

# Deep Convolutional Neural Network for Complex Wetland Classification Using Optical Remote Sensing Imagery

**Published in: IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing ( Volume: 11, Issue: 9, Sept. 2018)**

Presented by  
CSIE 3A 潘欣妤, CSIE 3B 鄭筠蓉

Authors: Mohammad Rezaee, Masoud Mahdianpari, Yun Zhang, and Bahram Salehi, Member, IEEE

---

**01**  
**INTRODUCTION**

**03**  
**RESULT and DISCUSSION**

**02**  
**METHOD**

**04**  
**CONCLUSION**

---

---

## 論文摘要

- 現今的空間分析多是利用低階特徵來進行分類，沒有充分利用高階特徵。
  - 本研究旨在提出**以卷積神經網路的分類框架來提供衛星影像圖的高階特徵分析**，使用光學遙測數據分析濕地影像的特徵。
  - 使用AlexNet模型來提高預判的精準度。結果顯示，卷積神經網路能夠以較少量的輸入特徵得到較優異的濕地映射結果。
- 

# 01

## INTRODUCTION

---

## 1. 溼地的危機與監測

- 濕地生態系為地球提供許多重要的貢獻，卻在過去幾十年內因為人為開發與自然的演替而大幅減少。
  - 遙感技術的進步得以讓我們獲得濕地面積變化檢測、水位監測等多個方面的重要數據。
  - 由於**濕地複雜的地貌特性**以及其**高度破碎化的性質**，導致不同的分類群可能擁有數個相似的光譜特性，使得分類更加困難。
- 

---

## 2. 利用光譜與空間特徵分類

- 現階段的評估方法多使用衛星影相中的低階特徵，例如相鄰像素的光譜或形態特徵。
  - **現行的分類方法**難以替不同的資料集設置適當的參數初始值，而這類方式也**僅適用於具體的案例研究和資料集、難以推廣**。
  - 因為CNN能夠提取較為抽象的資料特徵，還能夠在不同的層級之中提取不同等級的抽象資料特徵，因此能夠達到更強的分類與推廣能力。
-

---

### 3. 卷積神經網路的訓練方式

- 儘管 Full training 的具有高效率和穩定性，但因為需要大量的訓練數據，因此不適用於遙測資料分類。
  - 相較之下，**Fine-tuning 以及 Pre-training 被認為更適合應用在遙測資料分類應用**，因為這兩種模型會事先使用不同的數據進行網路的訓練。
- 

---

### 4. 現今的地貌分類方式

- 使用 CNN 來分類光譜影像圖已經廣泛地被應用在許多遙測研，但其應用於較複雜地貌的應用效率卻仍未被驗證。
  - 現今進行複雜地貌分類的方式，多是透過**建立龐大的訓練資料集來解決「相近光譜特性的不同地貌」的分類**；然而，這樣的做法需要耗費大量時間與成本在累積資料的階段。
-

---

## 5. 本研究的宗旨

有鑑於複雜地貌分類方法的主要弱點，本研究旨在：

1. 評估預訓練 CNN 在多光譜衛星圖像分類中的分類能力
2. 確立 CNN 對於複雜濕地分類的適用性
3. 為進一步的濕地製圖研究生成合適的分類模型

本論文將說明研究團隊於加拿大紐芬蘭-拉布拉多省的濕地分類結果，並說明 CNN 如何展現其在「相近光譜特性的不同地貌」的分類潛力。

---

# 02 METHOD

A. Study Area and Dataset

---

## 1. 研究區域之地理特徵

- 本研究區域位於加拿大的紐芬蘭-拉布拉多省，面積約700平方公里、由許多大小相異的生態區交錯組成，在水文、生態和地景等地理面向上各自有不同的特徵。
  - 而在此區域內共有 bog, fen, marsh, swamp 以及 shallow-water 等五種不同類型的溼地。
  - 而此省境內也有都會區、高原、深水區等不同的地貌可進行對照。
- 

