

작은 이미지의 효과적인 데이터 증강 기법

구승연, 노시동, 정기석*
한양대학교

rrxloyeon@hanyang.ac.kr, sdroh1027@hanyang.ac.kr, *kchung@hanyang.ac.kr

An Effective Data Augmentation Technique for Small Images

Koo, Seung-Yeon, Roh, Si-Dong, Chung, Ki-Seok*
Hanyang Univ.

요약

Deep neural networks 에서 발생하는 과적합 문제를 효과적으로 해결하는 방법 중 하나로 data augmentation 을 사용할 수 있다. Data augmentation 기법에는 회전이나 노이즈 추가 등의 단순한 방법 외에도 Mixup, Cutout, RICAP 과 같이 새로운 방법들이 연구되고 있다. 본 논문에서는 기존의 RICAP 과 Mixup 을 결합하여 기존의 RICAP 이 가지고 있었던 정보 손실 문제를 해결하고자 하였다. 본 논문에서 제안한 RIMix 는 먼저 입력 이미지를 두 영역으로 분할한 뒤, 각 영역에 Mixup 을 적용하여 결과적으로 하나의 입력에 총 네 개의 이미지가 포함되도록 구현한다. RIMix 는 CIFAR-10 와 CIFAR-100 데이터셋에 대해서 테스트했을 때 RICAP 과 비교해 각각 0.01%, 0.43%의 성능을 향상시킬 수 있음을 확인하였다.

I. 서론

Deep neural networks 에서 발생하는 과적합 문제는 데이터셋의 크기를 확대하여 해결할 수 있지만 순수하게 새로운 데이터를 추가하는 것은 많은 비용이 필요하다. 따라서 데이터의 추가 없이 data augmentation 을 통해 학습 데이터 양을 늘리는 것이 일반적이다. 컴퓨터 비전 영역에서는 주로 이미지를 회전하거나 노이즈를 추가하는 등 비교적 단순한 변환을 통해서 모델의 일반화 성능을 개선한다. 단순한 변환으로 데이터셋의 다양성을 강화하는 방법 외에도 Cutout [1]과 Mixup [2]과 같은 data augmentation 기법들이 연구되고 있으며, Takahashi et al.이 제안한 random image cropping and patching (RICAP) 은 여러 장의 이미지를 이어 붙여 새로운 input 을 만드는 방식의 data augmentation 기법으로 CIFAR-10 데이터셋에서 기존보다 오차율을 2.19% 향상시켰다 [3]. 그러나 기존의 RICAP 은 네 장의 이미지를 cropping 하여 이어 붙이기 때문에 CIFAR 데이터셋과 같이 원본의 크기가 작은 이미지의 경우에는 cropping 을 수행한 이후 이미지가 매우 작아지며 정보를 충분히 포함하지 못하는 경우가 발생한다 [4]. 본 논문에서는 이러한 문제를 해결하기 위하여 RICAP 과 Mixup 방식을 결합한 RIMix (RICAP and Mixup)를 제안한다. RIMix 는 input 을 두 구역으로 나누고, 한 영역 당 두 장의 이미지를 Mixup 방식으로 할당하여 총 네 장의 이미지를 하나의 input 에 포함시킬 수 있다. RIMix 는 기존의 RICAP 에 비해 CIFAR-10 에서는 0.01%, CIFAR-100 에서는 0.43% 성능을 향상시킬 수 있었다.

