

이미지 분류 네트워크에서의 효율적 훈련 기법에 대한 연구

배은지, 이성진

동서울대학교

ejbae25@du.ac.kr, sungjinlee@du.ac.kr

A Study on the efficient training methodology in the Image Classification Network

Eunjee Bae, Sungjin Lee

Dong Seoul University

요약

최근 영상인식 기술의 성능적 발전은 수많은 데이터 축적과 딥러닝 네트워크의 심층화에 기인하여 왔다. 하지만, 이런 다양한 데이터들을 딥러닝 네트워크에 훈련시키는 것은 다양한 문제들을 유발시킨다. 적은 데이터 량에서 기인하는 오버피팅, 클래스 간의 데이터 양 차이에서 오는 클래스 불균형 (Imbalance), 멀티클래스 훈련 문제 등이 그것이며 본 논문은 Pascal VOC 데이터에서 이런 문제들을 발견하고 분석하였으며 해결책을 제시하고 실험으로 성능을 분석해 보았다.

I. 서론

최근 들어, IT산업뿐 아니라 대부분의 산업 분야에서 딥러닝 기반 영상인식 기술이 주목받고 있으며, 이러한 관심은 ImageNet, Open Images Dataset, MS COCO 와 같은 수많은 데이터 셋 구축 [1][2][3] 과 딥러닝 네트워크 심층화로 이어졌다 [4][5][6].

하지만 실제로 이런 딥 뉴럴 네트워크를 커스텀 데이터 셋에서 훈련시켜 문헌에서 제공된 성능만큼을 얻어내기란 쉽지 않다. 보통 문헌에서 제공되는 성능은 매우 큰 데이터 셋에서 다양한 클래스를 분류해 내는 챌린지(예, ILSVRC (ImageNet Large Scale Visual Recognition Competition))에 최적화된 네트워크를 사용하지만, 실제 산업에서 사용되는 데이터 셋과 분류 클래스는 그에 비해 매우 작기 때문에 적합한 모델 선택에 실패할 시, 데이터 과적합 (Overfitting) 혹은 데이터 언더피팅 (Underfitting) 문제 등의 문제들이 발생하기 때문이다.

본 논문에서는 객체 분류 (Classification) 와 검출 (Detection) 성능 분석

에 많이 사용된 PASCAL VOC 데이터 셋을 이용하여 VGG16 네트워크로 학습하였을 때의 결과를 통해 이 학습의 문제들에 대한 해결책들을 제시하고, 효과를 분석하였다. Pascal VOC 데이터 셋은 ImageNet 데이터셋에 비해 클래스 별 이미지 양이 적고 클래스 간 이미지 양의 차이가 크며 한 이미지 내에 여러 클래스가 포함되어 있는 경우가 많기 때문에 훈련시키기 까다로운 편이다.

본 논문은 이런 Pascal VOC 데이터 셋에 대해 분석하고 해당 데이터셋을 훈련시키는데 발생하는 어려운 점들, 오버피팅, 클래스 불균형, 멀티 클래스를 분석하였다. 그 후 각 문제점들을 위한 해결책들을 제시하고 이를 실험적으로 검증하였다.

II. 데이터 분석과 훈련기법

실험을 위해 사람, 동물, 탈것, 실내와 관련된 20개의 클래스를 가지고 있는 PASCAL VOC 데이터 셋을 사용하였다. 그러나 이 데이터 셋을 그대로

