

<https://doi.org/10.7236/JIIBC.2017.17.5.97>

JIIBC 2017-5-13

컨볼루션 신경망을 기반으로 한 드론 영상 분류

Drone Image Classification based on Convolutional Neural Networks

주영도*

Young-Do Joo*

요약 최근 고해상도 원격탐사 자료의 분류방안으로 컨볼루션 신경망(Convolutional Neural Networks)을 비롯한 딥 러닝 기법들이 소개되고 있다. 본 논문에서는 드론으로 촬영된 농경지 영상의 작물 분류를 위해 컨볼루션 신경망을 적용하여 가능성을 검토하였다. 농경지를 논, 고구마, 고추, 옥수수, 깻잎, 과수, 비닐하우스로 총 7가지 클래스로 나누고 수동으로 라벨링 작업을 완료했다. 컨볼루션 신경망 적용을 위해 영상 전처리와 정규화 작업을 수행하였으며 영상 분류 결과 98%이상 높은 정확도를 확인할 수 있었다. 본 논문을 통해 기존 영상분류 방법들에서 딥 러닝 기반 영상분류 방법으로서의 전환이 빠르게 진행될 것으로 예상되며, 그 성공 가능성을 확신할 수 있었다.

Abstract Recently deep learning techniques such as convolutional neural networks (CNN) have been introduced to classify high-resolution remote sensing data. In this paper, we investigated the possibility of applying CNN to crop classification of farmland images captured by drones. The farming area was divided into seven classes: rice field, sweet potato, red pepper, corn, sesame leaf, fruit tree, and vinyl greenhouse. We performed image pre-processing and normalization to apply CNN, and the accuracy of image classification was more than 98%. With the output of this study, it is expected that the transition from the existing image classification methods to the deep learning based image classification methods will be facilitated in a fast manner, and the possibility of success can be confirmed.

Key Words : Deep Learning, Convolutional Neural Networks, Image Classification

1. 서론

인공위성, 항공기, 최근에는 드론에 이르기까지 탑재된 카메라를 통해 수집된 고해상도, 고품질의 영상은 다양한 응용분야에서 활용되고 있다. 그중에서 농경지 작물분류의 경우, 대개 광범위한 지역을 다루기 때문에 그들 고해상도 영상을 활용한 원격탐사 기법이 필수적이다. 최근까지도 서포트 벡터 머신(Support Vector Machine)과 의사결정 트리(Decision Tree) 같은 머신 러닝

(Machine Learning) 기법들이 영상 분류에 적용되어 만족할만한 결과를 보여왔다^[1-4]. 특히 지도학습(Supervised Learning)에서 기존의 머신 러닝은 학습 데이터의 특성과 이러한 특성을 기반으로 새로운 데이터를 구별하는 학습 모델을 사람이 사전에 정의했다. 그러나 본 논문에서 소개할 컨볼루션 신경망(Convolutional Neural Networks: CNN)같은 딥 러닝(Deep Learning) 기법은 사람이 학습하는 것을 컴퓨터가 스스로 할 수 있게 한다는 것이 기존 머신 러닝 방법과의 가장 큰 차이점

*정회원, 강남대학교 컴퓨터미디어정보공학부
접수일자: 2017년 8월 14일, 수정완료: 2017년 9월 14일
게재확정일자: 2017년 10월 13일

Received: 14 August, 2017 / Revised: 14 September, 2017 /

Accepted: 13 October, 2017

*Corresponding Author: ydjoo@kangnam.ac.kr

Dept. of Computer & Media Information, Kangnam University, Korea

이다. 스스로 자가 학습을 통해 새로운 데이터를 분류해 내는 비지도학습(Unsupervised Learning)에 해당하는 딥 러닝 기술은 최근 인식(Recognition), 구별(Segmentation), 분류(Classification), 분석(Analysis) 등 다양한 분야에서 괄목할 만큼 활발하게 연구되어지는 주제이다^{[5],[6]}.

본 논문에서는 딥 러닝 기법 중에 가장 널리 사용되는 CNN을 드론이 촬영한 영상기반 농경지 작물분류 문제에 적용하여 그 가능성을 검토하였다. 2장에서는 CNN에 대한 기본 개념을 설명하고 3장에서는 드론 영상에 CNN을 적용한 실험 결과를 기술하였다.