II. 본론

RICAP 은 CIFAR 데이터셋(32x32)과 같이 입력 이미지의 크기가 작은 데이터에 대해서는 이미지 분할 시 각 이미지 패치가 매우 적은 수의 픽셀만을 포함하기

때문에 object 의 정보가 누락되는 문제가 발생한다 [4]. 따라서 cropping 되는 영역의 크기를 확장하기 위하여 RICAP 과 Mixup 을 섞는 새로운 알고리즘 RIMix 를 제안한다. RIMix 는 RICAP 과 마찬가지로 하나의 input 에 네 장의 이미지를 포함하지만 input 을 두 영역으로 분할하고 Mixup 이 적용된 이미지를 cropping 하여 각 영역에 할당한다. (그림 1)

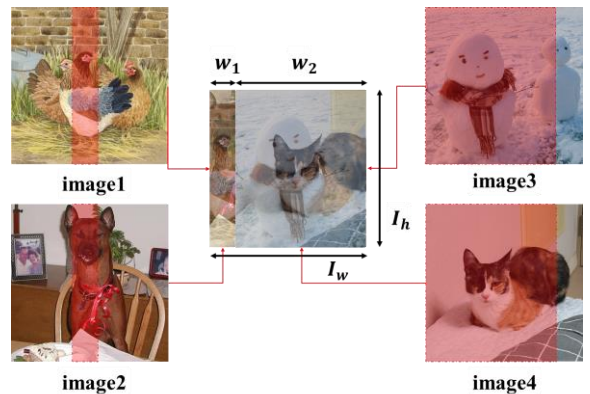


그림 1. RIMix 알고리즘 예시

이미지 패치의 너비 w_1 와 w_2 는 beta distribution 에 의해 무작위로 결정된다. RIMixHor 는 RIMixVer 과는 상반되게 너비를 고정하고 높이를 분할한다.

$$w_1 = I_w \text{Beta}(\beta, \beta)$$
$$w_2 = I_w - w_1$$

두 이미지 x_1 과 x_2 는 아래 식에 의해 Mixup 된다.

$$\tilde{x} = \lambda x_1 + (1 - \lambda)x_2$$
$$\lambda = \text{Beta}(\alpha, \alpha)$$

표 1. WideResNet 28-10 에서 CIFAR-10, 100 테스트 성능 비교 (%)

	baseline	RICAP	Mixup	RIMixVer	RIMixHor
CIFAR-10	3.707 ±0.10	2.653 ±0.09	2.827 ±0.04	2.643 ±0.09	2.760 ±0.12
CIFAR-100	18.503 ±0.35	16.917 ±0.39	17.300 ±0.21	16.487 ±0.12	17.433 ±0.13

이때, RIMix 된 input 의 label 은 one-hot vector y 에 의해 아래와 같이 weighted sum 연산으로 결정된다.

$$\hat{y} = \sum W_i(\lambda y_{2i-1} + (1-\lambda)y_{2i})$$

$$W_i = \frac{w_i}{I_w}$$

본 논문에서는 [2], [3]에 근거하여 영역 분할을 위한 beta distribution 의 β 를 0.3 으로 설정하였으며 Mixup 을 위한 beta distribution 의 α 는 1.0 으로 설정하였다. 실험에 사용한 데이터셋은 CIFAR-10 과 CIFAR-100 데이터셋으로 32x32 크기의 이미지 데이터로 이루어져 있으며 각각 10 개, 100 개의 class 로 구성되어 있다. 표 1 과 그림 2 는 WideResNet 28-10 [5]으로 테스트 성능을 측정한 결과이다. 표 1 의 결과는 3 회의 최소 테스트 오차율의 평균으로 기술했다. 그림 2 는 각 data augmentation 기법들의 차이를 분명하게 보여주기 위하여 y 축을 로그 스케일로 설정하였다.