---

## 2. 研究資料集

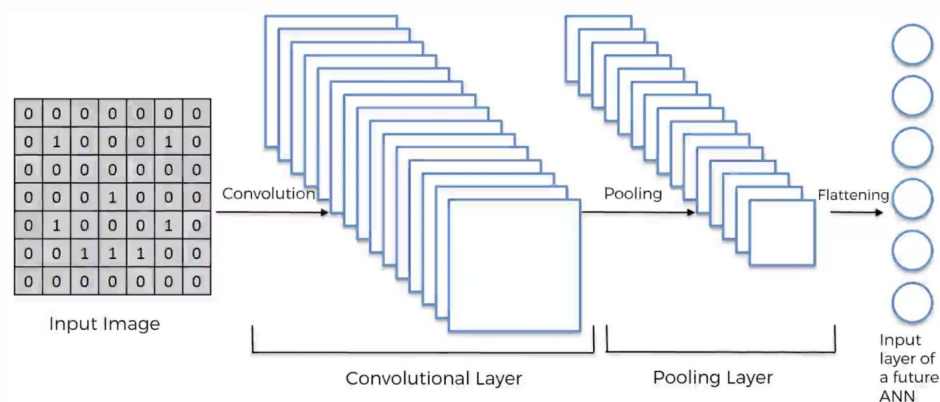
- 分別於 2015 年和 2016 年的夏季和秋季記錄了不同地貌的空間分佈狀態，並同步紀錄了每個採樣點的 GPS 定位，以便作為 CNN 的訓練資料集。
  - 而本研究的實驗數據是使用 RapidEye 衛星在 2015 年 6 月 18 日和 10 月 22 日獲得的兩幅 3A 級光學影像。包含五個不同光譜帶：藍光、綠光、紅光、red-edge 以及 near-infrared 的資訊。
-

