

사과 병해충 분류를 위한 CNN 기반 모델 비교

이수민, 이유현, 이은솔, 한세윤
서울여자대학교 정보보호학과

lsm971230@swu.ac.kr, dldbgus1234@swu.ac.kr, dmsthf1129@swu.ac.kr, spzh1020@swu.ac.kr

Comparison of CNN-based models for apple pest classification

Su-min Lee, Yu-hyeon Lee, Eun-sol Lee, Se-yun Han
Dept. of information security, Seoul Women's University

요 약

세계에서 가장 중요한 온대 과일 작물 중 하나인 사과의 생산성과 품질은 병해충 여부에 큰 영향을 받는다. 이를 진단하기 위해서는 효율적이고 많은 전문 지식과 상당한 시간이 필요하다. 그러므로 이를 해결하기 위해 효율적이고 정확하게 다양한 병해충을 진단하는 시스템이 필요하다. 본 논문에서는 이미지 분석에 큰 효율을 보인 딥러닝 기반 CNN 들을 비교 분석하여 사과의 병해충 여부를 판별하고 최적의 모델을 제시한다. 딥러닝 기반 CNN 구조를 가진 AlexNet, VGGNet, Inception-ResNet-v2, DenseNet 을 채택해 사과 병해충 분류 성능 평가를 진행했다. 그 결과 DenseNet 이 가장 우수한 성능을 보여주었다.

1. 서론

사과는 세계에서 가장 중요한 온대 과일 작물 중 하나이다. 병해충은 과수원의 전반적인 생산성과 품질에 큰 위협이 되는데 이를 예방하기 위한 질병 진단 과정은 수작업에 기반을 두고 있어 시간과 비용이 많이 든다. 최근 다양한 기계 학습 알고리즘을 이용해 식물 병해충 분류에 활용한 것이 성공적인 결과를 보여주었다. 이 중 딥러닝 기반의 CNN 아키텍처의 진화는 분류의 정확성을 크게 향상했다.

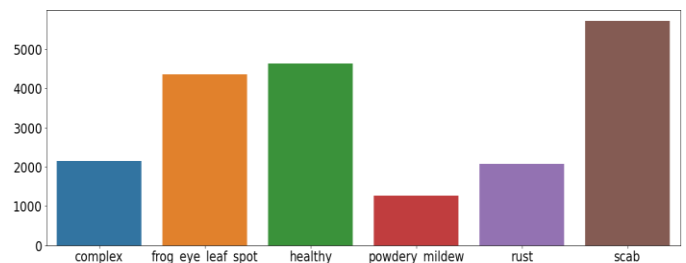
kaggle 에서 주최한 Plant Pathology 2021 - FGVC8[1]의 최종 목표는 머신 러닝 기반 모델을 개발하여 사과 병해충을 분류하는 것으로 WEI HAO KHOONG 의 Insect Augmentation et al.[2]가 0.898 점으로 가장 우수한 점수를 받았다. 그 외에도 우수한 성적을 거둔 모델이 다수 존재하였는데, 해당 모델들 중 다수가 작은 epoch 로 모델을 학습시켜 성능을 측정하였다. 이에 본 논문은 낮은 epoch 에서는 제대로 된 성능을 내지 못하는 모델이 있을 것으로 판단하였고 보다 높은 에포크를 학습시켜 각 모델들이 보이는 전체적인 성능을 비교하였을 때 어떠한 결과가 나오는지 분석하였다.

본 논문에서는 CNN 의 주요 아키텍처 중 AlexNet, VGGNet, Inception-ResNet-V2, DenseNet 을 채택하였다. 학습 모델의 선정 기준은 컬러 이미지 분류 시 기용 빈도가 높은가, 유의미한 성능 차이

로 기준점이 될 수 있는가였다.

2. 모델 학습

데이터 셋은 kaggle 의 Plant Pathology 2021 - FGVC8 에서 제공한 11 가지 사과 병해충 이미지와 건강한 사과 나뭇잎의 이미지를 사용하였다. 병해충의 종류별로 train 데이터와 test 데이터를 제공하며 데이터 셋의 세부 사항은 (그림 1)과 같다. 데이터 셋은 컬러 이미지로 구성되어 있으며 레이블은 'scab', 'frog_eye_leaf_spot', 'rust', 'powdery_mildew', 'complex', 'healthy' 가 존재하며, 'complex' 는 여러 질병이 섞인 경우가 포함되어 있다.



(그림 1) 데이터 셋의 질병 분포

여러 상황에 적용시키기 위하여 다른 배경의 이미지, 하루 중 다른 시간에 찍은 이미지, 초점이 잘 맞지 않는 일부 이미지, 복합적으로 병해충 질병을 가진 잎의 이미지를 포함하여 다양성을 향상시켰다.

