

딥 러닝을 활용한 씨앗 발아 확인 시스템

김우주, 권민서, 이재준, 류관희*, 홍장의, 나스리디노프 아지즈

*충북대학교 소프트웨어학과

e-mail : khyoo@chungbuk.ac.kr

Implementation of Seed Germination Confirmation System with Deep Learning

U Ju Gim, Min Seo Kwon, Jae Jun Lee, *Kwan Hee Yoo, Jang-Eui Hong, Aziz Nasridinov

*Dept of Computer Science, Chungbuk National University

요 약

최근 대두되고 있는 딥 러닝은 학습을 통해 사물이나 데이터를 군집화하거나 분류하는 데 사용하는 기술이다. 본 논문은 딥 러닝에 활용하기 위해 개발된 오픈소스 소프트웨어인 텐서플로 Inception V3를 사용해 연구를 진행했다. 딥 러닝을 활용한 씨앗 발아 확인 시스템은 기존의 영상 처리를 활용한 시스템에서 고안했으며, 씨앗 발아 여부의 정확성이 떨어지는 단점을 개선하고, 모든 종자들의 발아 여부를 확인할 수 있도록 구현해 사용자가 효과적으로 연구를 수행할 수 있도록 하는 목적에 있다.

키워드 : 딥 러닝, 텐서플로, Inception V3, 씨앗 발아 확인, 영상 처리 개선

1. 서론

유전적, 식물 발달 및 생리학 연구에서 씨앗의 사용이 증가함에 따라 발아 여부를 확인 할 필요성이 증가했다. 이에 우리는 전 세계에서 큰 인기를 얻고 있는 구글의 딥 러닝 라이브러리인 텐서플로를 이용해 씨앗 발아 여부를 확인하고자 한다. 텐서플로는 오픈 소스 소프트웨어인 만큼 기존 딥 러닝 라이브러리보다 사용이 쉽고 개발을 원하는 사람들이 어떤 제약도 받지 않고 유연하게 사용할 수 있는 기술이다.

본 논문은 텐서플로의 Inception V3를 활용해 씨앗 이미지를 학습시켜 씨앗의 발아 여부를 확인함으로써, 기존의 영상 처리로 진행되었던 씨앗 발아 확인 시스템에서의 정확성이 떨어지는 단점을 개선한다. 또한 다양한 종자들의 발아 여부를 확인할 수 있도록 구현함으로써 사용자가 본 시스템을 이용해 효과적으로 연구를 수행 할 수 있도록 하는 목적에 있다.

2. 관련 연구

본 장에서는 딥 러닝을 활용한 씨앗 발아 확인 시스템과 관련된 연구에 대해서 설명 한다. 현재 씨앗 발아 여부를 확인하는 실험은 대부분 영상 처리를 통해 진행되고 있다. 그 중 ‘A Software Package for High-throughput Scoring and Curve Fitting of Arabidopsis Seed Germination’[1]의 논문은 이미

지 처리를 사용해 시간에 따른 객체의 크기를 계산하고 가중치를 이용하여 씨앗 발아를 판별한다.

위와 유사한 논문으로 ‘영상처리를 활용한 씨앗 발아 확인 시스템 구현’[2]이 존재한다. 본 논문은 Arabidopsis 씨앗에 대해서만 이미지 처리를 통해 씨앗의 윤곽 및 색상을 검출하는 방법으로 발아 유무를 판단하고, 사용자에게 프로그램으로 제시한다.

대부분의 영상 처리를 통해 진행되는 씨앗 발아 확인 실험에 반해 텐서플로의 k-NN알고리즘을 활용해 연구한 ‘Using k-NN to Analyse Images of Diverse Germination Phenotypes and Detect Single Seed Germination in Miscanthus Sinensis’[3]의 논문에서는 다양한 발아 형태에 접근하고 씨앗 개개의 발아 여부 판별법을 제시한다.

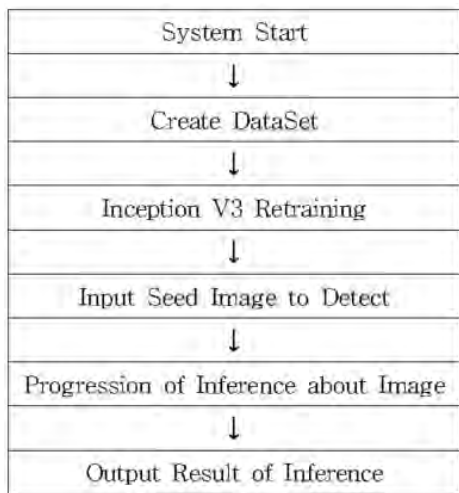
앞서 제시한 논문들은 모두 수작업으로 진행해 오던 씨앗 발아 여부 확인 실험을 자동화하여 사용자 편의성을 증가시킨다. 그러나 단지 자동화기기로 실험을 하거나 외부개체에 의한 정확도 감소는 해결해야 할 문제로 남아있다. 이러한 정확도 문제를 해결하기 위해 딥 러닝을 통해 씨앗 이미지를 학습하여 발아상태를 확인하는 시스템을 제안한다. 더욱이 기존 사용된 씨앗 뿐 아니라 기존에 연구되고 있는 다른 종자들에 대해서도 발아확인이 가능하게 한다.