## 02 METHOD

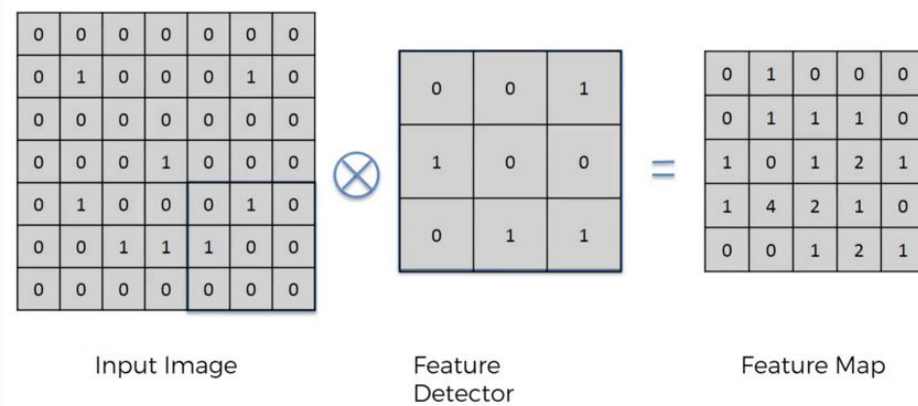
### B. Convolutional Neural Network

#### 1. 卷積神經網路與圖像分析

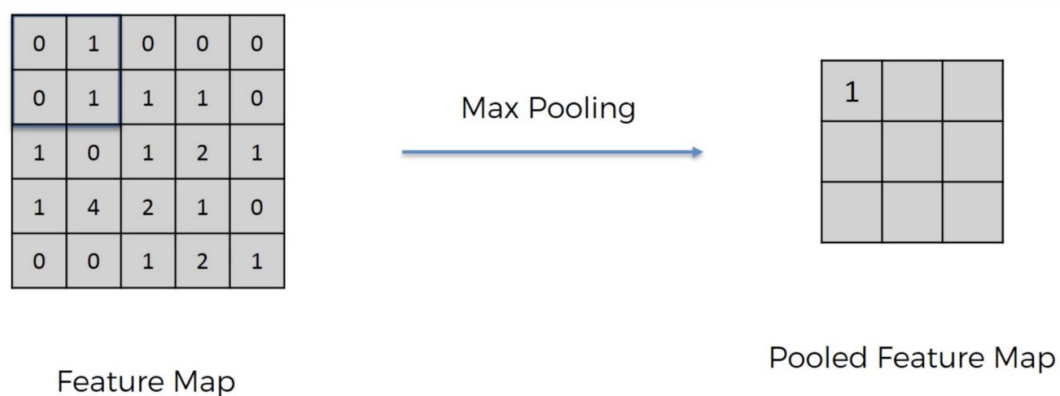
典型的CNN係由卷積層、池化層和全連接層三種架構組成。



## 2. 卷積神經網路架構說明——卷積層

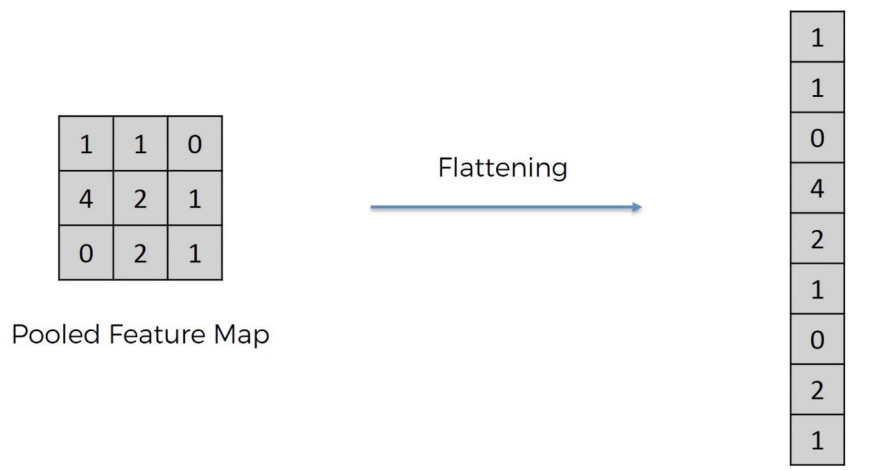


## 2. 卷積神經網路架構說明——池化層





## 2. 卷積神經網路架構說明——全連接層



# 02 METHOD

C. Patch-Based Image Labeling

## PBIL(Patch-Based Image Labeling)

- 目的

convert categorization problem to classification in order to make CNN **compatible** with remote sensing applications

- 作法

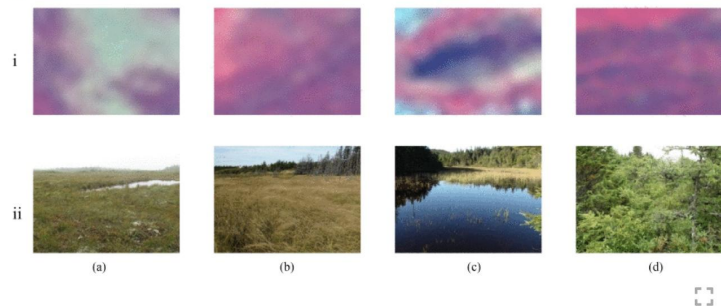
an input image is divided into several patches and a label is assigned to the center of each patch

$$P(n(M, i, w_m) | n(S, i, w_s))$$

$$f_{il}(s) = \exp(a_{il}(s)) / Z = P(m_i = l | s)$$

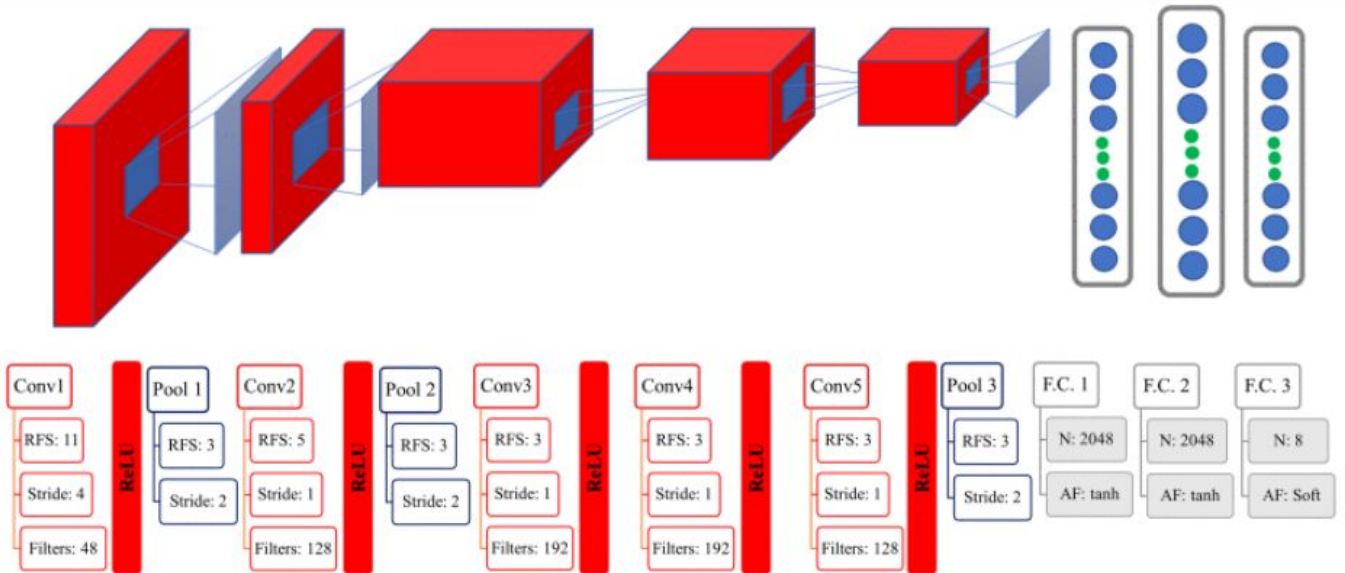
Fig. 3.