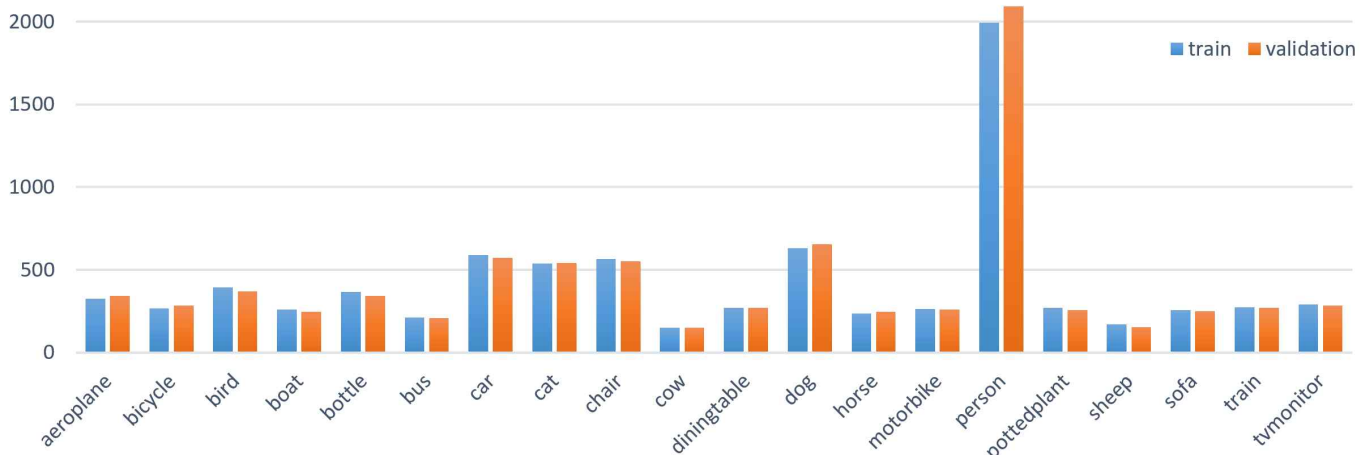


그림 1 Pascal VOC 데이터 셋의 클래스 구성

이용하여 학습하기에 여러 어려운 점들이 있다.

우리는 이러한 어려운 점들을 분석하고 해결하기 위해 데이터 전처리 과정에서 multi class 데이터를 제거해 보고, 부족한 훈련 데이터를 보충하여 실험해 보았으며, 클래스 간 데이터 비율을 평탄화하여 학습하였을 때의 정확도를 비교해보았고, 원 데이터 셋에서 가장 데이터가 많은 4개의 class를 선정해 학습하였을 때의 정확도도 비교해보았다.

1. Overfitting

그림 1에서 보듯이 Pascal VOC의 데이터 셋은 특정 클래스의 경우 이미지의 개수가 매우 적은 것을 확인할 수 있다. 이런 데이터 부족은 Overfitting 문제를 유발시켜 train 과 validation accuracy 의 차이를 발생 시킨다. 이 문제를 해결하기 위해 해당 클래스의 이미지가 최소 250여 장이 되도록 추가적인 데이터 약 2500여장 보충을 수행하였다. 또한 해당 데이터 들은 기존 이미지와는 겹치지 않게 새로운 데이터로 크롤링하여 보충하였다.

2. Class Imbalance Problem

훈련 시 클래스 간 이미지 개수를 유사하게 맞추는 클래스 균등화 (Class balancing) 는 인식 시에 클래스 간 편향성을 줄여줄 수 있는 전처리 단계이다. 그림 1에서 보듯이, Pascal VOC 데이터 셋은 클래스 간 imbalance 가 심하다. 학습 데이터 (training data) 가 총 8,331장으로 이루어져 있으나 그 중 1,994장이 person class에 속해있었으며, 가장 data가 적은 class (151장) 와는 13배가량 차이가 난다. 본 논문에서는 20개 클래스 간 이미지 개수를 유사하게 맞추기 위해 이미지가 부족한 클래스는 위에서 설명한 것처럼 추가적인 데이터 보충 단계를 진행하였고 일부 데이터가 너무 많은 클래스는 250여장으로 균일하게 줄여서 전처리 작업을 수행하였고 이에 대한 영향도 분석하였다.

3. Multi-Class Training

원 데이터 셋인 Pascal VOC 에서는 한 이미지 내 여러 클래스 (multi-class) 가 다수 존재한다. 이런 다중 클래스를 포함하는 이미지는 딥러닝 네트워크 훈련 시 인식률에 악영향을 미칠 수 있게 된다. 그래서 이런 다중 클래스를 포함한 이미지를 제거하고 훈련함으로의 인식을 변화를 확인하였다.