3. 제안한 방법

이번 장에서는 딥 러닝을 활용한 씨앗 발아 확인 시스템을 구현 시 적용한 방법을 제안한다. 먼저 시스템의 전체적인 흐름을 설명한다. 다음으로 시스템 구현에 사용되는 텐서플로의 이미지 인식 및 추론, Inception V3 모델의 알고리즘에 대해 설명한다.

3.1. 시스템 흐름도

[그림 1]은 본 시스템의 전체적인 흐름을 나타낸다. 우선 Inception V3 모델 알고리즘을 이용하기 위해 우리가 추론하고자 하는 이미지 데이터들의 데이터 셋을 생성한다. 만들어진 데이터 셋을 기반으로 Inception V3 Retraining 기법을 이용해 이미지를 학습한다. Retraining 기법은 구글이 미리 정의한 Inception 모델의 모델 구조와 파라미터들을 재사용해 우리가 생성한 데이터 셋에 적용해주는 효율적인 기법이다. 사용자가 검출하려는 종자 이미지를 Input으로 입력하면 Retraining으로 생성된 모델에 대한 이미지 추론을 진행한다. 최종적으로 추론 결과를 출력해줌으로 사용자에게 결과를 제공해준다.



[그림 1 시스템 전체 흐름도]

3.2. 이미지 추론

본 시스템에서 이미지 추론을 진행하는 알고리즘인 Inception V3는 ImageNet의 Large Visual Recognition Challenge에서 2012년 데이터를 사용하여 훈련된 모델이다. 이 모델은 전체 영상을 “얼룩말”, “식기 세척기”와 같은 1000여개의 클래스로 분류하는 Computer Vision의 표준 작업이며 구글에서 이를 이용하여 Image Recognition API를 제공 중이다. 최근 Inception V3 모델에 몇 가지 기능을 추가해 업그레이드한 모델인 Inception V4가 사용되고 있지만 본 논문에서는 Inception V3을 사용한다.

이미지 추론 진행 과정

```

Input: 추론할 이미지 경로
Output: 가장 높은 5개의 예측 값에 해당하는 label
if 이미지 경로가 존재하지 않으면
    '파일이 존재하지 않습니다' 출력
end if
이미지 데이터 = 이미지 경로
추론 실행
그래프 생성
텐서플로 세션 실행:
    그래프 = 이미지 데이터
    소프트 맥스 실행
    가장 높은 5개의 예측 값에 해당하는 label 반환
    
```

[그림 2 이미지 추론 진행 과정 수도코드]

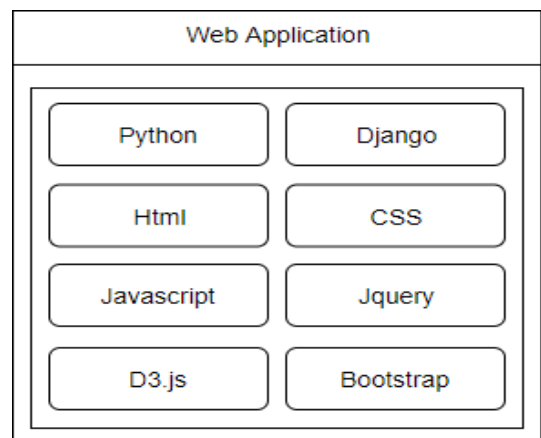
[그림 2]는 이미지에 대한 추론을 진행하는 과정을 수도코드를 활용해 설명한 그림이다. 추론할 이미지를 설정하고 이미지 데이터에 경로를 넣어 추론을 시작한다. 데이터 셋의 그래프 파일로부터 그래프를 생성하고 텐서플로 세션을 실행한다. 각 세션은 그래프에 이미지 데이터를 저장하고, 입력 데이터를 2진수로 나타내 확률로 해석할 수 있도록 도와주는 소프트맥스 기법을 실행 한다. 각각의 소프트맥스 노드에 대해 라벨을 변환 해주고, 가장 높은 5개의 예측 값을 얻어 각 라벨에 해당하는 값을 반환해준다.

4. 구현 및 실험

4.1 구현

본 시스템은 임의적인 종자 데이터 및 실제 종자 데이터를 수집한 후 이에 대해 텐서플로를 사용하여 학습한다. 이 때 Inception V3 모델을 활용하여 발아된 종자와 발아가 되지 않은 종자를 구별해 학습한다. 학습된 모델 기반으로 판단하려하는 이미지에 대해 추론을 진행한 결과를 웹 시스템에서 출력해준다.