» Show in Context



Sample patches (i) and field surveying images (ii) of four wetland classes, namely (a) bog, (b) fen, (c) marsh, and (d) swamp.

## Alexnet



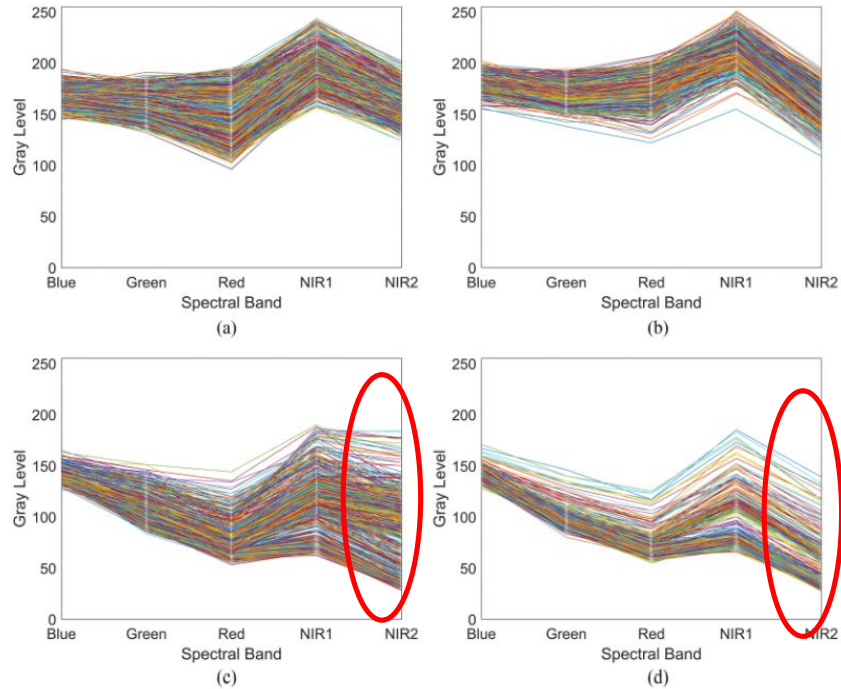
## Preprocessing Step

1. High standard deviation

high degree of misclassification would occur in case of exclusive use of spectral information

**解決:**

- a. band selection technique
- b. supplementary data
- c. extract the spatial information from images



## Preprocessing Step

2. High degree of overlap between patches

In the PBIL approach, the step parameter was adjusted to 1 pixel

解決:

normalized images

(the mean of each patch was set to zero by subtracting the mean value of each patch )

---

## Preprocessing Step

### 3. Cloud masking

This is because the CNN was not trained to classify the cloud, and, importantly, it could not have an unclassified label.

---

---

## Training Step

- main challenges
    1. the limited number of training samples  
解決:  
pretrained network can be utilized
    2. determine an optimum patch size  
解決:  
patch size of 30 pixels (i.e., 150m on the ground)  
found to be an optimum value
-

## Loss function

- negative log likelihood

$$L(s, m) = \sum_{all\ patches} \sum_{i=1}^{w^2m} (m_i \ln(f_i(s)) + (1 - m_i) \ln(1 - f_i(s)))$$

## Optimization

- stochastic gradient descent with minibatches

## Testing Step

- independent of the training dataset and both groups had roughly comparable pixel counts
- make sure that the network is not overfitted

Training data: first satellite imagery (June 18th)

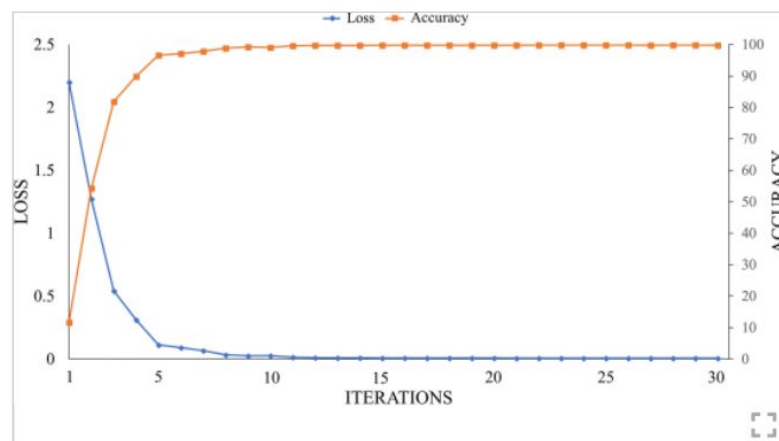
Testing data: second satellite imagery (October 22)

-> This procedure illustrates the reliability of results.