II. 컨볼루션 신경망

CNN은 일반 신경망(Neural Network)과 매우 유사하나 이미지가 가지고 있는 특성을 인코딩 할 수 있어 입력이 이미지인 영상 처리에 주로 사용된다. CNN 모델은 그림 1에서 보는 바와 같이 컨볼루션층(Convolutional Layer), 맥스풀링층(Maxpooling Layer), 완전연결층(Fully-connected Layer) 등으로 구성된다^[7]. 컨볼루션층은 컨볼루션 신경망을 이루는 핵심 요소이며 학습 가능한 필터들과 이 필터를 통해 계산 값을 비선형으로 바꾸어 주는 활성화 함수(Activation Function)로 이루어진다. 입력 이미지에 필터를 통과시키면 특징 맵(Feature Map)이 생성되며, 이 특징 맵에 활성화 함수를 적용한다. 컨볼루션층은 주로 은닉층에 사용되며, 활성화 함수로는 ReLU(Rectified Linear Unit)가 주로 사용된다. 컨볼루션층이 입력층이라면 입력값은 입력 이미지의 픽셀값이 되고, 입력층이 아니라면 입력값은 이전 컨볼루션층의 특징 맵의 값이 된다.

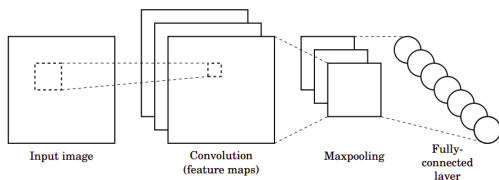


그림 1. 컨볼루션 신경망 개념도

Fig. 1. Conceptual Diagram of CNN

(출처 : <https://www.vaetas.cz/assets/img/convolutional-neural-network.png>)

컨볼루션층의 가장 큰 특징은 지역적 연결성을 가지

고 있으며, 연결 가중치를 공유하고 있다는 것이다. 일반적인 다층신경망(Multi-layered Neural Network)에서는 인접한 두 개의 층간의 뉴런(Neuron)들이 모두 연결된 완전 연결 구조를 갖는다. 이러한 완전 연결 구조를 이용하면 이론적으로 입력 영상에 있는 위상학적 구조를 찾을 수는 있으나 현실적으로는 입력 영상이 다차원으로 이루어져 있기 때문에 문제 복잡도가 매우 크고 일반화가 되지 않아 과적합(Overfitting)이 생기는 문제점이 있다. CNN에서는 이전 층과 지역적으로 제한된 영역의 뉴런들이 연결되어 있으며, 이러한 구조로 이미지의 위상학적 구조를 일반화시켜 학습시킬 수 있다. 또한 위치가 바뀌더라도 그 연결 가중치가 공유되기 때문에 위치에 무관하게 특징을 추출할 수 있다. 컨볼루션층을 통과하면서 필터 계산에 의해 경계부분이 소실되어 특징 맵이 축소된다. 깊은 층을 가진 모델인 경우 특징 맵이 계속 작아서 정보가 많이 손실되므로 필터를 통과하더라도 원본 사이즈가 유지될 수 있도록 패딩(Padding) 처리를 수행한다.

맥스풀링층은 이전 층에서 생성된 특징 맵의 해상도를 줄여주며, 이 과정에서 최대치만 취하는 연산을 수행한다. 이 층은 차원을 축소할 뿐만 아니라 특징의 이동 및 왜곡에 강인한 모델을 만드는 효과가 있다. 컨볼루션 신경망 모델에서는 컨볼루션층과 맥스풀링층이 반복적으로 조합하여 특징을 추출한다. 마지막으로 컨볼루션층과 맥스풀링층을 통해 생성되는 특징을 분류할 수 있도록 완전연결층과 조합한다^{[8],[9]}. 출력은 완전연결층의 가장 마지막 층에서의 활성화 함수에 의해 정해진다. 본 논문에서는 다중 클래스 분류를 다루기 때문에 소프트맥스(Softmax) 함수를 사용한다. 소프트맥스 함수는 입력 특징에 대해서 각 클래스 별로 확률을 나타내며, 그 확률들의 합은 1인 특성을 가지고 있다.

