

Class-Balanced Loss를 이용한 이미지 분류

*박지희 **황원준

아주대학교

*rosebird@ajou.ac.kr

Image Classification using Class-Balanced Loss

*Jihee Park **Wonjun Hwang

Ajou University

요약

Long-tail problem은 class 별로 sample의 개수에 차이가 있어 성능에 안 좋은 영향을 미치는 것을 말한다. 본 논문에서는 cost-sensitive learning 중 Class-Balanced Loss를 이용해 성능을 개선하여 Long-tail problem을 해결하려고 한다. 먼저, balanced data set과 imbalanced data set의 성능 차이를 살펴보도록 할 것이다. 그 후, Class-Balanced Loss를 3가지 버전으로 이용해 그 성능을 측정하고 분석해 볼 것이다.

1. 서론

카메라가 내장된 모바일 기기의 등장으로 이미지, 영상 데이터의 수가 폭발적으로 증가하게 되었고 이를 처리할 인공지능의 필요성 또한 증가하게 되었다. 현재 여러 산업분야에서 이미지, 영상을 처리하는 컴퓨터 비전을 필요로 하고 그 활용영역이 점차 확장되고 있다. 이러한 컴퓨터 비전의 대표적인 기술 중 가장 기초적인 응용 분야라고 할 수 있는 classification은 이미지 속 객체를 인지하여 그 클래스를 분류해 내는 기술이다. 예를 들어 강아지 사진 10000 장, 고양이 사진 10000 장이 있을 때 총 20000 장의 사진을 학습한 후 새로 입력되는 사진이 강아지인지 고양이인지 분류하는 것이 바로 classification이다. 하지만 실제 산업에서 필요한 classification은 이렇게 간단하지 않을 것이다. 분류해야 할 클래스의 개수(class label 수)도 훨씬 많을 것이고, 학습할 때 사용될 클래스 별 이미지의 개수(class sample 수)도 변화가 있을 것이다. 이미지 분류를 위해 모델을 학습시킬 때 학습 데이터가 가진 class label 수나 class sample 수 등 학습 데이터의 특징에 따라 모델의 성능이 바뀌는 것은 쉽게 생각할 수 있을 것이다.

그렇다면 과연 class 별로 학습 데이터의 sample 개수가 다를 경우 어떻게 될지 의문을 가질 수 있다. 예를 들어 주변에서 학습 데이터를 직접 수집한다고 생각해보았을 때 주위에서 쉽게 볼 수 있는 강아지, 고양이, 비둘기 등의 학습 데이터는 얻기가 쉽기에 sample의 개수를 많이 수집할 수 있을 것이다. 반면 개미핥기, 캥거루 등 주위에서 쉽게 보기 어려운 class의 학습 데이터는 sample의 개수를 적게 수집할 수밖에 없을 것이다. 실제 데이터들을 살펴보면 class

별로 sample의 개수에 차이가 있는 imbalance한 문제를 발견할 수 있고 이 차이가 class 별로 심하게 나타나는 현상을 long-tailed problem이라 부른다. 이렇게 클래스 별로 sample 개수가 차이가 나는 long-tailed data를 학습한다면 sample의 개수가 많은 클래스의 경우 학습을 더 많이 할 수 있어 분류를 잘 할 수 있겠지만 sample의 개수가 적은 클래스의 경우 학습을 많이 하기 어려워 분류를 잘하지 못할 것이고 이는 모델의 전체 성능을 떨어뜨리는 방향으로 작용할 것이다.

본 연구는 imbalance한 data set을 학습할 때 성능을 떨어뜨리지 않는 방법으로 Class Balanced Loss를 적용하여 학습한 결과를 분석해볼 것이고, loss function에 class-balanced term을 기존의 term의 제공근과 제공으로 바꾸어 성능을 확인하는 연구를 수행하고자 한다.

2. Class Balanced Loss

long tail problem을 해결할 수 있는 방법에는 re-sampling, cost-sensitive learning, logit adjustment, transfer learning, data augmentation, representation learning, classifier design, decoupled training, ensemble learning 같이 여러 방법이 존재한다 [1]. 본 연구에서는 long tail problem 해결을 위한 방법으로 cost-sensitive learning 중 Class-Balanced Loss를 이용한 방법을 채택하였다[2].

