

딥러닝을 이용한 양파 밭의 잡초

검출 연구

(Deep learning-based Automatic Weed Detection on Onion Field)

김서정*, 이재수**, 김형석***

(Seo jeong Kim, Jae Su Lee, Hyong Suk Kim)

요약

이 논문은 양파 밭에서 딥러닝 기반 자동 잡초 검출기의 설계 및 구현을 제시합니다. 이 시스템은 컨볼루션 뉴럴 네트워크를 기반으로 제안된 영역을 선택합니다. 검출기는 양파 밭에서 직접 찍은 데이터 셋을 가지고 훈련됩니다. 학습이 완료된 후에, 잡초가 될 확률이 매우 높은 후보 지역을 잡초로 간주합니다. Non-maximum suppression을 통해 오버랩된 박스가 최대한 적게 남게 됩니다. 다른 양파 농장을 통해 수집된 데이터를 통해 제안된 분류기를 평가합니다. 분류 정확도는 고려된 데이터 셋에서 약 99%를 보여주며, 제안된 방법이 양파 밭에서 잡초 검출과 관련하여 우수한 성능을 나타낼 수 있습니다.

■ 중심어 : 잡초 분류; 잡초 검출; 컨볼루션 뉴럴네트워크; 딥러닝;

Abstract

This paper presents the design and implementation of a deep learning-based automated weed detector on onion

fields. The system is based on a Convolutional Neural Network that specifically selects proposed regions. The detector initiates training with a dataset taken from agricultural onion fields, after which candidate regions with very high probability of suspicion are considered weeds. Non-maximum suppression helps preserving the less overlapped bounding boxes. The dataset collected from different onion farms is evaluated with the proposed classifier. Classification accuracy is about 99% for the dataset, indicating the proposed method's superior performance with regard to weed detection on the onion fields.

■ keywords : weed classification ; weed detection ; convolutional neural network(CNN) ; deep learning

I. 서론

잡초의 인식은 부분적으로 잡초를 캐거나 제초제를 살포하여 제거하기 위해 필요한 일이다. 이렇게 하면, 제초제 살포량을 최소화 할 수 있으며, 작물들이 손상되지 않고, 화학약품에 노출을 최소화 할 수 있게 하여 건강한 작물을 키울 수 있다. 이를 위해서는 작물과 구별하여 잡초를 식별하는 기술개발이 필수적이다. 기존의 영상 분류기술은 작물과 잡초 구별 기술이 매우 떨어지기 때문에 전혀 효과를 보지 못하였다.

최근 CNN(Convolution Neural Network), DNN 등 딥러닝 기반의 스마트 인공지능 기술이 다양한 영역에 활용되고 있다. 이 딥러닝 기술 사용에 있어서 우선적으로 중요한 일은 학습 데이터를 준비하는 일이다.

이를 위해서 잡초를 포함하는 영상에 라벨을 달아주는 일이 꼭 필요하다. 예를 들면, 잡초를 포함하는 영역에 대해서는 양의 값으로 주석을 달아주고, 잡초를 제외한 모든 영역 즉, 흙이나 양파 등에 대해서는 음의 값으로 주석을 달아 준다. 말하자면, 어떤 영역에 음의 값으로 주석이 달려있다면, 그 영역 내에는 잡초가 존재하지 않다는 표시를 한다는 것이다

신경회로망의 입력은 일정하게 고정되어 있으므로, 입력 영상은 신경회로망의 입력 크기에 맞게 정해져야 한다. 이렇게 신경회로망 입력 영상을 규격화하기 위해서는, 원 입력 영상 위에 일정한 크기의 슬라이딩 윈도우를 중첩하여 윈도우와 같은 크기의 부영상을 읽어내어 입력영상으로 한다. 이 때, 이 부영상들은 잡초인지 여부에 따라 양 혹은 음의 레이블

