

이미지 증강에 대한 고찰

Discussion On the Image Data Augmentation

이지민 (Jimin Lee, leejeemin0608@gmail.com) 세종대학교 컴퓨터공학과

방식이다.

I. 머리말

데이터 증강(Data Augmentation)이란 이미지에 다양한 변화를 주면서 기존의 이미지와 같은 특성을 가지는 이미지를 생성하는 방법이다. 주로 많은 양의 데이터를 필요로 하는 딥러닝 시 부족한 훈련 이미지의 수를 늘리기 위한 해결 방법으로 쓰인다. 이미지의 분류 및 인식을 진행할 때, 충분하지 않은 이미지 데이터의 수로 인해 과적합이 일어날 확률이 높기 때문이다.

과거에 주로 쓰인 데이터 증강 방식은 주로 전통적인 데이터 증강 기법으로 불린다. 이에 해당하는 예로는 이미지 회전, 블러, 반전 등의 방식으로 원본 이미지에 물리적인 변환을 가하는 방식이다. 하지만 이런 방식을 통해 증강된 이미지 셋은 주는 다양성이 적다고 한다[1]. 이런 한계를 개선하기 위해 GAN 등의 알고리즘을 이용해 기존 이미지를 학습하여 그와 비슷한 이미지를 생성하는 방식이 많이 쓰이는 추세이다.

II. 이미지 데이터 증강 기법 동향

이미지 데이터 증강은 크게 회전, 픽셀 단위 변환, Random Erase, 좌우 반전과 같은 전통적인 방법과 GAN과 같은 생성 알고리즘을 이용한 증강 방법으로 구분할 수 있다.

Pixel-Level Transforms는 이미지를 픽셀 단위로 이미지를 변환시키는 방법으로 이미지를 흐리게 만드는 Blur, 노이즈를 주는 Noise, 색을 변환시키는 RGBShift 등 이미지에 적용시키는 기법이다. 흔히 알려진 가우시안 블러, Contrast Jitter 등이 이에 해당한다[2]. 그림 1의 앵무새 그림과 같이 이미지 자체의 색이나 외면에 직접 변환을 가하는



그림 1. Pixel-Level Transforms [2]

연구 [3,4]에서는 회전, 밝기 조절, 이미지 이동 등의 기법으로 생성된 이미지를 사용해 원형 반점 바이러스 등의 질병을 분류하는 모델을 제안했다. 실험 결과, 증강된 데이터를 사용한 모델의 정확도는 약 13.07%까지 상승했다. 이러한 결과로 보아, 증강된 데이터가 포함된 훈련 셋이 증강되지 않은 원본 데이터보다 딥러닝 훈련 데이터로서 더 효과적이라는 결론을 도출할 수 있었다.

연구 [5]에서는 2장의 이미지 A, B를 훈련 셋에서 랜덤하게 추출한다. 이렇게 얻게 된 2장의 이미지를 평균을 구해내서 Mixed Patch를 생성한다. 이 때 라벨은 이미지 A의 라벨로 사용한다. 이렇게 만들어진 이미지는 A, B의 특징이 섞여있다. 이런 방식의 데이터 증강은 이미지 분류 시 굉장히 높은 정확도를 이끌어냈다고 한다. 이 연구에서 더 발전시켜 단순 2가지 이미지를 평균내는 방식이 아닌

8종류의 이미지를 섞는 방식을 제안한 연구 [6]가 있다. 그림 2는 이에 대한 그림으로 두 가지의 이미지를 입력으로 받고 Cutmix, 모자이크 등등의 기법을 이용해 8가지의 이미지를 생성한 뒤 이 8가지 이미지를 섞어서 평균을 내는 방식을 제안했다. 이 방식 또한 실험에서 높은 성능을 도출했다고 한다.