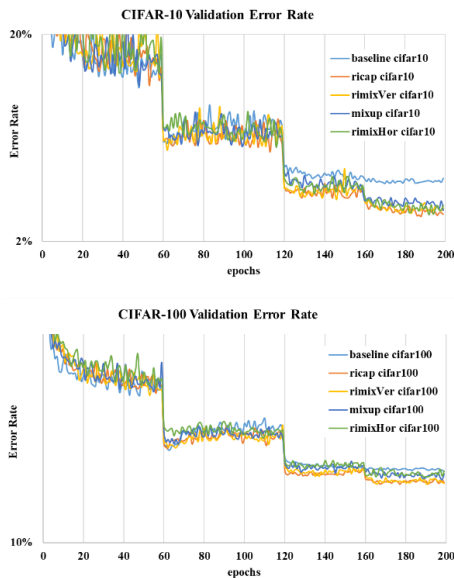


그림 2. CIFAR-10, 100 실험 결과

표 1 의 결과를 보면, 제안하는 RIMixVer 가 다른 방법과 비교해 가장 낮은 오차율을 보임을 알 수 있다. RIMixVer 의 오차율은 CIFAR-10 에서 RICAP 보다 0.01%, Mixup 보다 0.184% 감소했으며 CIFAR-100 에서 RICAP 보다 0.43%, Mixup 보다 0.813% 감소했다. 반면 RIMixHor 의 경우 CIFAR-10 에서 Mixup 보다 오차율이 0.067% 감소하였지만 RICAP 보다 0.107% 증가하였고 CIFAR-100 결과 Mixup 과 RICAP 에 비하여 각각 0.133%, 0.516% 만큼 증가한다. 이를 통해 데이터의 특성에 따라 RIMix 의 이미지 분할 방향이 모델의 성능에 영향을 미칠 수 있음을 확인하였다. 결과적으로 RIMix 는 적절한 분할 방향을 설정할 경우 이미지 분할 과정에서의 정보 손실 문제를 해결함으로써 기존의 RICAP 보다 성능 면에서 명확하게 개선됨을 알 수 있다.

III. 결론

과적합 문제는 딥러닝 전반에서 발생하는 문제이며 data augmentation 기법으로 해소할 수 있다. 효율적인 data augmentation 을 위해 다양한 연구들이 이루어지고 있으며 그 중 RICAP 은 CIFAR-10 데이터셋에서 최신 성능을 달성하였다. RICAP 은 이미지를 분할하여 이어 붙이는 방법이기 때문에 자칫 필요한 정보를 포함하지 못하는 문제점이 있다. 이를 해결하기 위하여 본 논문에서는 RICAP 과 Mixup 방식을 결합한 RIMix 를 제안한다. Input 을 네 영역으로 구분하는 RICAP 과는 다르게 RIMix 는 두 영역으로만 나누어 하나의 이미지 패치가 더 넓은 영역을 가질 수 있으며 Mixup 을 통해 RICAP 과 동일한 수의 class 를 하나의 input 에 포함시킬 수 있다. CIFAR-10, 100 데이터셋과 WideResNet 28-10 으로 분류 성능을 비교했을 때 RIMixVer 의 오차율은 CIFAR-10 에서 RICAP 보다 0.01%, Mixup 보다 0.184% 향상되었고 CIFAR-100에서 RICAP 보다 0.43%, Mixup 보다 0.813% 감소하며 성능이 개선됨을 확인할 수 있었다.

ACKNOWLEDGMENT

이 논문은 2022 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(No.2021-0-00131, 제조검사장비 경량화를 위한 지능형 엣지컴퓨팅 반도체 개발)을 받아 수행된 연구임

참 고 문 헌

- [1] T. DeVries, and G.W. Taylor. "Improved regularization of convolutional neural networks with cutout," arXiv:1708.04552, 2017.
- [2] H. Zhang, M. Cisse, Y.N. Dauphin, & D. Lopez-Paz, "mixup: Beyond empirical risk minimization," in Proceedings of the International Conference on Learning Representations, 2018.
- [3] R. Takahashi, T. Matsubara and K. Uehara, "Data Augmentation Using Random Image Cropping and Patching for Deep CNNs," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 9, pp. 2917-2931, 2020.
- [4] 구승연, 노시동, 정기석, "이미지 패치 연결 방식을 통한 데이터 증강 기술의 경향성에 관한 연구," 대한임베디드공학회 추계학술대회 논문집, pp.278-281, 2022.
- [5] S. Zagoruyko and N. Komodakis. "Wide Residual Networks," in Proceedings of the British Machine Vision Conference (BMVC), pp. 87.1-87.12. BMVA Press, 2016.