III. 실험 연구 및 분석

이 장에서는 본 연구에서 사용된 데이터 정보 및 전처리 과정을 간략히 소개하고, 실제 드론 영상분류에 CNN을 적용 분석하였다. 사용된 데이터는 대전광역시에서 촬영한 영상을 사용하였으며, 3DR-SOLO 드론을 이용하여 30m 상공에서 촬영한 농경지 데이터이다. 표 1은 수집 데이터의 정보를 나타낸다.

표 1. 데이터 수집 정보

Table 1. Summary of Data Collection Information

항목	정보
드론명	3DR-SOLO
촬영고도	30m
촬영지역	대전광역시
수집기간	2016. 7. 8. - 2016. 8. 4.
영상개수	114장

우선 촬영한 영상 데이터의 기하왜곡을 보정하고 데이터 분류를 위해 클래스 별로 나누는 작업을 수행한다. 영상 데이터의 전처리를 위한 전체적인 수행 과정은 그림 2와 같다.

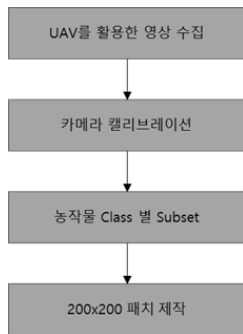


그림 2. 이미지 데이터 전처리 과정

Fig. 2. Pre-processing Steps of Image Data

드론에서 사용하는 카메라는 시야각(Field of View)이 넓은 광각 렌즈나 초광각 렌즈를 사용하여 넓은 범위를 볼 수 있지만 이로 인해 상대적으로 영상 왜곡이 심해지는 문제가 발생하기 때문에 이러한 왜곡을 보정하기 위한 과정이 필요하다^[10]. 그림 3과 같은 다수의 흑백 사각형이 정렬되어 있는 체커 보드를 이용하여 영상왜곡을 보정하기 위한 기하학적 정보를 수집한다. 그림 4는 카메라 보정을 수행하기 전과 보정 후의 영상을 비교하여 보여준다.

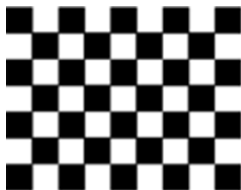


그림 3. 카메라 영상 왜곡을 보정하기 위한 체커 보드

Fig. 3. Checker Board for Camera Calibration

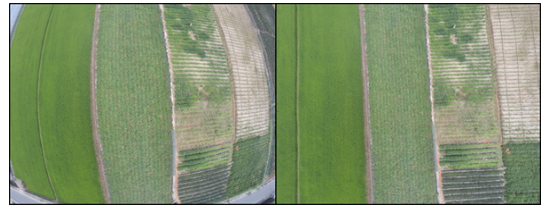


그림 4. 카메라 영상 왜곡 보정 전후 영상

Fig. 4. Before and After Camera Image Distortion Calibration

CNN 모델을 준비한 데이터로 훈련시키기 위해서 이미지와 이미지에 해당하는 라벨이 입력 데이터로 필요하다. 라벨이 붙여진 데이터 셋으로 만들기 위해 카메라 영상 왜곡을 보정한 영상에서 클래스 별로 영상을 절취한다. 클래스는 논, 고구마, 고추, 옥수수, 깻잎, 과수, 비닐 하우스로 7가지로 나누었으며, 그림 5와 같이 트레이닝을 위해 영상의 크기를 200x200으로 정규화하였다. 오버피팅을 피하기 위해 트레이닝, 밸리데이션, 테스트 셋을 표 2와 같이 구성하였다.



그림 5. 이미지 사이즈 정규화

Fig. 5. Image Size Normalization

표 2. 데이터 셋 구성

Table 2. Organization of Data Set

구분	Train	Validation	Test
corn	191,712	60,544	60,544
fruit_tree	198,998	62,203	62,203
pepper	198,532	65,120	65,120
perilla_leaf	190,144	60,428	60,428
rice_paddy	192,789	65,446	65,446
sweet_potato	193,612	60,519	60,519
vinyl_greenhouse	199,516	62,366	62,366

본 연구에서 사용한 CNN 모델은 그림 6과 같다.