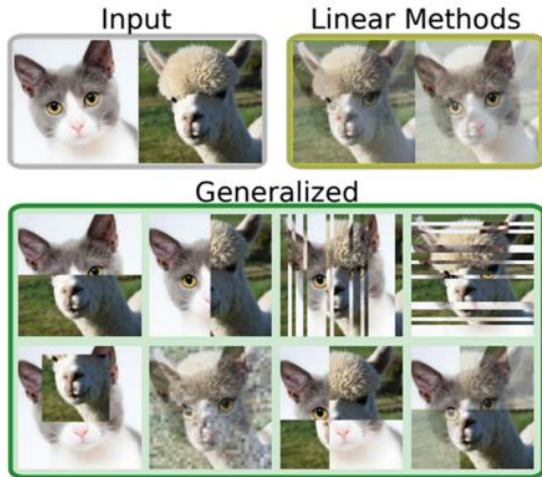


그림 2. Improved Mixed-Example Data Augmentation

연구 [7]은 이미지에서 랜덤 크롭한 이미지를 합쳐서 한 장으로 만드는 Random Image Cropping and Patching을 제안했다. 또한 이미지 4장의 라벨 4개를 합쳐진 이미지에서 차지하는 면적의 비율에 따라 섞어 소프트 라벨을 만들었다. 그림 3이 그 예시로 4장의 이미지에서 이미지의 일부를 합쳐 1장의 이미지를 도출한다. 하지만 이 방식에도 단점이 존재한다. 예로 들어, 배의 그림에서 배가 제외된 채 크롭된 이미지가 새롭게 만들어지는 이미지에 포함됐을 때, 그 이미지에는 배에 해당하는 물체가 없지만 라벨은 배가 부여될 수도 있다는 점이다.

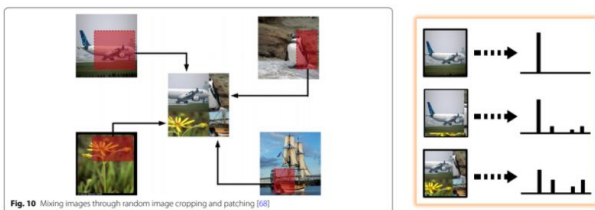


그림 3. Random Image Cropping and Patching

연구 [8]은 Random Erase 방식을 제안했다. 기존의 섞거나 물리적인 변환을 주던 방식과는 다르게 이미지에서 랜덤하게 몇몇 부분을 지우는 방식이다. 이와 비슷한 방식은 연구 [9]로, 이미지에 랜덤한

바운딩 박스를 그린 뒤, 박스 내의 값을 0으로 채우는 방식을 제안했다. 이런 방식을 Cutout이라 한다. 그림 4가 그 예시로 이미지에 회색 바운딩 박스가 있는 것을 볼 수 있다.



그림 4. Cutout 예시

지금까지의 연구가 입력된 이미지의 일부를 지우거나, 회전하는 방식을 통해 증강하는 방식이었다면 연구 [10]에서는 학습된 Image-to-Image 네트워크의 가중치와 활성화를 변화시키는 방식으로 증강 기법을 제안했다. Deep Augment를 사용하면 기존의 기법들로 구현하기 어려운 다양한 이미지를 생성할 수 있다. 하지만 회전, 밝기 조절과 같은 전통적인 데이터 증강 방법은 데이터의 변화와 다양성이 떨어진다는 단점이 있다[11,12].

연구 [11]에서는 AR-GAN(Attentive Recurrent Generative Adversarial Network)을 사용해 토마토 식물의 Gray Mold, Leaf Mold 등의 병해충 이미지를 증강했다. 증강 후 분류 학습을 시행한 결과, 비증강 데이터나 전통적인 증강 기법을 사용한 모델보다 AR-GAN을 이용해 증강한 모델의 분류 정밀도는 평균 약 0.05%가 향상됐다.

AR-GAN은 그림 4와 같이 데이터의 합성과 클래스 밸런싱을 위해 AR-GAN을 사용하는 합성 데이터 증강 블록과 이미지 인식을 위해 딥 컨벌루션 신경망을 사용하는 인식 시스템의 두 블록으로 구성된다.