또한 추가적으로 데이터를 확장하여 입력 이미지를 증가시켜 정확도를 높였으며 epoch 를 증가시켜 각 모델이 확장된 데이터 셋을 효율적으로 활용할 수 있도록 유도했다.

3. 심층 CNN 모델

이미지 분석 알고리즘(CNN, Convolutional Neural Network)이란 데이터에서 이미지를 직접 학습하고 패턴을 사용해 이미지 분류를 해주는 알고리즘으로 딥러닝의 신경망 중 하나로 DNN(Deep Neural Network)의 한 종류이다. CNN 은 입력 데이터에 대해서 필터를 통해 이미지의 특징을 추출, 강화하고 이미지의 크기를 축소하는 과정을 반복하여 처리된 결과를 출력해 준다.[3]

3-1 모델 소개

1. AlexNet[4]

AlexNet 은 2012 년 개최된 ILSVRC 대회에서 1위를 차지한 모델로 당시 뛰어난 성능으로 CNN 의 우수함을 전 세계적으로 알린 모델이다. 3 개의 fully-connected layer 와 15 개의 convolutional layer 로 이루어져 있으며, 위아래로 filter 를 절반씩 나누어 병렬연산을 수행하는 것이 가장 큰 특징이다. 또한, 학습 속도를 높이기 위해 비선형성 뉴런인 ReLU 함수를 사용했고 과적합을 방지하기 위해 Dropout 을 시킨다.

2. VGGNet[5]

VGGNet 은 2014 년 이미지 넷 이미지 인식 대회 ILSVRC 에서 준우승을 한 모델로 간단한 구조와 좋은 성능 덕분에 같은 대회에서 우승을 한 복잡한 형태의 GoogleNet 보다 더 인기를 얻었다. 또한, 네트워크의 깊이를 깊게 만든 모델의 성능이 어떻게 좋아지는지 확인하고자 진행한 연구로 만들어져 네트워크의 역사를 봤을 때 해당 모델부터 깊이가 깊어졌다. VGGNet 은 매우 간단한 구조를 가지면서 꽤 좋은 성능을 보이기 때문에 비교군으로 혹은 테스트를 할 때 자주 이용된다.

3. Inception-ResNet-v2[6]

Inception-ResNet 는 Inception network 에 Residual Connection 을 적용하여 Inception 의 계산 효율을 보장해 주면서 residual 의 이점을 모

두 취한 네트워크로 v1 버전과 v2 버전이 있다. Residual Connection 이란, 2015 년에 등장한 ResNet 에서 등장한 개념으로 output 에 input 이미지를 합하여 리턴 시키는 과정으로 학습 속도를 크게 향상시킨다. Inception v4 모델에 ResNet 장점을 흡수시켜 학습 속도가 빠른 모델로 Inception 신경망을 좀 더 효과적으로 깊고 넓게 만들기 위해 고안되었다.

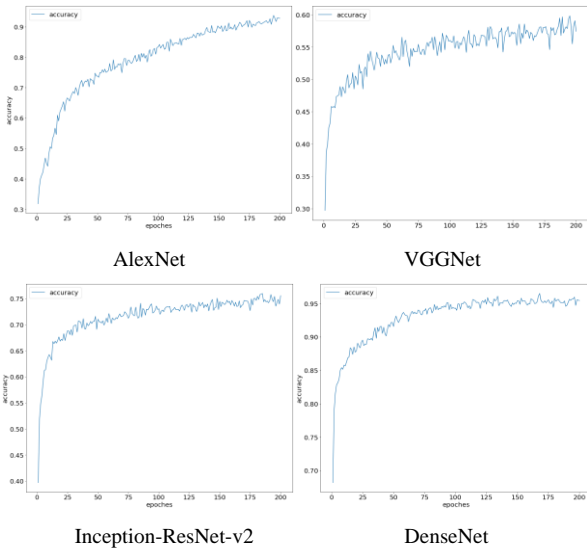
4. DenseNet[7]

DenseNet 은 이전 layer 의 feature map 을 다른 모든 layer 의 feature map 에 연결한다. 즉, 처음 layer 의 feature map 을 마지막 layer 까지 연결하여 정보를 잃는 것을 막는다. 해당 모델에서 각 layer 를 연결할 때는 skip connection 이 이루어지는데, concatenation 을 통해 이루어진다. 따라서 feature map 의 크기가 동일해야만 연결을 수행할 수 있다. 정보 손실을 막을 수 있다는 점 이외의 DenseNet 의 장점은 적은 파라미터 수를 이용한다는 점이다. 각 layer 의 feature map 을 연결하여 다음 layer 로 전달을 하면, 적은 채널 수의 feature map 을 생성한다. 그리고 각 feature map 과 합쳐서 다음 layer 로 전달을 하게 된다.