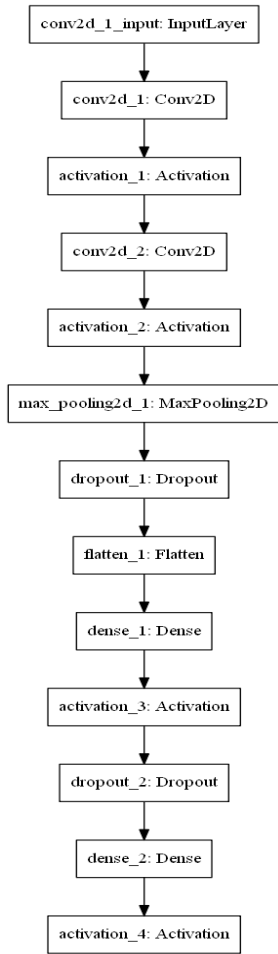


그림 6. 컨볼루션 신경망 모델 순서도
Fig. 6. Flowchart of CNN Model

입력 영상으로는 클래스 별로 잘려진 200x200 크기의 패치 영상을 사용하며, 컨볼루션층에서는 3x3 크기의 커널을 사용하여 특징 맵을 생성한다. 활성화 함수는 $\max(0,x)$ 와 같이 음수에 대해서만 0으로 처리하는 ReLU 함수를 사용했다. CNN 모델에서 주로 사용하는 풀링을 위해서, 샘플 값에서 가장 큰 값을 추출하는 맥스 풀링(Maxpooling)을 사용하였고, 풀링 크기는 3x3 이다. 본 논문에서는 7가지 농작물을 분류하는 분류기를 구성하기 때문에 최종 출력으로 작용하는 완전연결층은 각 클래스에 해당하는 7개의 노드로 구성된다. 따라서 학습을 위해 마지막 활성화(Activation) 함수는 소프트맥스 알고리즘을 적용하였다^[11]. 표 3에 모델에 대한 정보를 정리하였다.

표 3. CNN 모델 정보

Table 3. Summary of CNN Model Information

Layer (Type)	Output Shape	Parameter
Convolution2D	(2, 198, 198)	320
Activation	(32, 198, 198)	0
Convolution2D	(32, 196, 196)	9248
Activation	(32, 196, 196)	0
MaxPooling2D	(32, 98, 98)	0
Dropout	(32, 98, 98)	0
Flatten	(307328)	0
Dense	(128)	39338112
Activation	(128)	0
Dropout	(128)	0
Dense	(7)	903
Activation	(7)	0

그림 7은 트레이닝 과정에서의 정확도(Accuracy)와 손실(Loss)의 변화를 보여주고 있으며, 테스트 결과의 민감도(Sensitivity)와 특이도(Specificity)의 관계를 확인하기 위해 Receiver-Operating Characteristic Curve 그래프를 그림 8에서 보여준다.

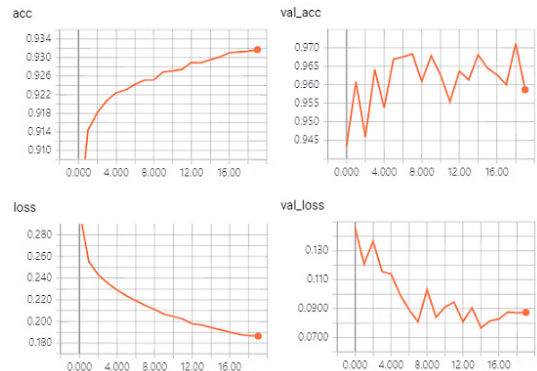


그림 7. 트레이닝 과정에서의 정확도와 손실 변화

Fig. 7. Accuracy and Loss Changes during Training

본 연구에서 수행한 실험 결과에 따르면 CNN으로 훈련된 분류기는 98% 이상의 정확도를 보여주었다. 따라서 영상분류에 있어 기존 머신 러닝 기법보다 우수한 성능을 보인다는 것을 드론으로 촬영한 농경지 분류 실험을 통해서도 입증 가능했다.