* 학생회원, 전북대학교 전자공학부

** 정회원, 농공학부, 농업과학원, 농촌진흥청

*** 정회원, 지능형 로봇 연구소, 전북대학교

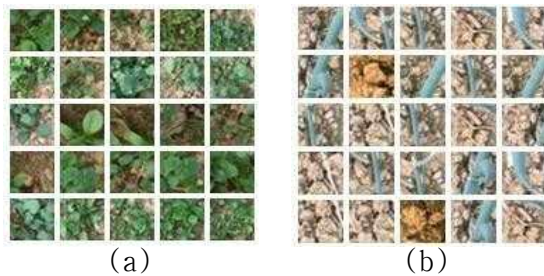


그림1. 양파 밭의 잡초(양의 레이블링)와 작물(양파, 음의 레이블링)의 영상

을 달아 놓는다. 각 부 영상에 대한 양 혹은 음의 레이블링을 한 예는 그림 1 과 같다.

이 논문에서는 CNN 학습을 사용하였다. CNN 은 영상을 표현하는데 사용되는 매우 큰 특징들 집합으로부터 매우 작은 특징들을 생산해 내는 고도화된 특징 추출장치이다.

이 논문은 다음과 같이 구성되었다. II 절에서는 이 연구와 관련된 기존의 연구내용들이 기술되어 있고, III 절에서는 학습데이터 세트의 구조를 설명하고, IV 절에서는 제안하는 신경회로망 구조와 학습 방법을 설명한다. 또, V 절과 VI 절은 각각 실험결과와 결론 부분이다.

II. 관련 연구

영상인식은 농업분야에서 있어서도 잡초를 제거하기 위한 목적으로 많이 사용되어왔는데, 주로 RGB 및 다 채널영상을 사용하여 작물과 잡초를 식별 하려는 연구들이었다[1][2][3]. 그러나, 고전적인 방법으로는 영상인식에 한계가 있어서, 진척을 보지 못하다가 근래에 인공지능과 함께 큰 발전을 하기에 이르렀다. Hall 등은 Flavia leaf 의 데이터세트를 대상으로 기존의 수작업을 통하여 분류하는 방법과 CNN을 사용한 연구를 수행하였고, CNN 사용방법이 강건성과 성능면에서 기존의 방법에 비해 우월하다는 결론을 얻었다 [4]. Chen 등도 CNN 을 사용하여 다양한 조명 상태 및 잎과 과일들이 겹쳐지는 복잡 환경하에서 과일의 수를 세는 연구를 발표하였다 [5]. 이를 위해서 먼저 색깔 영역을 추출하였으며, 두 번째 CNN 을 사용하여 각 영역 내에서의 과일을 추출한 다음, linear regression 모델을 사용하여 과일 수를 세었다.

또, Potena 등은 [6] RGB+NIR 영상을 대상으로 작물과 잡초 구별연구를 수행하였다. 이 연구에서는 2 개의 다른 CNN 을 사용하였는데, 첫 번째의 얇은 CNN 에서는 식물체 영역을 검출하고, 두 번째의 깊은 CNN 에서는 이 식물체를 정밀 분류할 수 있도록 했다. 이를 위해서, 우선 픽셀 단위의 분류를 수행하였고, 이를 통해 얻은 식물체의 무더기들이 어떤 종류에 해당하는지 분류

할 수 있게 하였다. 이 연구결과로 약 97%의 식물 영역 검출 성능을 얻었는데, 이 결과는 NDVI(Normalized Difference Vegetation Index)에서 얻은 문턱치 사용결과와 견줄만한 것이었다. 또, 작물과 잡초의 분류 결과로는 98%의 성능을 얻었다. 또, Lee [7]등과 Reyes[8] 등도 CNN 을 사용하여 식물 분류 시스템 연구결과를 발표하였다.