Class-Balanced Loss란 기존의 loss function에 effective number의 역을 곱한 것이고, 수식1과 같이 표현할 수 있다.

$$CB(p, y) = \frac{1}{E_{n_y}} L(p, y)$$

수식 1 Class-Balanced Loss

effective number란 중복되지 않는 sample의 expected volume이다. 예를 들어 크기가 N인 feature sample S에서 각 volume이 1인 sample을 뽑는다고 가정했을 때 data를 sampling하면 할수록 그 data들은 S를 더 많이 cover하게 될 것이고 sampling한 데이터들은 S를 더 잘 나타내게 된다. 실제로 data를 sampling할 때 뽑은 data가 기존에 sampling되었던 data에 포함될 수도, 포함되지 않을 수도, 겹칠 수도 있을 것이다. sampling된 data들은 겹칠 수도 있기에 이때 중복되지 않는 sample의 volume을 effective number로 정의한다.

본 연구에서는 먼저 class 별로 sample 수에 차이가 없는 일반적인 data set과 class 별로 sample 수에 차이가 나는 Imbalance한 data set을 이용해 학습을 진행시켜 그 차이를 살펴볼 것이다. 그 후 일반적인 Class-Balanced Loss, 기존의 Class-Balanced term을 제공근한 버전인 Squared Root Class-Balanced Loss, 기존의 Class-Balanced term을 제공한 버전인 Squared Class-Balanced Loss 총 3가지 버전으로, Loss function은 Softmax, Sigmoid, Focal 총 3가지를 이용하여 성능을 측정하고 분석할 것이다.

3. 성능 분석

3.1 성능 분석 환경

성능을 분석하기 위해 google의 colab에서 런타임 유형을 gpu로 설정하여 환경을 세팅하였다. 학습모델로는 ResNet-32 모델을 사용하였고, train epoch은 200, evaluation epoch은 2로 설정하였다. train batch size는 128, evaluation batch size는 100으로 설정하였고, momentum은 0.9로 설정하였다. data set으로는 CIFAR-10 data set을 사용하였고, imbalance한 data version에서는 imbalance factor로 0.02를 사용했고, balance한 data version에서는 imbalance factor를 None으로 설정했다. learning rate의 경우 초기엔 0.1로 설정하였다가 160 epoch, 180 epoch에서 0.01단위로 감소하도록 설정하였다. Balanced data set의 경우 beta는 0.0, gamma는 1.0으로 설정한 뒤 loss function의 종류를 softmax, sigmoid, focal로 변경해가며 성능을 측정하였다. Class-Balanced term을 적용하지 않은 imbalanced data set의 경우 beta는 0.0, gamma는 1.0, imbalance factor는 0.02로 설정한 뒤 loss function의 종류를 softmax, sigmoid, focal로 변경해가며 성능을 측정하였다. 3가지 버전의 Class-Balanced term을 적용한 imbalanced data set의 경우 beta는 0.9999, gamma는 1.0, imbalance factor는 0.02로 설정한 뒤 loss function의 종류를 softmax, sigmoid, focal로 변경해가며 성능을 측정하였다.

3.2 성능 분석 결과

Dataset Name	CIFAR-10	
Imbalance	X	O(0.02)
Softmax	0.8535	0.769
Sigmoid	0.9286	0.7724
Focal(gamma=1.0)	0.8581	0.7727

표 1 long-tailed CIFAR-10과 일반 CIFAR-10에 대해 다양한 loss function으로 학습시킨 ResNet-32 모델의 Classification accuracy

표 1을 보면 Imbalance한 long-tailed CIFAR-10보다 balance한 일반 CIFAR-10 dataset의 accuracy가 모든 loss function에 대해 높은 것을 확인할 수 있다. 이를 통해 class 별로 sample수가 다른 imbalance한 data set의 경우 전체 성능이 떨어지는 long tail problem을 확인할 수 있다.

Dataset Name	Long-Tailed CIFAR-10(Imbalance=0.02)			
Class-Balanced term	X	CB	\sqrt{CB} (square root ver.)	CB^2 (squared ver.)
Softmax	0.769	0.7805	0.7692	0.6086
Sigmoid	0.7724	0.7854	0.7782	0.6402
Focal(gamma=1.0)	0.7727	0.7901	0.7737	0.6595

표 2 long-tailed CIFAR-10에 대해 다양한 Class-Balanced term을 적용시킨 loss function으로 학습시킨 ResNet-32 모델의 Classification accuracy

표 2를 보면 Imbalance한 long-tailed CIFAR-10에 대해 loss function에 effective number의 역을 곱한 일반 Class Balanced loss를 적용시킨 버전, effective number 역의 제공근을 곱한 square root 버전, 아무것도 곱하지 않은 기본 loss 버전, effective number 역의 제공을 곱한 squared 버전 순으로 accuracy가 높은 것을 확인할 수 있다.