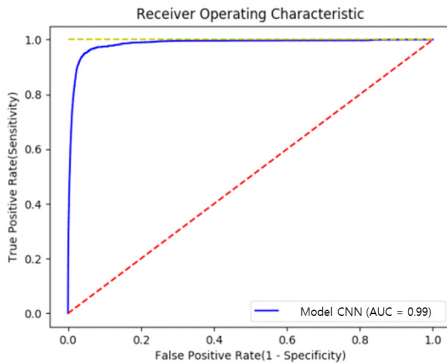


그림 8. ROC 곡선
 Fig. 8. Receiver-Operarting Characteristic Curve

IV. 결론

본 논문에서는 드론으로 촬영한 농경지 분류에 최근 딥 러닝 기법 중에 일반적으로 사용되는 컨볼루션 신경망을 적용하여 그 가능성을 검토하였다. 학습 훈련을 통해 농경지를 논, 고구마, 고추, 옥수수, 갯잎, 과수, 비닐하우스의 7가지 클래스로 구분해내는 분류기는 정확도에 있어서 98% 이상의 성능을 보였다. 현재 수작업을 통해 생성된 농경지 분류 지도를 자동화할 수 있음은 물론 분류 지도의 갱신 주기도 혁신적으로 앞당길 수 있어 선진 스마트 농업정책을 펼치는데 결정적인 역할을 할 것으로 예상된다. 따라서 향후 다양한 딥 러닝 기법의 적용을 통해 농경지 분류를 위한 최적 분류기를 생성하는 것을 기대하고 있다.

References

[1] M. Langkvist, A. Kiselev, M. Alirezaie, and A. Loutfi, "Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks", *Remote Sensing*, Vol. 8, pp. 329-350, 2016.
 DOI: 10.3390/rs8040329

[2] G. Fu, C. Liu, R. Zhou, T. Sun, and Q. Zhang, "Classification for High Resolution Remote Sensing Imagery Using a Fully Convolutional

Network", *Remote Sensing*, Vol. 9, pp. 498-519, 2017.
 DOI: 10.3390/rs9050498

[3] M. Castelluccio, G. Poggi, C. Sansone, and L. Verdoliva, "Land Use Classification in Remote Sensing Images by Convolutional Neural Networks", Available online: <https://arxiv.org/abs/1508.00092>(accessed on 12 April 2016).

[4] D. Choi and J. Park, "A Study on Security Authentication Vector Generation of Virtualized Internal Environment using Machine Learning Algorithm", *The Journal of the Institute of Internet Broadcasting and Communication(JIIBC)*, Vol. 16, No. 6, pp. 33-42, 2016.
 DOI: <https://doi.org/10.7236/JIIBC.2016.16.6.33>

[5] A. Karpathy, "Neural Networks Part 1: Setting Up the Architecture." Notes for CS231n Convolutional Neural Networks for Visual Recognition, Stanford University, <https://cs231n.github.io/neural-networks-1/>, 2017.

[6] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 55, Issue. 2, pp. 645-657, 2017.
 DOI: 10.1109/TGRS.2016.2612821

[7] Deep Learning, Available online: https://en.wikipedia.org/wiki/Deep_learning (accessed on 3 May 2017).

[8] Y. Zhong, F. Fei, Y. Liu, B. Zhao, H. Jiao and L. Zhang, "SatCNN: Satellite Image Dataset Classification using Agile Convolutional Neural Networks", *Remote Sensing Letters*, Vol. 8, Issue 2, pp. 136-145, 2017.
 DOI: <https://dx.doi.org/10.1080/2150704X.2016.1235299>

[9] R. Socher, B. Huval, B. Bath, C. D. Manning, and A. Y. Ng, "Convolutional-recursive Deep Learning for 3D Object Classification", *Advances in Neural Information Processing Systems: Curran Associates: North Miami Beach, FL, USA*, pp. 665-673, 2012.

- [10] Z. Zhang, "A Flexible New Technique for Camera Calibration", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, Issue. 11, pp. 1330-1334, 2000.
DOI: 10.1109/34.888718
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet Classification with Deep Convolutional Neural Network", Advances in Neural Information Processing Systems 25, pp. 1106-1114, 2012.

저자 소개

주영도(정회원)



- 1983년 한양대학교 전자통신공학과 학사
- 1988년 미국 University of South Florida 컴퓨터공학과 석사
- 1995년 미국 Florida State University 전산학과 박사
- 1996년 ~ 2000년 KT 연구개발본부

선임연구원

- 2000년 ~ 2005년 시스코 시스템즈 코리아 상무
- 2005년 ~ 2006년 화웨이 기술 코리아 부사장
- 2007년 ~ 현재 강남대학교 컴퓨터미디어정보공학부 교수
<관심분야: 신경망, 머신러닝, 정보검색>

※ 본 연구는 (2015학년도) 강남대학교 교내연구비 지원에 의해 수행되었음