영상 분류에 있어서 위치불변 특징과 위치가변 특징들 사이에는 서로 배치되는 면이 있다. 즉, 딥 CNN은 위치적으로 많이 변화해 있다 해도 인식이 가능하다는 특징이 있다. 그러나, 물체의 회전은 위치적 변화를 야기한다. 즉, 물체가 변화하면 바운딩 박스간에 적절한 상관관계가 형성되어 있어야 한다

본 연구에서는 작물영상에서 잡초들을 검출하여 그 위치를 찾아내는 연구를 수행하였다. CNN 의 깊이가 깊으면, 위치 변화에 덜 민감한 것으로 가정하여, 이 논문에서는 잡초의 존재를 검출하는 분류기를 사용하였다. 이 분류기는 물체의 위치를 찾아 내는 선택적 탐색과 특징 검출로 시작되는데, 마지막 단계에서는 Softmax를 사용하여 잡초를 분류하게 하였다.

특히, 본 연구는 식물들이 심하게 겹쳐있는 환경에서 CNN 을 사용하여 양파와 작물을 구분하는데 기여를 하였다.

III. 제안 접근법

1. 학습 데이터 준비

제안한 분류기를 위한 학습 데이터는 여러 다른 위치의 양파 농장으로부터 촬영한 영상을 사용하였다. 이 영상들은 28x28 크기로 잘라서 잡초 영역과 양파 영역으로 구분하였다. 이 크기는 학습 시의 계산 복잡성을 줄이면서, 잡초와 양파의 분류가 시각적으로 가능한 크기이다. 잘라낸 패치 영상들은 잡초영역 인지 양파영역인지로 분류하여 라벨을 붙이게 하였다. 그 다음, 물체가 바운딩 박스에 포함되어 있는지 확인하고, 잡초의 넓이가 영상의 80%보다 작고, 잡초가 길이 방향으로 80% 이하인지 확인한다. 또, 잡초의 넓이나 높이가 영상의 60% 이상이 되도록 한다. 슬라이딩 윈도우가 한 번에 8 픽셀을 움직이게 하고, 줌 레벨은 1.414 를 최대가 되게 하여 관심물체가 포함 되도록 한다. 특히 밀집한 잡초가 있는 곳에서 집중하여 데이터를 추출하였다. 이렇게 준비된 학습 영상의 수는 양파 부의 라벨이 공히 2293 개의 동일한 수를 준비하였다.

2. 분류기의 구성

CNN 은 영상의 특징추출 시 인간의 개입이 필요하여 시간 및 노력의 소모가 과도했던 기존 알고리즘 방법과는 달리, 영상으로부터 학습 과정에 의해 자동적으로 영상의 특징을 추출해 내는 방법이다. 제안한 CNN 구조에서는 다른 컨볼루션 구조에서처럼 컨볼루션과 풀링 층이 교대로 연결되어 있는데, 컨볼루션 층은 주어진 데이터들을 학습하기 위한 학습 가능한 모듈로 구성되어 있다. 이 컨볼루션 층은 칼라정보나, 에지, 무늬등의 특징정보를 학습하는 역할을 한다.

다음에 오는 풀링 층은 특징정보를 줄이기 위한 것인데, 특징들을 분할하고 각 영역의 통계 정보를 얻는 역할을 한다. 그림2는 연구에서 사용한 분류기 구조이다.