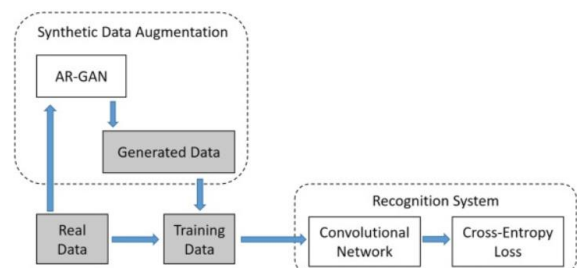


그림 4. AR-GAN

그림 5는 AR-GAN의 구조도이다. A, B는 2개의 도메인으로 2개의 Generator는 한 도메인에서 다른 도메인으로 이미지를 변환한다. 각 도메인의 Discriminator는 이미지가 해당 도메인에 속하는지 여부를 판단한다. 데이터는 $A \rightarrow B \rightarrow A$ (녹색), $B \rightarrow A \rightarrow B$ (빨간색)로 흐른다. AR-GAN은 주기 일관성을 유지하기 위해 특징 추출 네트워크로 구성된 활성화 재구성 모듈을 도입했다.

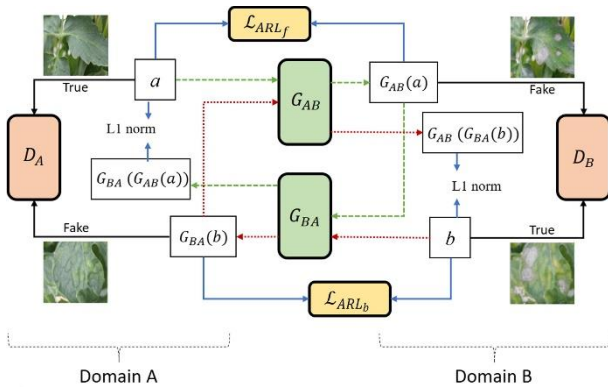


그림5. AR-GAN의 구조도

연구 [13]에서는 토마토 식물의 병해충 이미지를 증강시키기 위한 모델로 C-GAN(Conditional GAN)을 선택했다. 사용된 데이터 셋은 PlantVillage[14]에서 제공된 16,012장의 토마토 식물 잎 이미지로 C-GAN을 통해 증강시켜 4,000장의 이미지를 생성했다. 증강된 이미지가 포함된 훈련 데이터 셋으로 학습한 결과 정확도가 평균 2.57% 상승했다. 연구 결과로 미루어 보아, GAN 알고리즘을 사용하여 훈련 데이터를 증강시키는 방법이 전통적인 데이터 증강 방법에 비해 병해충 진단에 효과적이라는 결론을 도출했다.

C-GAN은 기존의 GAN에 몇 가지의 데이터를 입력하여 성능을 향상시킨 모델이다. 기존의 GAN은 Generator에 노이즈, Discriminator에 이미지를 입력했다. C-GAN의 경우 그림 6처럼 라벨을 가지는 노이즈와 이미지를 이용한다. 라벨이라는 추가적인 정보를 이용해 성능을 향상시키는 방식이다[15].

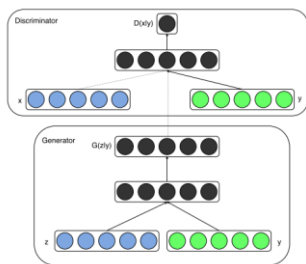
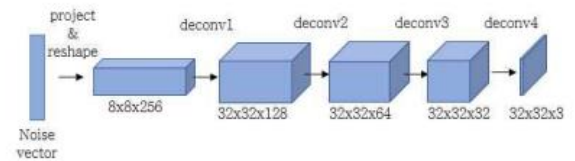


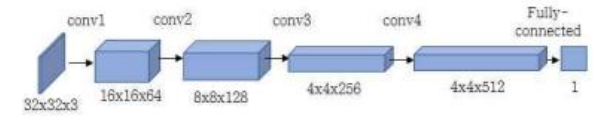
그림 6. C-GAN

연구 [16]은 DC-GAN(Deep Convolutional Generative Adversarial Networks)을 이용하여 토마토 질병 분류를 진행하였다. DC-GAN은 기존의 GAN에 이미지 특징을 잘 추출할 수 있는 CNN 구조를 더함으로써 원본과 비슷한 이미지를 생성할 수 있다는 특징이 있다. 이를 이용해 DC-GAN 모델에 토마토 질병 이미지를 학습시키고 생성했다.