III. 실험

본 논문에서는 데이터 셋의 문제점들을 면밀히 관찰하기 위해 가장 보편적으로 사용되는 VGGNet16을 사용하였고 ImageNet에 Pretrain 된 가중치에 Header 만 추가하여 재학습 시켰다. 분석 데이터 셋의 조합은 다음과 같다.

- **ORG** : Pascal VOC 기본 데이터 셋
- **DS** : ORG 데이터 셋에 각 클래스 별로 300장 이상 이 되도록 데이터 보충
- **DSB** : DS 데이터셋에서 모든 클래스들이 균일하게 300장이 되도록 조정
- **NMC** : ORG 데이터 셋의 훈련 데이터에서 여러 클래스가 아닌 한 개의 클래스 만 포함하도록 조정한 데이터 셋
- **NMCB** : NMC 데이터셋에서 모든 클래스들이 균일하게 250장이 되도록 추가

그림 2에서 알 수 있듯이 ORG 데이터 셋에 비해 DS 데이터 셋에서 성능이 많이 향상된 것을 알 수 있다. 이로서, 데이터 양이 확보되는 것이 정확도 향상에 중요하다는 것을 알 수 있다.

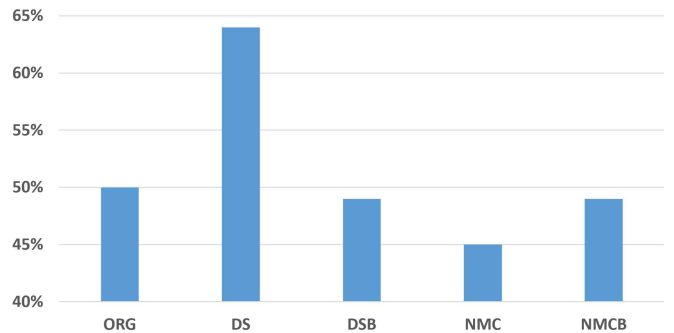


그림 2 Pascal VOC 데이터 셋 조합에 따른 VGGNet19 성능

두 번째로 클래스 별 불균형을 해결하기 위한 DSB 성능을 보면 오히려 조금 줄어든 것을 확인할 수 있다. 그 원인은 데이터 양이 많았던 일부 클래스의 데이터 양의 크기가 줄어들어서 해당 클래스의 정확도 감소가 일어났기 때문이다. 이로서 클래스 균형 보다는 데이터 양 확보가 정확도에서는 더 중요하다는 것을 알 수 있다.

세 번째로 ORG 훈련 이미지에서 멀티 클래스 이미지를 제거한 NMC의 성능이 제일 줄어든 것을 알 수 있다. 그 이유는 NMC에서 데이터 양의 줄어들기 때문에 거기서 오는 성능 열화가 상당한 것으로 분석된다. 두 번째의 경우와 마찬가지로 데이터 양이 적으면 클래스 균등화 혹은 멀티 클래스 제거와 같은 효과는 크지 않거나 오히려 열화의 원인이 된다.

IV. 결론

본 논문에서는 커스텀데이터 셋들에서 발생할 수 있는 데이터 부족, 클래스 불균형, 멀티 클래스 인식에 관한 어려운 점들을 분석해 보았고 실험결과를 통해 데이터 셋 양 확보가 성능에 제일 중요한 요인이라는 것을 보였다.

ACKNOWLEDGMENT

이 논문은 2020년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기본연구사업(No. NRF-2019R1F1A1062878)

참 고 문 헌

- [1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A Large-Scale Hierarchical Image Database. IEEE Computer Vision and Pattern Recognition (CVPR), 2009
- [2] <https://opensource.google/projects/open-images-dataset>
- [3] Tsung-Yi Lin and etc, "Microsoft COCO: Common Objects in Context"
- [4] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", NIPS 2015
- [5] Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition"
- [6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition"