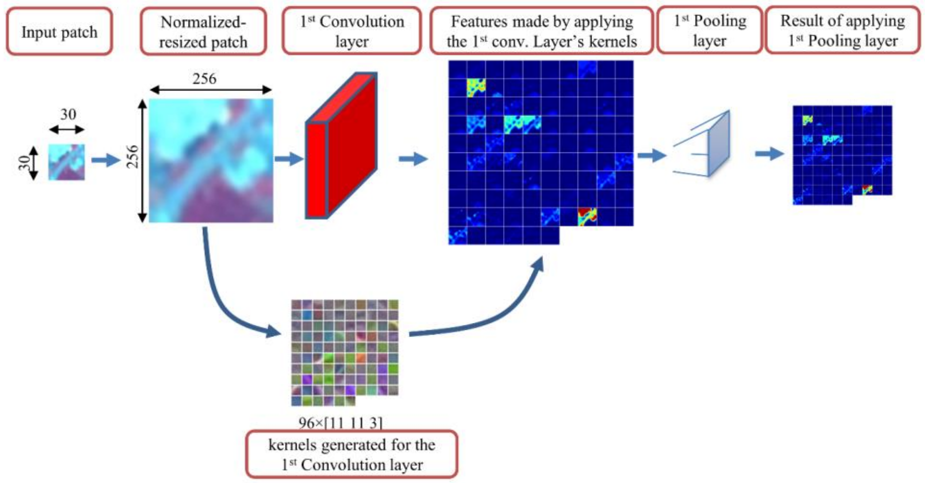

Figure 2: The first convolution layer, its designed kernels, and generated features.

oceanic climate of the Avalon area facilitating extensive peatland formation (i.e., bog and fen). On the other hand, the dominant non-wetland class is upland, which is defined as forested dry land. Notably, the classified map produced by CNN is realistic and demonstrating the detailed spatial distribution of all land cover classes presented in the study area. For example, the classified map shows the predominance of bog and upland classes, while marsh and swamp are less prevalent. These observations fit well with fields notes recorded during field data collection. Confusion matrices for these classified maps are presented in Tables 2. The classification overall accuracy of about 72% was obtained for RF, all wetland classes produce an accuracy of less than 50%. In particular, bog is only correctly classified in 45% of cases, fen in 32% of cases, swamp in 47% of cases, marsh in 37% of cases, and shallow-water in only 33% of cases. Thus, the overall accuracy of 72% is due to the high classification accuracy for non-wetland classes, such as deep-water and urban classes. However, the classification overall accuracy of about 95% is achieved using CNN by incorporating three input features. This is of great significance taking into account the complexity of similar wetland classes and the large number of pixels, which were correctly classified. In particular, all land cover classes have high producer's accuracies of greater than 77%, excluding the fen class. More precisely, bog correctly classified in 89% of cases, fen in 62% of cases, swamp in 78% of cases, marsh in 77% of cases, and shallow-water in 95% of cases.

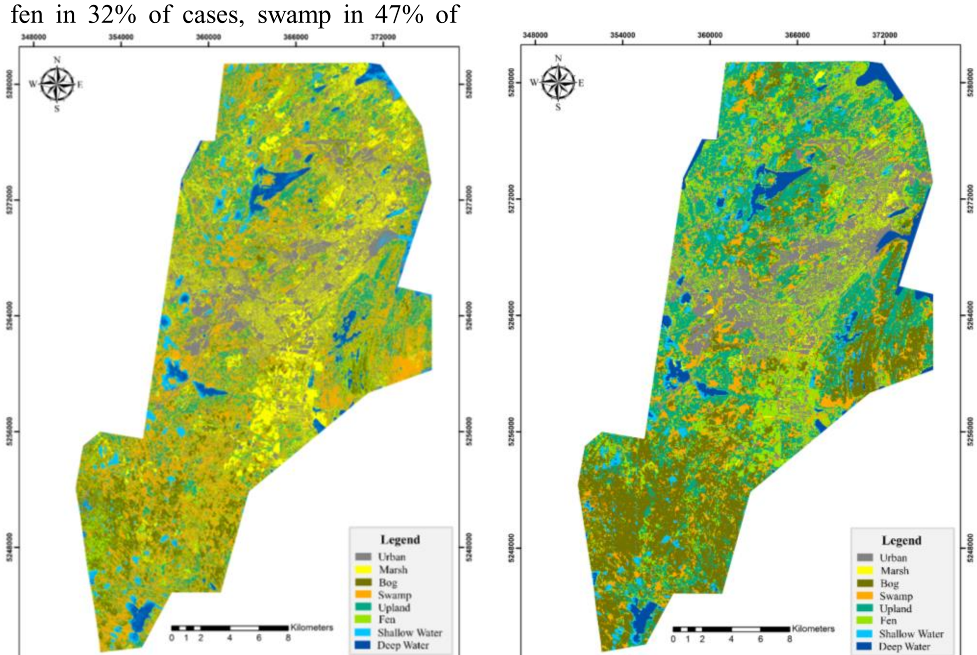

Figure 3: the classification maps obtained by three input feature (left): RF and (right) CNN

Table 2: Confusion matrix of CNN: Overall accuracy: 94.82%, Kappa coefficient: 0.93.