DC-GAN은 유사한 이미지를 생성하는 Generator와 생성된 이미지가 진짜 이미지와 유사한지 판별하는 Discriminator 부분으로 구성된다. 이는 기존의 GAN과 유사한 구조이나 DC-GAN은 기존 GAN의 신경망을 CNN으로 대체한 구조이다. Generator는 그림 7의 (a)와 같으며 Fully-Connected Layer와 Deconvolution Layer로 구성된다. Discriminator는 그림 (b)와 같고 4개의 Convolution Layer와 스트라이드의 크기를 2로 설정함으로써 차원 축소가 진행된다.



(a)



(b)

그림 7. (a) DCGAN Generator

(b) DCGAN Discriminator

생성된 4가지의 토마토 질병 이미지 분류기는 Precision, Recall 등을 통해 평가되었고 모든 평가 지표에서 성능 향상을 보였다.

III. 결론

현재의 이미지 데이터 증강 기술은 전통적인 데이터 증강 방법과 생성 알고리즘을 이용한 증강 방법으로 나눌 수 있다. 전통적인 데이터 증강 방법은 다양성이 떨어진다고 알려져 있지만 단순히 물리적 변환을 주는 것이 아닌, 여러가지 증강 방법을 혼합하여 다양성을 높일 수도 있다. GAN을 이용한

이미지 생성 기법은 이미지 증강 시 다양성을 줄 수 있다는 장점이 있지만, 원본 이미지를 학습하여 유사 이미지를 생성하는 방법이기엔 오히려 부족한 방법으로 학습하면 비슷하지 않은 이미지가 생성될 수도 있다는 단점이 있다. 이런 두 가지 카테고리의 증강 방식을 이용해 각자의 단점을 보완할 수 있는 연구가 진행되어야 한다.

참고문헌

- [1]Haseeb Nazki, Unsupervised Image Translation using Adversarial Networks for Improved Plant Disease Recognition, Computers and Electronics in Agriculture, 2019
- [2]<https://hoya012.github.io/blog/Image-Data-Augmentation-Overview/>
- [3]Prem Enkvetchakul, Effective Data Augmentation and Training Techniques for Improving Deep Learning in Plant Leaf Disease Recognition, Applied Science and Engineering Progress 10.14416, June 2022
- [4]J Arun Pandian, Data Augmentation on Plant Leaf Disease Image Dataset Using Image Manipulation and Deep Learning Techniques, 2019 IEEE 9th International Conference on Advanced Computing (IACC), December 2019
- [5]Hiroshi Inoue, Data Augmentation by Pairing Samples for Images Classification, ICLR 2018, 2018
- [6]Cecilia Summers, Michael J. Dinneen, Improved Mixed-Example Data Augmentation, 2019 IEEE Winter Conference on Applications of Computer Vision, 2018
- [7]Ryo Takahashi, Takashi Matsubara, Kuniaki Uehara, Data Augmentation using Random Image Cropping and Patching for Deep CNNs, IEEE Transactions on Circuits and Systems for Video Technology Volume: 30 Issue: 9, 2019
- [8]Zhun Zhong, Liang Zheng, Guoliang Kang, Random Erasing Data Augmentation, The Thirty-Fourth AAAI Conference on Artificial Intelligence, 2017
- [9] Terrance DeVries, Graham W. Taylor, Improved Regularization of Convolutional Neural Networks with Cutout, 2017
- [10]Dan Hendrycks, Steven Basart, Norman Mu, The Many Faces of Robustness: A Critical Analysis of Out-of-Distribution Generalization, CVPR 2020, 2020
- [11]HaseebNazki, Unsupervised image translation using adversarial networks for improved plant disease recognition, Computers and Electronics in Agriculture, Volume 168 105117, January 2020
- [12]Amy Zhao, Data augmentation using learned transformations for one-shot medical image segmentation, Computer Vision and Pattern Recognition (cs.CV) 9 pages, February 2019
- [13]Amreen Abbas, Tomato plant disease detection using transfer learning with C-GAN synthetic images, Computers and Electronics in Agriculture Volume 187, August 2021
- [14]PlantVillage, <https://plantvillage.psu.edu/>
- [15]Mehdi Mirza, Simon Osindero, Conditional Generative Adversarial Nets, 2014
- [16]Ji-yeon Park, Hyeon-ji Kim1, Kyungbaek Kim, Assessing Impact of DCGAN Image Data Augmentation for CNN based Tomato Disease Classification, Journal of Digital Contents Society Vol. 21, No. 5, pp. 959-967, May 2020