작은 dataset에 대한 효율적인 CNN 학습방법 연구

*나성원 **배효철 ***윤경로

건국대학교 컴퓨터공학과

*securityin4@naver.com, **th1g@nate.com, ***yoonk@konkuk.ac.kr

A study of efficient learning methods of CNN for small dataset

*Seong-Won, Na **Hyo-Churl Bae ***Kyoungro, Yoon

Department of computer science and engineering, Konkuk University

요약

최근 이미지 처리 및 인식 문제를 해결하는데 많이 사용되고 있는 CNN(Convolution Neural Network)를 이용하여 작은 dataset에서 Overfitting을 감소시키며 학습 할 수 있는 방법인 Dropout과 이미지를 왜곡하여 data를 늘리는 방법을 사용하여 보다 효율적으로 학습할 수 있는 방법을 연구 하였다. Batch별 처리속도를 기준으로 두 네트워크의 구조를 다르게 구현하여 비슷한 처리 시간을 수행하게 되도록 실험환경을 만들고 진행 하였다. Tensorflow로 네트워크를 구성하였고. Dataset은 Cifar_10을 사용 한다. 실험결과에 의하면 dropout의 경우 더 빨리 정확도가 향상되지만 이미지 왜곡을 사용하는 경우 저 높은 정확도로 수렴하였다.

1. 서론

최근 이미지 인식 분야에서는 Deep Learning의 한 방법인 CNN(Convolution Neural Network)[1]이 정확도 측면에서 가장 인상적인 성능을 보여주고 있으며, 이에 관련된 많은 연구가 활발하게 진행되고 있다. 이미지 인식 대회 중 하나인 ILSVRC(ImageNet Large-Scale Visual Recognition Challenge)[1]대회에서 기존의 feature기반의 Classifier보다 성능이 뛰어난 Deep Convolution Neural Network가 우승을 차지하면서 이후 매 대회 Convolution기반의 Network들이 선보이고 있으며, 점차 정확성도 높아지고 있다. 하지만 DNN(Deep Neural Network) 모델을 학습시키기 위해서는 충분한 양의 data가 요구 되며 이는 ILSVRC와 같이 Dataset을 제공해주는 경우도 있지만 사용자의 관심 분야가 다를 경우 직접 Database를 준비해야 한다. 이와 같은 경우 직접 이미지를 촬영하거나 인터넷 검색을 통해 수집해야 하기 때문에 많은 시간이 소비 될 뿐 아니라 다양하고 많은 data를 수집하기 힘들다.

또한, CNN 모델의 학습 시 충분한 data가 제공 되지 않으면 정확도를 높이는 것이 힘들 뿐 아니라 Overfitting 문제가 발생 할 수 있기 때문에 이를 해결 할 수 있는 이미지 왜곡과 Dropout을 사용하여 작은 dataset에서 좀 더 효율적으로 학습 할 수 있는 방법을 실험하고 효과를 확인 하고자 한다. 논문의 구성은 다음과 같다. 먼저 Overfitting과 사용 된 Dataset에 대해 간단하게 설명하고, 그 다음 각각의 네트워크 구성에 대해 설명한다. 세 번째로 실험 환경 및 결과를 분석 하고, 마지막으로, 결론을 논의 한다.

2. 본론

2.1. Dataset

이 논문에서 실험에 사용한 Cifar_10은 자동차, 오토바이, 비행기 등과 같은 10개의 클래스를 갖는 Dataset으로 크기는 32x32 Color image로 클래스 당 6000개의 이미지로 구성 되어 있다. Training에 사용 되는 data는 50000개, Test에 사용하는 data는 10000개로 나뉘어져 있다.

2.2. Overfitting

Overfitting은 DNN(Deep Neural Network)에서 Training data에 대해 과도하게 학습 되는 것을 뜻한다. 즉, Training dataset에 대해서는 정확도가 점점 증가하게 되지만 실제 일반 data에서는 정확도가 학습 시 보다 감소하는 것을 말한다. 이유는 Training 시 data의 수가 적어 비슷한 data를 반복적으로 학습하고 적응하게 되어 사소한 특징까지 학습을 했기 때문이다. 이러한 문제를 해결하기 위해서는 대표적으로 data를 증가 시켜 다양한 입력 값을 주는 것과 Dropout[2]을 적용하여 완화하는 방법이 있다. 본 논문에서는 data부족으로 인해 발생하는 Overfitting 문제에 대해 이미지를 왜곡하여 dataset를 늘리는 방법과 Network unit들을 무작위로 제거하여 입력 값의 변화를 주는 Dropout 방법의 실험을 통해 작은 Dataset에서 좀 더 효율적으로 Training 할 수 있는 방법을 확인 하고자 한다.

2.3. CNN 구성

배치별 처리 속도를 기준으로 Dropout과 이미지 왜곡 방법의 네트워크를 각각 구성하였다.