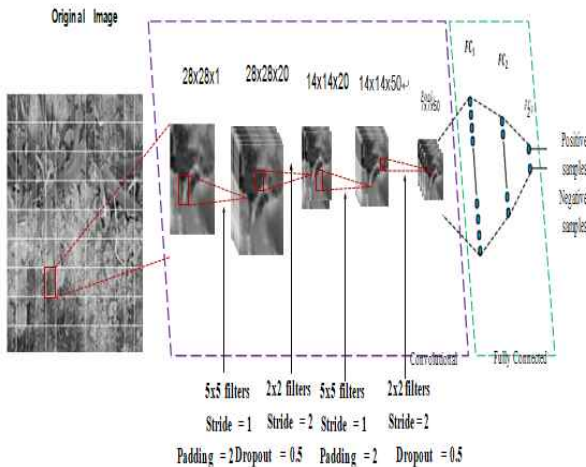


그림 2 사용한 딥 컨볼루션서널 Network 구조

제안한 분류기의 입력은 3x3 크기의 20 개의 필터로 구성되고, 활성화 함수로 RELU 함수를 사용한다. 이 층은 28x28 pixels 을 가진 영상을 입력 받아, 1 stride, 2 padding 에 의한 5x5 filtering 을 하여, 28x28x20 크기의 특징맵을 생성한다. 그 후, max-pooling 과정을 거쳐 14x14x20 크기의 특징맵으로 축소함으로써, 계산량을 줄인다. 이 축소과정에서 50% "dropout" 하여 over fitting 문제를 완화한다. 컨볼루션, activation, max pooling 과정을 반복하여, 원래 크기의 1/2 이되게 한다 (면적으로는 1/4). 그 후, 2 개의 완전 연결 층 (fully connected layer)을 연결한다. 이를 위해서 전 단의 영상은 1 열로 정렬 "flattened"하며, 중간 층에 128 개의 뉴런, 출력 층에 2 개의 뉴런을 갖는 신경망에 연결한다. 출력 층에서의 한 개의 뉴런은 잡초, 다른 뉴런은 비잡초를 표시하게하여, softmax 에 의해 결정하게 한다. 예, 양의 값은 잡초의 존재, 음의 값은 잡초가 아님을 의미함. 이 회로망에서 사용한 총 파라미터 수는 1,252,072 개 였는데, 자세한 파라미터는 Table 1 에 보여준다. 여기서 일반화 특성을 개선하고, overfitting을 완화시키기 위해서 "dropout"을 사

용하였다.

제안한 CNN 기반의 분류기는 28x28 크기의 4586 개의 학습 영상에 대해 학습시켰다. 이 중 절반 (2293 개)은 잡초이고, 나머지 반은 잡초가 아닌 데이터가 되게 하였다. 이 중 잡초와 작물 중 각 20% 를 무작위로 추출하여 테스트 데이터로 사용하였다. 이 분류기를 테스트하기 위해서 28x28 크기의 sliding window 를 사용하여 입력 영상을 취하였으며, 물체가 검출되는 경우 이 좌표를 따로 기록해 두었다.

이 때, 각 입력 영상마다 그 영상의 위치 좌표와 함께 입력 영상이 잡초일 확률과, 영상에서 각 위치 점이 잡초일 확률을 기록하였다. 이 실험의 초기에는 CPU 를 사용했지만, GPU 를 사용함으로써 학습을 고속화하고, 프로그램을 단순화 하였다. Table 1 에서 보는 바와 같이, 이 구조는 1.25 M 개의 파라미터로 네트워크를 구성할 수 있었다.

표1. 학습에 사용한 네트워크의 파라미터

Layer(type)	Output Shape	Parameter #
Conv 2D	(None,20,28,28)	520
Activation	(None,20,28,28)	0
MaxPooling	(None,20,14,14)	0
Dropout	(None,20,14,14)	0
Conv2D	(None,50,14,14)	25050
Activation	(None,50,14,14)	0
MaxPooling	(None,50,7,7)	0
Flatten	(None,2450)	0
Dense	(None,500)	1225500
Activation	(None,500)	0
Dense	(None,2)	1002
Activation	(None,2)	0

3. 대상 영역 추출

제안한 구조의 학습이 완료되면, 테스트 영상을 입력한다. 앞 절에서 기술한 바와 같이 입력 영상은 sliding window 를 사용하여 부분영상을 취하고, 이를 신경회로망에 입력하여 잡초 여부를 결정하여, 잡초 여부를 이진 값으로 출력한다. 이 값은 영상의 위치와 함께 따로 저장함으로써, 후에 bounding box 를 생성할 수 있게 한다. 그림 3 은 영상처리 과정을 설명한 것이다. 225x225x3 크기의 원영상 R, G, B 에 대한 가중치 합산을 통하여 100x100x3 크기로 줄인다.