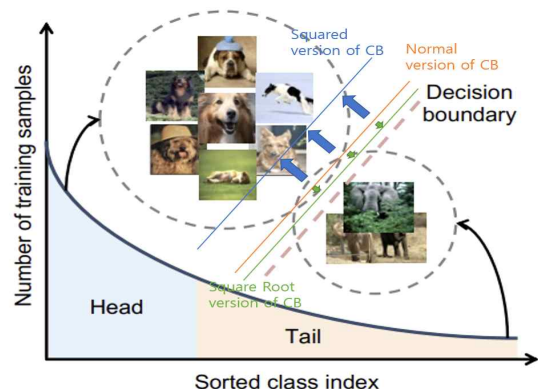


그림 1 Decision boundary

Class Balanced loss를 적용시키는 방법의 경우 effective number를 통해 decision boundary를 이동시켜 성능을 높이는 것으로 이해할 수 있다. Loss function에 effective number의 역을 곱한 일반 Class balanced loss의 성능이 가장 높은 것으로 보아 일반버전의 CB가 decision boundary를 가장 올바른 위치에 가깝게 이동시켰다고 이해할 수 있을 것이다. Loss function에 effective number 역의 제곱근을 곱한 Square root 버전의 CB의 경우 일반버전의 CB보다 성능이 낮지만 기존의 아무것도 곱하지 않은 일반 loss 버전에 비해 성능이 높은 것을 확인할 수 있다. 이는 일반버전의 CB보다 더 적은 값의 weight를 곱한 것이기에 일반버전의 CB보다 decision boundary를 덜 이동시켜 일반 버전의 CB와 일반 loss 버전의 사이에 위치하는 것으로 이해할 수 있을 것이다. Loss function에 effective number 역의 제곱을 곱한 squared 버전의 CB의 경우 나머지 세경우에 비해 월등히 성능이 떨어진 것을 확인할 수 있다. 이는 CB term의 제공한 값을 weight에 곱한 것이기에 너무 큰 값을 weight으로 곱한 것이다. 따라서 적은 sample의 class에 매우 큰 weight을 주고, 많은 sample의 class에 매우 적은 weight을 주어 오히려 반대 방향으로 imbalance를 촉진시켜 decision boundary를 반대방향으로 매우 이동시켰다고 이해할 수 있을 것이다.

4. 결론 및 향후 연구

class 별로 sample의 개수가 다른 imbalance dataset을 가지고 훈련을 시킬 때 성능이 낮아진다. 이러한 imbalance dataset을 학습시킬 때 성능이 낮아지지 않도록 하는 많은 선행 연구들이 존재한다. 본 연구는 많은 선행 연구들 중 class-balanced loss based on effective number of samples라는 논문을 기반으로 한 연구로 해당 논문에서 제시한 class-balanced term을 변형해보며 모델의 성능에 어떠한 영향을 미칠 지 알아보려고 하였다.

long tail problem을 해결하여 학습 모델의 성능을 높이기 위해 본 연구에서는 cost-sensitive한 re-weighting 방법을 적용하였다. 적은 sample을 보유한 class에 높은 weight을 주고, 많은 sample을 보유한 class에 낮은 weight을 주어 학습의 balance를 맞추는 기법을 채택하였다. effective number의 역을 loss function에 곱해 long tail data set을 학습하는 모델의 성능을 높였던 기존의 연구에서 더 나아가 loss function에 곱하는 class balanced term을 effective number 역의 제곱근, effective number 역의 제곱으로 변형하여 성능을 측정해 보았다. 측정 결과 기존의 class balanced loss를 적용시킨 버전, class balanced의 square root 버전, 일반 loss를 적용시킨 버전, class balanced의 squared 버전 순으로 accuracy가 높은 것을 확인할 수 있었다.

class balanced loss보다 더 좋은 성능을 갖는 새로운 term을 만드는 것을 연구과제로 삼을 수도 있고, long tail problem을 해결하는 방법을 cost-sensitive learning이 아

닌 다른 방법으로 변경하여 연구과제를 발전시킬 수 있을 것이라 기대한다.

5. 사사

본 연구는 2022년 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학사업의 지원으로 수행되었음 (2022-0-01077)

참고문헌

[1] Yifan Zhang, Bingyi Kang, Vryan Hooi, Schuichen Yan, "Deep Long-Tailed Learning: A Survey", 2021.

[2] Yin Cui, Menglin Jia, Tsung-Yi Lin, Yang Song, Serge Belongie, "Class-Balanced Loss Based on Effective Number of Samples", 2019.