4. 실험 결과

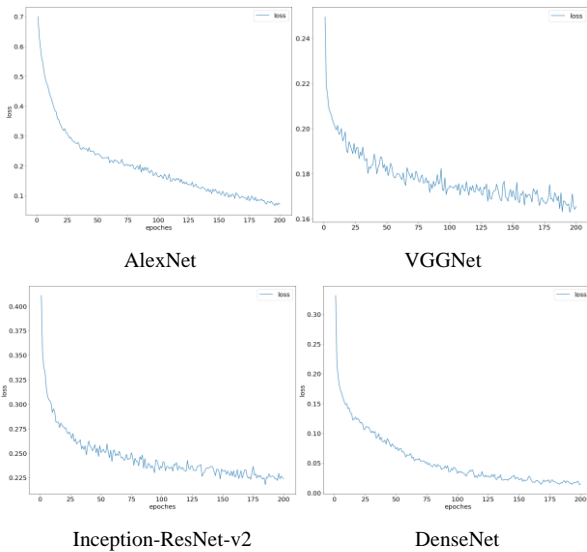
AlexNet, VGGNet, Inception-ResNet-v2, DenseNet 을 같은 학습 환경 설정과 데이터 셋을 이용, epoch 를 200 으로 설정하여 실험했다. 이후 각각의 학습 속도, 정확도(accuracy), loss 값, 정밀도(precision), 재현율(recall)을 성능 비교의 지표로 삼았다. 본 논문에서는 학습 속도는 1 epoch 당 학습에 걸리는 시간, 정확도는 True 값은 True 으로 False 값은 False 로 예측한 것의 비율, loss 값은 실제 정답과 모델이 예측한 값 사이의 차이, 정밀도는 모델이 True 라고 분류한 것 중 실제 True 값을 가진 것의 비율, 재현율은 실제 True 값인 것 중 모델이 True 라고 예측을 한 것의 비율로 정의한다.

1 epoch 당 학습에 걸리는 속도는 AlexNet 이 34 초, VGGNet 이 56 초, Inception-ResNet-v2 이 58 초, DenseNet 이 70 초로 AlexNet 이 다른 모델들보다 약 2 배 빠른 속도를 보였다.



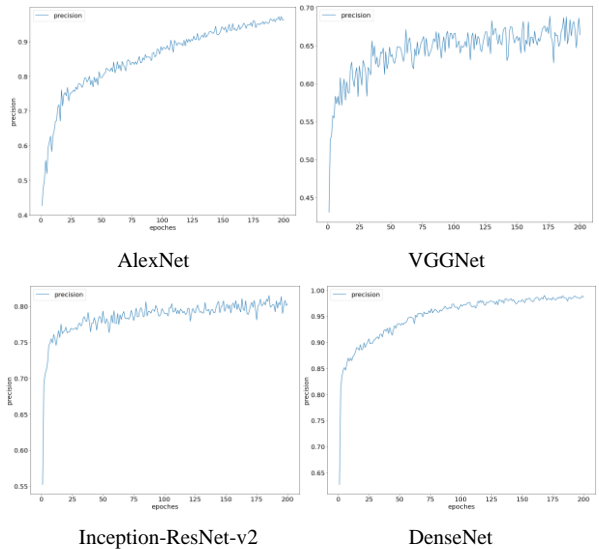
(그림 2) epoch 200 일 경우 각 모델의 정확도

(그림 2)은 epoch 에 따른 각 CNN 모델의 학습 정확도를 나타내는 그래프이다. DenseNet, AlexNet, Inception-ResNet-v2, VGGNet 순으로 높은 정확도를 보였다. 최대 정확도는 DenseNet 은 약 96%, AlexNet 은 약 92% 로 나타났으며 Inception-ResNet-v2 는 약 76%, VGGNet 은 약 59%로 비교적 낮은 정확도를 보였다.