	Class	Reference Data: Bog	Fen	Swamp	Marsh	Upland	Urban	Shallow-water	Deep-water	Tot.	User Acc.
Classified Data	Bog	15237	1810	4	11	1320	1183	0	0	19565	77.88
	Fen	256	7094	26	920	436	8	54	0	8794	80.67
	Swamp	203	128	7623	156	978	403	0	0	9491	80.32
	Marsh	125	71	773	4015	168	86	0	0	5238	76.65
	Upland	1259	2187	1259	59	85114	0	0	0	89878	94.70
	Urban	0	21	0	0	931	66173	0	0	67125	98.58
	Shallow-Water	0	0	0	0	0	0	5461	218	5679	96.16
	Deep-water	0	0	0	0	0	0	228	88966	89194	99.74
	Total	17080	11311	9685	5161	88947	67853	5743	89184	294964	
	Prod. Acc.	89.21	62.72	78.71	77.80	95.69	97.52	95.09	99.76		

4. CONCLUSION

Canada is one of the richest countries in terms of different wetland types. However, both anthropogenic and natural processes have expedited wetland loss during the last two decades. Thus, much effort needed to preserve the diversity of species within wetland ecosystems. Satellite imagery has provided a unique source of data from several inaccessible wetlands to map and monitor these productive habitats using cost- and time-efficient tools. In this study, we utilized a classification framework based on the current state-of-art deep convolutional neural network (CNN) for wetland classification. More specifically, a pre-existing convolutional neural network (AlexNet) was applied for classification of satellite imagery. The results demonstrated the higher potential of CNN for classification even by incorporating the less number of input features. Specifically, the overall accuracy of CNN was 94.82% suggesting an improvement of 15.73% compared to the Random Forest (RF) classifier for all land cover types. Moreover, an average improvement of by about 30% was attained for wetland classes when CNN was employed. The proposed classification framework provides an insight into the significance of incorporating high-level spatial features to the classification scheme to reduce the confusion between spectrally similar classes.

5. REFERENCE:

[1] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, and B. Brisco, "An assessment of simulated compact polarimetric SAR data for wetland classification using random Forest algorithm," *Canadian Journal of Remote Sensing*, *43*(5), pp.468-484, 2017.

[2] M. Mahdianpari, B. Salehi, F. Mohammadimanesh, and M. Motagh, "Random forest wetland classification using ALOS-2 L-band, RADARSAT-2 C-band, and TerraSAR-X imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 13–31, 2017.

[3] B. Brisco, F. Ahern, K. Murnaghan, L. White, F. Canisus, and P. Lancaster, "Seasonal Change in Wetland Coherence as an Aid to Wetland Monitoring," *Remote Sens.*, vol. 9, no. 2, 2017.

[4] S. Wdowinski, S.-W. Kim, F. Amelung, T. H. Dixon, F. Miralles-Wilhelm, and R. Sonenshein, "Space-based detection of wetlands' surface water level changes from L-band SAR interferometry," *Remote Sens. Environ.*, vol. 112, no. 3, pp. 681–696, 2008.

[5] B. Scholkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Cambridge, MA, USA: MIT Press, 2001.

[6] M. Mahdianpari, B. Salehi, and F. Mohammadimanesh, "The effect of PolSAR image de-speckling on wetland classification: introducing a new adaptive method," *Canadian Journal of Remote Sensing*, *43*(5), pp.485-503, 2017.

[7] W. Zhao and S. Du, "Learning multiscale and deep representations for classifying remotely sensed imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 113, pp. 155–165, 2016.

[8] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[9] K. Nogueira, O. A. B. Penatti, and J. A. dos Santos, "Towards Better Exploiting Convolutional Neural Networks for Remote Sensing Scene Classification," *Pattern Recognit.*, vol. 61, pp. 539–556, 2017.

[10] Y. Chen, Z. Lin, X. Zhao, G. Wang, and Y. Gu, "Deep learning-based classification of hyperspectral data," *IEEE J. Sel. Top. Appl. earth Obs. Remote Sens.*, vol. 7, no. 6, pp. 2094–2107, 2014.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[12] Y. Jia *et al.*, "Caffe: Convolutional Architecture for Fast Feature Embedding," *arXiv Prepr. arXiv1408.5093*, 2014.