[그림 3]은 구현을 위한 시스템 개발 환경을 나타낸 구성도이다. 웹 시스템은 Python으로 이루어진 오픈소스 프레임 워크인 Django 기반으로 구현했다.

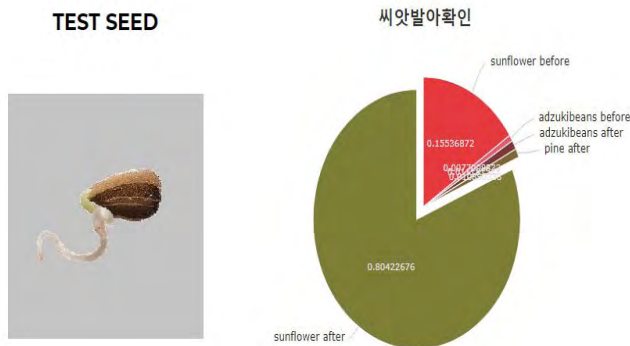


[그림 3 시스템 개발 환경]

4.2 실험 결과

본 연구에서 제안하는 시스템을 실험하기 위해 총 4000여 장의 종자 이미지를 Inception V3를 통해 학습하였다. 이 때, 종자 이미지는 250x250픽셀이며, 아기 장대(Arabidopsis), 팥(AdzukiBean), 해바라기(Sunflower), 소나무(Pine), 완두콩(Pea) 종자로 구성되어 있다. 학습이 완료되면 Retraining된 그래프 형식의 파일과 라벨들이 저장된다. 이 때, 라벨은 종자별로 받아낸 종자인 after와 받아가지 않은 종자인 before로 구성되어 있다. 받아 여부를 확인하려는 종자 이미지를 입력하면 시스템이 Inception V3를 활용해 Retraining된 그래프 파일을 불러와 이미지 추론을 진행한다. 입력한 이미지에 대해 학습된 결과를 바탕으로 가장 높은 확률을 가진 5개의 종자 라벨과 그 확률을 원 그래프 형태로 웹에서 출력해준다.

[그림 4]는 받아낸 해바라기 종자에 대해 이미지 추론을 진행한 결과를 웹에 출력해준 그림이다. 해당 종자가 0.66369444의 가장 높은 확률로 받아낸 해바라기 종자 라벨과 일치하는 것을 확인할 수 있다.



[그림 4 해바라기 종자 이미지 추론 진행 결과]

[표 1]은 본 시스템의 정확성을 검증하기 위해 종자 별로 10개의 테스트 이미지를 사용해 실험한 결과이다. 이 때, 테스트에 사용한 이미지는 받아 된 종자 이미지 5개, 받아가지 않은 종자 이미지 5개이다. 테스트 하려는 종자들에 대해 이미지 추론을 진행해 나온 결과 값으로 이미지 추론 정확도를 계산했다. 정확도를 계산하기 위해 입력한 테스트 이미지와 가장 높은 확률에 해당하는 결과 라벨이 일치하면 이미지 추론에 성공했다고 가정했다. 가장 정확한 추론 결과를 보인 종자는 완두콩과 소나무이며, 가장 부정확한 결과를 보인 종자는 팥이다. 결과적으로 실험에 사용된 종자들은 평균 85.71%의 정확도를 보였다.

[표 1 시스템 정확도 실험 결과]

순번	테스트횟수	종자	정확도
1	10	팥	60%
2	10	완두콩	100%
3	10	소나무	100%
4	10	해바라기	80%
5	10	애기장대	90%
			평균
			85.71%

5. 결론

기존 영상 처리를 활용한 씨앗 발아 시스템을 개선해 연구한 딥 러닝을 활용한 씨앗 발아 확인 시스템은 텐서플로 Inception V3를 사용해 받아 여부를 정확하게 확인할 수 있다. 또한 다양한 씨앗 품종과 저수준 이미지에서도 적용 가능하도록 구현하였다.

접근성이 뛰어난 웹 기반 소프트웨어로써 사용자가 발아 확인 여부를 원하는 플랫폼에서 사용 가능하고, 결과를 직관적으로 확인할 수 있다. 하지만 시스템에서 생성되는 오차율을 낮춰 보안을 해나가야 할 필요성이 있다. 식물 분야 연구를 진행 할 때 필요한 인적, 시간적 자원을 위 시스템을 활용하면 더욱 효과적으로 개선할 수 있을 것으로 예상된다.

'ACKNOWLEDGEMENT

"본 논문은 교육부가 지원하고 충북대학교가 수행하는 지역선도대학육성사업의 지원을 받아서 수행되었습니다."

참고문헌

[1] Danny Awty-Carroll et al. "Using k-NN to analyse images of diverse germination phenotypes and detect single seed germination in Miscanthus sinensis", 2018
 [2] Jae Jun Lee et al. Implementation of seed germination confirmation system with image processing, 2018
 [3] Ronny Joosen et al. GERMINATOR: a software package for high-throughput scoring and curve fitting of Arabidopsis seed germination, 2010