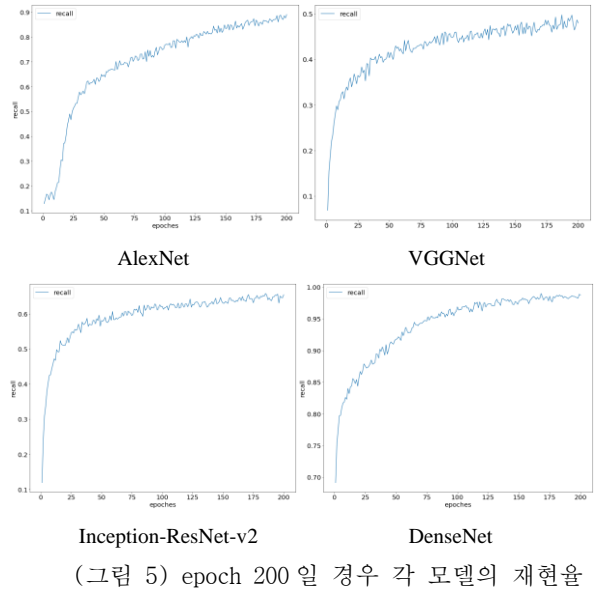


(그림 3) epoch 200 일 경우 각 모델의 loss 값

(그림 3)는 epoch 에 따른 각 CNN 모델의 loss 값을 나타내는 그래프이다. DenseNet, AlexNet, VGGNet, Inception-ResNet-v2 순으로 낮은 loss 값을 보였다. DenseNet 이 높은 정확도와 낮은 loss 값을 가지는 이유는 Dense block 과 Transition block 로 이루어져 있는 DenseNet 의 모든 layer 가 전부 연결되어 있어 정보 상실을 막을 수 있다는 Dense block 의 특성 때문으로 보인다.

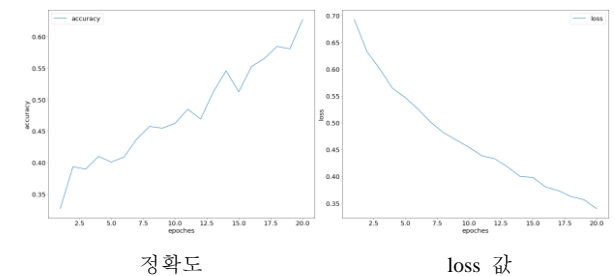


(그림 4) epoch 200 일 경우 각 모델의 정밀도



(그림 5) epoch 200 일 경우 각 모델의 재현율

(그림 4)과 (그림 5)는 epoch 에 따른 각 CNN 모델의 정밀도와 재현율을 나타내는 그래프이다. 이는 정확도와 정비례하여 (그림 2)과 유사한 양상을 보였다.



(그림 6) epoch 20 일 경우 AlexNet 의 정확도, loss 값

(그림 6)에서 알 수 있듯이, epoch 를 20 으로 낮출 경우, AlexNet 을 제외한 다른 세 모델은 이전 실험 결

과와 큰 차이를 보이지 않았다. 그러나, AlexNet 은 20 epoch 에서 유독 낮은 정확도와 높은 loss 값을 보여 epoch 수에 따라 확연하게 다른 결과를 보였다. AlexNet 의 성능은 epoch 값이 커질수록 개선되었는데, 이것은 AlexNet 을 활용하는 경우 어느 정도 이상의 epoch 가 필요하다는 것을 보여준다.

5. 결론

본 논문에서는 CNN 구조를 가진 AlexNet, VGGNet, Inception-ResNet-v2, DenseNet 의 사과 병해충 분류 성능을 다양한 epoch 에 따라 측정하였다. 비교 결과 DenseNet 이 정확도, 손실도, 정밀도, 재현율 모든 측면에서 가장 우수한 성능을 보여줬으며 epoch 값에 상관없이 일정하게 좋은 결과가 나타났다. 학습 속도 측면에서는 AlexNet 이 가장 뛰어났다. 높은 정확도를 요구하는 경우 DenseNet 을 사용하는 것이 가장 효율적인 선택이지만, epoch 값이 크고 빠른 학습 속도가 필요할 경우 AlexNet 을 사용하는 것 역시 사과 병해충 분류에 유용한 선택이 될 것이다.

참고문헌

- [1] kaggle Plant Pathology 2021 - FGVC8, <https://www.kaggle.com/c/plant-pathology-2021-fgvc8>
- [2] Insect Augmentation et al, <https://www.kaggle.com/code/khoongweihao/insect-augmentation-et-al>
- [3] Y. LeCun et al. "Backpropagation applied to handwritten zip code recognition". Neural computation v.1 no.4, 1989. pp.541 - 551
- [4] A. Krizhevsky, I. Sutskever and G. E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks", 2012.
- [5] Karen Simonyan, Andrew Zisserman (Visual Geometry Group, Department of Engineering Science, University of Oxford), "Very Deep Convolutional Networks for Large-Scale Image Recognition", International Conference on Learning Representations 2015, San Diego, CA, May 2015.
- [6] Szegedy, C., Ioffe, S., Vanhoucke, V. and Alemi, A. A. "Inception-v4, inception-resnet and the impact of residual connections on learning." In Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, California USA. 2017.
- [7] Gao Huang et al., "Densely connected convolutional networks.", Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, Honolulu, 2017., pp. 4700-4708