그 후, 28x28 크기의 sliding window 를 사용하여 부분 영상을 취한다. 이 영상 조각을 분류기의 입력으로 인가하는데, convolution-activation-pooling 과정을 거쳐서 50 개의 feature 들을 만든다. 이 특징 값들은 fully connected network 에 인가되어 최종 분류되게 한다. 이 때, 출력 값이 99%이상이면, 출력은 이진 수 1 이 되게 하여 잡초로 결정하고, 이 이외의 값일 때는 0 으로 하여 비 잡초 영역임을 표시한다.

그림 3 에서 보는 바와 같이, 제안한 구조는 영상에 있는 대부분의 잡초를 잘 인식할 수 있었다. 그림3에서 볼 수 있듯이 처음 영상을 적당한 크기로 바꾼 뒤에 RGB인 영상을 흑백 영상으로 바꾸는 작업을 하였고 Sliding-window를 통해 잡초가 있을 것이라고 생각하는 여러 후보 지역에 박스를 치고 Non-Maximum Suppression을 통해 가장 유력한 지역(잡초 부분)에 박스를 쳐주는 것을 확인 할 수 있다. 그림에서 보이는 박스들은 잡초의 존재할 99% 이상 확실한 곳이다. 이렇게 문턱 치를 높게 잡은 이유는 학습 데이터에는 잡초 영상과 비잡초 영상의 비율이 거의 동일하게 해야 학습이 잘 되지만, 실제 테스트 영상에서는 잡초의 영역이 훨씬 드물게 나타나기 때문이다. 만약, 50%를 문턱치로 사용한다면,분류기는 잡초에 대해서 false positive 결과를 많이 출력하게 된다. 또,잡초 detecting 과정에서“non-maximum suppression”기법을 사용했는데, 입력영상 추출 시에 동일한 영역이 다른 입력 영상 범위에 여러 번 포함되기 때문에 인접한 위치의 bounding box 가 겹쳐서 나타나는 현상을 방지하기 위해서 이다. 이 기술은 여러 개의 bounding box 가 겹쳐 나타나는 구역 내에서 가장 큰 확률을 갖는 bounding box 만 남기고 다른 bounding box 들을 제거하는 기술이다.

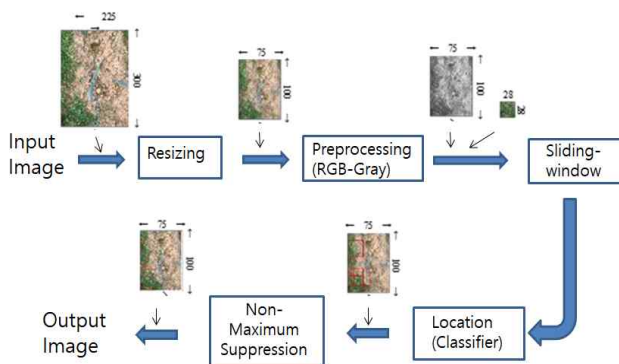


그림 3. 잡초 검출 절차

IV. 실험결과

CNN 을 사용한 딥러닝 학습framework에는 속도, 구현의 용이성, 구조변경 가능성 및 새로운 계층의 생성 용이성 등에 따라 달라진다. 이 논문에서는 구현이 용이한 Keras library 를 사용해서 컨벌루션 망을 구성했다. 수행한 실험 결과, 구성된 분류기가 정확하고, 시스템 구현이 용이하며, 기존에 배운 내용을 활용해서 용이하게 구현할 수 있다는 것을 보여준다. 수행한 연구에서는 Stochastic Gradient descent

(SGD)과 mini-batch size 50 그리고, epoch size 를 1000 로 하였다. Weight decay 값은 0.000001 로 하였으며, momentum 은 0.9, 학습률은 0.01 이었다. table 2 는 총 4000 회를 학습시킨 후, precision, recall, F1 score 를 측정된 결과

이다.