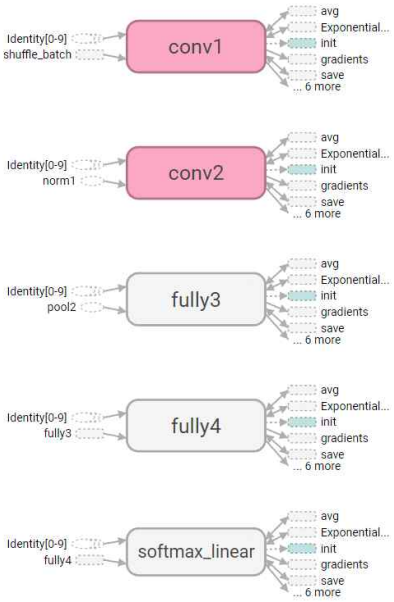

그림 1. 이미지 왜곡 CNN 구성도

이 논문에서 배치는 일괄 처리하는 이미지의 수를 나타내며 실험에서 배치 크기는 128로 설정 하였다 또한 32x32의 원본 이미지를 무작위로 24x24 크기로 잘라내어 사용 한다. 위 구성을 살펴보면 이미지 왜곡을 사용하는 네트워크의 첫 번째와 두 번째 레이어는 Convolution과 Rectified linear units(Relu)[3] 활성화 함수를 사용하였고, filter는 5x5, stride 1으로 설정, 개수는 64개로 구성 되었다. Convolution층 후에는 MaxPooling 레이어가 뒤따르게 되는 구조이며 3x3, stride 2의 filter를 사용한다. 3, 4번째 층은 Relu 활성화 함수를 포함한 완전 연결 레이어로 구성 되었으며, 각각 512, 192개의 크기를 갖는 결과를 갖게 되고, 마지막 층은 SoftMax함수를 통해 분류를 진행하게 된다.

이미지 왜곡 시 좌, 우 flip, brightness, contrast와 같은 image processing 시간이 포함되며 배치 당 약 0.2 초가 소요된다. Dropout에서는 Convolution 레이어 하나를 추가 하였고, filter의 수를 각각 96, 128, 256으로 설정 하였다. 또한 Dropout CNN은 input으로부터 Softmax함수 전까지 drop rate를 0.9로 설정 하였다.

3. 실험 및 결과

표 1. 구현 시스템 환경

CPU	Intel i7 Quad Core(3.4GHz)
GPU	Geforce GTX 980Ti
Language	Python, Tensorflow
Tool	Pycharm

위 표는 실험을 위해 구현한 시스템 환경을 나타내며 반복적인 연산에서 큰 성능향상을 볼 수 있는 GPU를 사용하여 학습 하였다.

그림 2에서 X축은 step수를 나타내는데 여기서 step은 배치 실행 횟수를 나타낸다. 두 가지 방법의 스텝 당 처리속도는 0.2초로 거의 비슷하게 설정되었다. Dropout의 경우 정확도가 86.2%에 수렴 하였으며, image 왜곡의 경우는 86.5%에서 수렴 하였다.

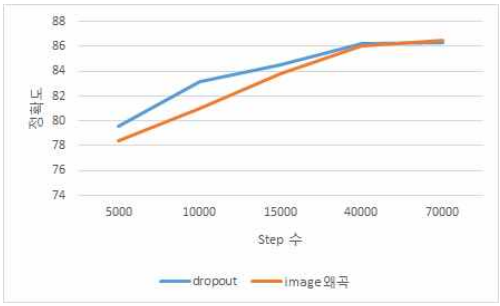

그림 2. step 수에 따른 Dropout과 Image왜곡의 정확도 변화(Dropout rate 0.9)

4. 결론

위 실험 결과를 통해 Dropout의 경우 적은 step만으로도 빠르게 정확도가 증가하는 것을 볼 수 있었다. 하지만 Dropout의 경우 32x32 이미지를 24x24로 잘라내게 되면 image size자체가 작아 drop rate를 변경하더라도 83%정도의 정확도에 수렴하는 것을 알게 되었고, 0.7이하로 설정 할 경우 정확도가 크게 떨어지는 것을 알게 되었다. 다양한 크기로 실험한 결과 26x26이 가장 좋은 성능을 보였다. 크기가 작은 이미지의 학습 시에는 Drop rate를 잘 선택해야 image의 객체를 손상시키지 않을 수 있다고 판단된다. 또한 이미지 왜곡방법의 CNN구조에서 단순히 Dropout만 추가하였을 경우 image size자체가 작아 아무리 입력 값을 무작위로 탈락 시키더라도 반복이 거듭 될수록 data에 적응하여 수렴하게 되었다. 이전 실험을 토대로 배치별 처리속도를 비슷하게 맞추는 수준내에서 Convolution 레이어와 filter를 추가하여 더 좋은 성능을 달성할 수 있었다. 이는 더 다양한 weight에 대해 학습을 할 수 있게 되어 더 나은 결과를 가져온 것으로 판단된다. 반면 이미지 왜곡의 경우 Dropout 보다 천천히 정확도가 증가하였지만 Dropout에 비해 더 높은 정확도를 달성하였다. 이는 Image processing을 통해 학습을 위한 다양한 데이터를 제공하는 것이 성능향상에 중요한 요소가 된다는 것을 의미한다고 판단된다.

감사의 글

이 논문은 2017년 정부(미래창조과학부)의 재원으로 정보통신기술진흥센터의 지원을 받아 수행된 연구임 (UHD 방송콘텐츠 기반 지능형 Dynamic Media 생성, 분배 및 소비 기술 개발)

참고문헌

[1] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

[2] Srivastava, Nitish, et al. "Dropout: A simple way to prevent neural networks from overfitting." The Journal of Machine Learning Research 15.1 (2014): 1929-1958.

[3] Nair, Vinod, and Geoffrey E. Hinton. "Rectified linear units improve restricted boltzmann machines." Proceedings of the 27th international conference on machine learning (ICML-10). 2010.