

마약용 양귀비와 관상용 양귀비의 분류 알고리즘 모델에 관한 연구

강운석, 김우현, 이성호, 윤병주
경북대학교 IT 대학 전자공학부

2001kys@gmail.com

Classification Algorithm Model for Narcotic Poppies and Ornamental Poppies

Yun seok kang, Woo Hyun Kim, Seong Ho Lee, and Byoung-Ju Yun
School of Electronics Engineering, Kyungpook National University

요 약

본 논문은 컨볼루션 신경망(CNN)을 사용하여 마약용 양귀비와 관상용 양귀비를 구별하는 분류 알고리즘을 제시하였다. 마약용 양귀비의 비밀 재배가 증가함에 따라 마약 성분이 없는 장식용 품종과 이를 구별할 필요가 있다. 7 개의 양귀비 종에 대한 데이터를 수집하였고, 이 데이터셋을 확장하기 위해 두 가지 데이터 증강 기법을 적용하였다. 이후 다양한 버전의 EfficientNet 을 기반으로 모델을 구축하였다. 훈련된 모델은 약 89.21%의 정확도로 양귀비 종류를 구분하였고 97.93%의 정확도로 마약용과 관상용 양귀비를 구분하여 그 효과를 입증하였다. 미래 연구에서는 양귀비의 꽃, 열매, 줄기 부분을 별도로 탐지 및 비교 분석하여 정확도를 향상시키는 것을 목표로 하였다.

I. 서 론

본 논문은 Convolutional Neural Network(CNN)를 활용하여 마약용 양귀비와 관상용 양귀비를 구분하는 분류기 알고리즘을 제시한다. 현재는 전문가의 육안 검사를 통해 관상용 양귀비와 마약용 양귀비를 구분하지만, 이는 비효율적이며 오류 가능성이 있다. CNN 은 이미지 인식 분야에서 뛰어난 성능을 보이는 딥러닝 기법으로, 본 연구에서는 이를 바탕으로 두 종류의 양귀비를 정확하게 구분하는 모델을 개발하는 것을 목적으로 한다.

이 연구는 불법 마약 재배 방지와 농작물 관리 효율성 향상에 기여할 것으로 기대되며, 향후 양귀비 분류 시스템 구축에 유용한 데이터를 제공할 수 있을 것이다.

II. 본론

본 논문에서는 양귀비를 마약 양귀비 2 종 나도 양귀비 (Papaver Setigerum), 양귀비 (Papaver somniferum)와

나머지 관상용 양귀비 5 종 두메 양귀비 (Papaver radicum), 고산 양귀비 (Papaver alpinum), 개 양귀비 (Papaver rhoeas), 속근 양귀비 (Papaver orientale), 줌 양귀비 (Papaver dubium)로 총 7 개의 카테고리로 분류한다.



그림 1. 나도 양귀비, 양귀비, 두메 양귀비, 고산 양귀비, 개 양귀비, 속근 양귀비, 줌 양귀비

양귀비 이미지 데이터 셋은 GBIF (Global Biodiversity Information Facility) 웹사이트에서 API 를 호출하여 다운로드 하였다[1]. 다운로드한 데이터 셋에서 꽃과

관련이 없는 이미지를 제거하고 가로 세로 비율을 1:1 로 맞추어 이미지를 잘라주는 과정을 진행하였다.

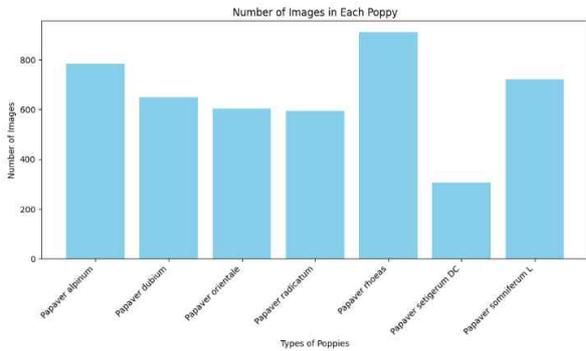


그림 2. 데이터셋의 이미지 수

그 결과 최종적으로 확보한 데이터셋의 수는 그림 2 와 같다. 본 연구에서는 양귀비 이미지를 분류하기 위하여 EfficientNet 모델로 전이 학습을 수행하였다[2]. EfficientNet 에는 다양한 버전이 있는데 이중 우리는 b3, b4, v2_s 를 선정하여 비교하였다.

표 1. 10epoch 학습 후 EfficientNet 버전 별 정확도

	Train Accuracy	Valid Accuracy	Params
b3	96.25	88.80	12.2M
b4	91.43	86.09	19.3M
v2_s	91.02	85.43	21.5M

각 버전별 10 epoch 학습 후 정확도는 표 1 과 같이 나왔다. 표 1 에서 제일 정확도가 높게 나온 b3 를 학습 모델로 선정하였고, b3 모델을 개선하여 더 높은 정확도로 분류할 수 있게 하였고 최종적으로 89.21%의 검증 정확도를 달성하였다.

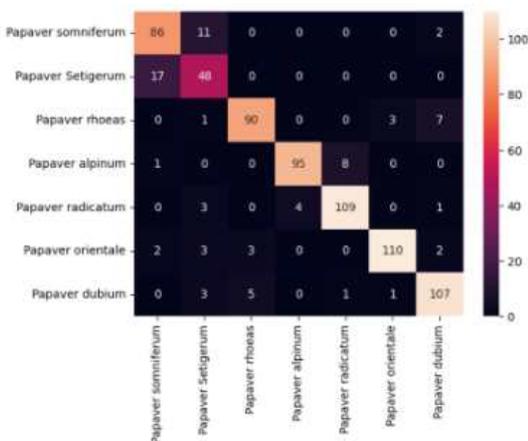


그림 3. 양귀비 분류 모델의 혼동행렬

그림 3 은 학습된 양귀비 분류 모델이 각 이미지를 어떤 레이블로 분류했는지를 나타낸 혼동 행렬이다. 그림 3 에서 대체적으로 모델이 입력을 잘 분류하였으나 양귀비 (Papaver somniferum)와 나도 양귀비 (Papaver setigerum)를 구분하는 과정에서 일부 혼동이 발생한 것으로 보인다. 그러나 이 두 종은 모두 마약성 양귀비로 연구의 목적인 마약성 양귀비의 구분에 두 종의 혼동은 영향을 주지 않는다고 판단하였다. 그 결과 본 연구에서 학습시킨 모델은 입력된 양귀비 사진이 어떤 종인지를 89.21%의 정확도로 분류할 수 있으며, 해당 양귀비가 마약용인지 관상용인지를 97.93%의 정확도로 분류할 수 있다.

III. 결론

본 논문에서는 GBIF 웹사이트에서 수집한 양귀비 데이터셋을 EfficientNet-b3 모델에 맞춰 전처리하고, 데이터 증강 기법으로 데이터 품질과 양을 향상시켰다. 학습 과정에서는 StepLR 학습률 스케줄러와 조기 종료 기법을 사용해 과적합을 방지했으며, 그리드 검색을 통해 최적의 하이퍼파라미터를 선정하였다. 그 결과 양귀비 종 분류에서는 89.21%, 마약용과 관상용 양귀비 분류에서는 97.93%의 높은 정확도를 달성했다.

참 고 문 헌

- [1] Web Page : Global Biodiversity Information Facility, (<https://www.gbif.org/>)
- [2] Mingxing Tan, Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", 2020 ([arXiv:1905.11946v5](https://arxiv.org/abs/1905.11946v5))