표 2. Precision, Recall, F1-Scores and support.

	Precision	Recall	F1—score	support
class 0 (negative samples)	0.99	0.98	0.99	513
class 1 (positive samples)	0.98	0.99	0.99	405
Avg/total	0.99	0.99	0.99	918

table 2를 분석해 보면 precision, recall, f1-score 의 모든 값이 0.98에서 0.99 구간의 값을 갖고, 정확도 (accuracy)도 0.99가 됨을 알 수 있다. 학습의 질을 입증할 수 있는 또 다른 유용한 tool은 그림 4과 같은 confusion matrix이다.

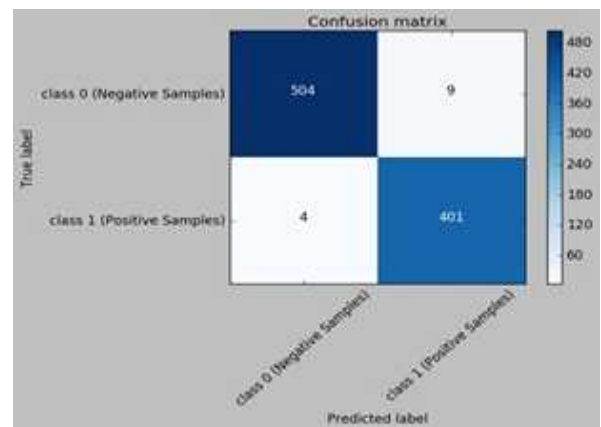


그림 4. 컨퓨전 매트릭스

여기서 True positive 는 어떤 물체라고 결정한 것이 실제로 true 인 경우이고, False positive 는 어떤 물체라고 한 것이 틀린 경우이고, True Negative 는 어떤 물체가 아니라고 주장한 말이 맞는 경우이고, False Negative 는 어떤 물체가 아니라고 주장한 말이 틀린 경우이다. 실험 결과, 총 513 개의 잡초가 아닌 식물에 대해서 504 개가 실제로 잡초가 아니었다. 또, 401 개의 식물에 대해서 잡초를 잡초라고 했고, 4 개가 아닌 것을 잡초라고 했다. 결과적으로 99%의 분류 정확성을 보였다. 그림 5 는 제안한 분류기의 최종 출력들의 예를 보여준다. 잡초들의 영역이 양파와 분리되어 잘 검출 됨을 알 수 있다.



그림 5. 제안한 분류기에 의해 양
파밭에 잡초 검출 예

V. 결론

잡초를 작물과 구별하는 것은 여러 가지 점에서 장점이 있다. 잡초의 양에 따라서 잡초들을 적절히 다룰 수 있을 뿐 아니고, 작물에 해를 주는 잡초들에 대해서 선택적으로 제거를 할 수 있다. 제조제를 사용하는 경우에도 잡초에 대해서만 소량의 제조제를 사용하여 제거할 수 있으므로, 환경에 주는 영향을 최소화 할 수 있다. CNN은 물체의 시각적 인식에 매우 매우 효과적인 신경 회로망이다. 제안한 시스템은 이 CNN 을 양파와 잡초를 구별하는데 응용하였다. 이를 위하여, 입력 영상을 인가할 때, 신경회로망이 적절히 동작할 수 있도록 초기화를 하였다. 구성된 신경회로망 구조는 파라미터 수가 많지 않는 구조로 구성되었다. 원 영상에서 28x28 크기의 부분영상 중, 2293 개는 잡초영상에서 같은 수를 비 잡초 영상에서 취하여 준비하였다. 그 중 무작위로 추출한 405 개의 잡초영상과 513 개의 비 잡초 영상을 테스트 영상으로 하였다.

실험 결과 0.99 의 정확도를 갖는 결과를 얻음 으로서, 제안한 시스템의 유용성이 입증되었다. 이 시스템은 제조제의 사용을 최소화할 수 있는 효과가 있을 것이다.