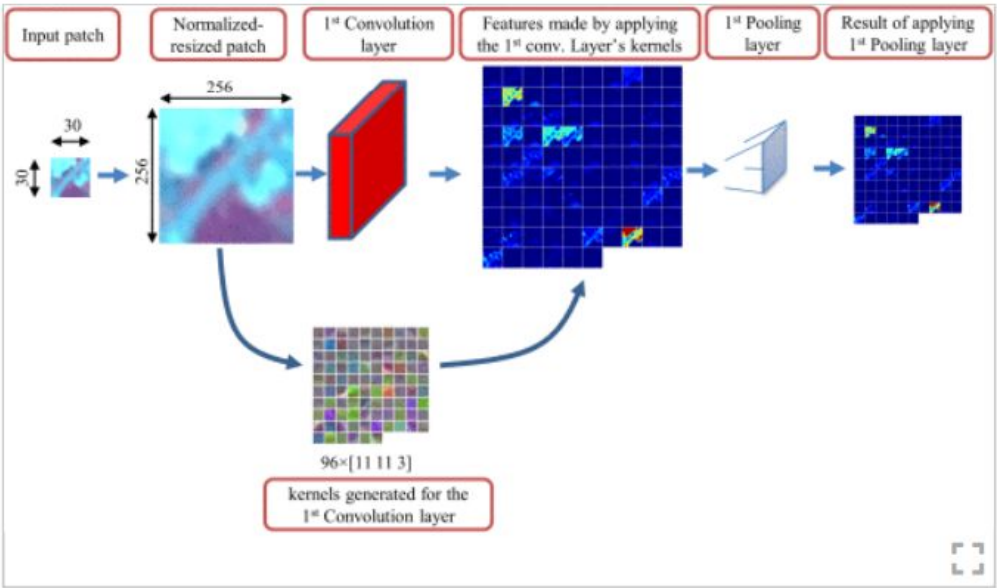
# 03

## RESULTS and DISCUSSION

The training step was completed using 30,000 iterations in approximately 12 hours



**Fig. 4.** Value of validation accuracy and loss as a function of epochs.



**Fig. 5.** First convolution layer, its designed kernels, and generated features.

These features highlight the area that is related to the corresponding class

Order	Original patch	Kernels of the first conv. layer	The result of the first conv. layer	The result of the second conv. layer	Label
1					7
2					4
3					5
4					1



To evaluate the efficiency of CNN for wetland mapping, the classification results of CNN were compared with the RF classifier

**TABLE II** Confusion Matrix of CNN: Overall Accuracy: 94.82% Kappa Coefficient: 0.93

Classified Data	Class	Reference Data								Tot.	User.Acc
		Bog	Fen	Swamp	Marsh	Upland	Urban	Shallow-water	Deep-water		
	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		

**TABLE III** Confusion Matrix of RF: Overall Accuracy: 79.11% Kappa Coefficient: 0.73

Classified Data	Class	Reference Data								Tot.	User.Acc
		Bog	Fen	Swamp	Marsh	Upland	Urban	Shallow-water	Deep- water		
	Bog	11620	2457	249	691	4337	17	194	0	19565	59.39
	Fen	1950	5907	108	649	180	0	0	0	8794	67.17
	Swamp	749	57	6234	376	1996	79	0	0	9491	65.68
	Marsh	188	192	51	4392	291	17	107	0	5238	83.85
	Upland	7913	3768	3106	39	59981	15071	0	0	89878	66.74
	Urban	775	818	591	97	4916	59928	0	0	67125	89.28
	Shallow-Water	11	8	0	516	0	0	3661	1483	5679	64.47
	Deep- water	0	0	0	24	0	0	7559	81611	89194	91.50
	Total	23206	13207	10339	6784	71701	75112	11521	83094	294964	
	Prod. Acc.	50.07	44.73	60.30	64.74	83.65	79.78	31.78	98.22		

# 04

## CONCLUSION

1. superiority of CNN relative to RF even by incorporating a smaller number of input features
2. significance of incorporating high-level spatial features into the classification scheme to reduce confusion between similar wetland

**Any questions?**