REFERENCES

- [1] Guo, W., Rage, U. K. and Ninomiya, S., "Illumination invariant segmentation of vegetation for time series wheat images based on decision tree model," *Computers and Electronics in Agriculture*, vol. 96, pp. 58 - 66, Aug. 2013.
- [2] Hamuda, E., Glavin, M. and Jones, E., "A survey

of image processing techniques for plant extraction and segmentation in the field," *Computers and Electronics in Agriculture*, vol. 125, pp.184 - 199, July 2016.

- [3] Torres-Sanchez, J., López-Granados, F. and Peña, J., "An automatic object-based method for optimal thresholding in uav images: Application for vegetation detection in herbaceous crops," *Computers and Electronics in Agriculture*, vol. 114, pp. 43 - 52, July 2015.
- [4] Hall, D., McCool, C., Dayoub, F., Sunderhauf, N. and Upcroft, B., "Evaluation of features for leaf classification in challenging conditions," *2015 IEEE Winter Conference on Applications of Computer Vision*, pp. 797 - 804, Marriott Resort, USA, Jan. 2015.
- [5] Chen, S., Skandan, S., Dcunha, S., Das, J., Qu, C., Taylor, C. and Kumar, V., "Counting apples and oranges with deep learning: A data driven approach," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 781-788, Jan. 2017.
- [6] Potena, C., Nardi, D. and Pretto, A., "Fast and accurate crop and weed identification with summarized train sets for precision agriculture," *Proc. of Int. Conf on Intelligent Autonomous Systems (IAS)*, pp. 105-121, 2016.
- [7] S. H. Lee, C. S. Chan, P. Wilkin, and P. Remagnino, "Deep-plant: Plant identification with convolutional neural networks," *Proc. Of the IEEE International Conference on Image Processing (ICIP)*, pp. 452 - 456, Dec. 2015.
- [8] K.Reyes, J.C.Caicedo, and J.E. Camargo, "Fine-tuning deepconvolutional network for plant recognition," *Working Notes of CLEF 2015 conference*, vol.1391, pp.121, 2015
- [9] 강현중, 이준욱, 강성수, "스마트 디바이스 기반의 농업환경관리 지원 서비스," *스마트미디어저널*, 제1권, 제1호, 46-51쪽, 2012년 3월
- [10] 문해민, 박진원, 반성범, "역전파가 제거된 CNN과 LDA를 이용한 얼굴 영상 해상도별 얼굴 인식률 분석," *스마트미디어저널*, 제5권, 제1호, 24-29쪽, 2016년 3월
- [11] 박동주, 김병우, 정영선, 안창욱, "Deep Neural Network 기반 프로야구 일일 관중 수 예측 : 광주-기아 챔피언스 필드를 중심으로," *스마트미디어저널*, 제7권, 제1호, 16-23쪽, 2018년 3월
- [12] 박선, 김종원, "오픈 소스 기반의 딥러닝을 이용한 적조생물 이미지 분류," *스마트미디어저널*, 제7권, 제2호, 34-39쪽, 2018년 6월

저 자 소 개

**김서정(학생회원)**

2018년 전북대학교 전자공학부
학사 졸업.

<주관심분야 : 인공지능, 딥 러닝,
영상처리>

**이재수**

2010년 전북대학교 생물산업기계공학과 학사졸업.

2012년 동 대학원 석사졸업.

2015년 동 대학원 박사수료.

2018년 현재 국립농업과학원 농업공학부 농업연구사. 관심분야는 스마트

팜, 농업 ICT 융복합, IoT 및 환경제어시스템 개발.

**김형석(정회원)**

1979년 한양대학교 전자공학과
학사 졸업.

1981년 전북대학교 전기공학과
석사 졸업.

1992년 University of Missouri,
Columbia 박사.

<주관심분야 : 인공지능, 딥러닝,
영상처리>