

## 격자Mix

한형근, 이정우

서울대학교

hygnhan@snu.ac.kr, junglee@snu.ac.kr

## LatticeMix

Hyeonggeun Han, Jungwoo Lee

Seoul National Univ.

## 요약

본 논문에서는 새로운 형태의 data augmentation 방식인 격자Mix(LatticeMix)를 제안한다. LatticeMix는 여러 이미지를 겹치지 않고 mixup하는 방법으로 이미지들의 크기만을 조절하여 합쳐 새로운 데이터를 생성한다. 실험을 통해 기존 baseline 모델의 정확도를 향상시킴을 보인다.

## I. 서론

Mixup은 널리 알려진 data augmentation 기법 중 하나로 input image와 target label의 convex combination을 통해 새로운 가상의 학습 이미지들을 생성한다.[1] 간단한 interpolation에도 불구하고 얻어지는 가상의 이미지들은 효율적으로 학습 분포의 support를 크게 만들었고 이를 통해 calibration, accuracy, robustness 측면에서 큰 성능향상을 가져왔다. 이후 많은 Mixup과 관련된 연구들이 활발히 진행되고 있다.

Mixup의 성능을 더욱 향상시키고 단점을 보완하기 위해 여러 가지 많은 변형 Mixup들이 제안되었다. CutMix는 cut-and-paste 방법으로 기존 interpolation 방법의 Mixup이 갖고 있던 단점 중 하나인 부자연스러운 Mixup samples를 해결한다.[2] CutMix에서는 무작위로 patch를 잘라내어 이미지 내 물체에 대한 정보가 온전히 보존되지 못하는 문제가 발생하는데 이를 해결하기 위해 saliency 정보를 이용하여 patch를 잘라내는 방법들도 제안되었다.[3, 4] 기존 interpolation, cut-and-paste 기법에서 벗어나 resize-and-paste하는 방법 또한 제안되었다.[5]

본 논문에서는 새로운 형태의 Mixup인 격자Mix(LatticeMix)를 제안한다. LatticeMix는 앞에서 나열된 방법들과는 달리 이미지들을 interpolation하지 않으며 source image를 patch형태로 만들어서 target image에 paste하지도 않는다. 여러 이미지들을 격자무늬 형태로 동등하게 이어붙임으로써 mixup되는 이미지들의 정보가 온전히 합성된 가상 이미지에 누락 없이 들어가게 된다.

## II. 본론

Source image에서 무작위로 patch를 잘라내어 target image에 붙임으로써 새로운 가상의 이미지를 만들어내는 기존 cut-and-paste 방법과 달리 본 논문에서 제안하는 방법은 여러 이미지를 공간적 차원(spatial dimension)에서 이어붙임으로써 mixup을 수행한다. 자세히 설명하면, 기존 mixup 방법들과 달리 LatticeMix는 4개의 이미지를 사용하여 mixed sample을 만든다. 이미지의 높이와 너비를  $H, W$ 라고 할 때, 먼저 무작위로 이미지 내의 위치  $(h, w)$  하나를 선택한다. 여기서  $h, w$ 는 각각  $0 \leq h \leq H, 0 \leq w \leq W$ 를 만족해야한다. 이를 위해 beta

distribution  $B(\alpha, \alpha)$ 에서  $\tau$ 를 sample한다.  $\tau$ 를  $H, W$ 에 곱하여  $h, w$ 를 얻는다. 위치를 선택하고 나면 이를 기준으로 4개의 이미지를 크기 조절을 통해 공간적 차원을 따라 격자무늬로 이어 붙인다. 첫 번째 이미지는  $(h, w)$ 로 크기를 조절하고 두 번째 이미지는  $(h, W-w)$ 로, 세 번째 이미지는  $(H-h, w)$ 로, 마지막 네 번째 이미지는  $(H-h, W-w)$ 로 크기를 조절한다. 그 다음에 이 4개의 이미지를 공간적 차원에 따라 이어 붙여 합성된 이미지의 크기가  $(H, W)$ 가 되도록 만든다. 아래 그림 Fig 1.은 LatticeMix의 mixup방법을 시각적으로 보여준다.

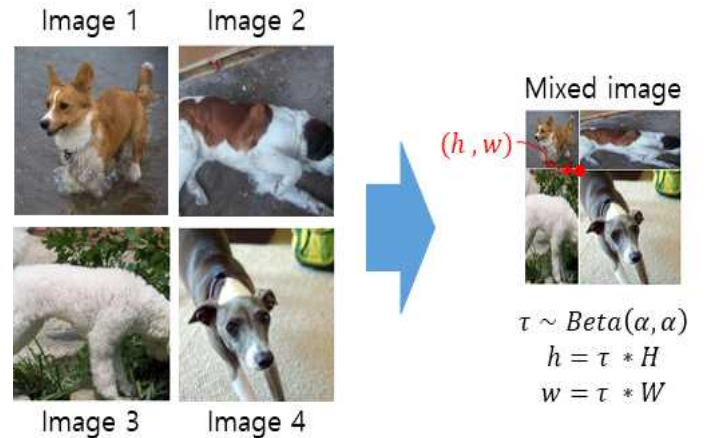


Fig 1. LatticeMix

합성된 가상의 이미지에 대한 합성된 target label은 기존 방법인 CutMix를 따라 얻는다. 전체 이미지에 대한 각 이미지들의 면적 비율을 가중치로 사용하여 각 이미지들의 target label의 weighted sum을 합성된 target label로 사용한다. LatticeMix에 사용된 각 이미지  $x_1, x_2, x_3, x_4$ 와 그에 상응하는 target label들을  $y_1, y_2, y_3, y_4$ 라고 하면, 다음의 가중치를 얻을 수 있다.

$$w_i = \frac{Area(x_i)}{H \times W}$$

여기서  $Area(x_i)$ 는 mixed image에서  $x_i$ 가 차지하는 면적을 의미한다. 이를 이용해 mixed image에 상응하는 target label은  $\sum_i w_i y_i$ 로 얻는다.

LatticeMix가 정말 유의미한 data augmentation 기법으로 모델의 성능향상을 가져오는지 확인하기 위해 WideResNet-28-10 모델과 CIFAR-10 데이터를 이용하여 accuracy를 측정했다. 실험결과는 아래 표에 나타나 있으며 전체 200 epochs에 대해 early stopping을 사용하였다. Table 1.에서 볼 수 있듯이 LatticeMix는 유의미한 성능향상을 가져왔다.

	CIFAR-10 (Accuracy)
Baseline	95.78
LatticeMix	96.70

Table 1. Baseline, LatticeMix 성능 비교

### III. 결론

본 논문에서는 mixup 기반의 새로운 형태의 data augmentation 방식으로 LatticeMix를 제안하였다. 또한 실험을 통해 accuracy 측면에서 성능향상을 이끌어냄을 보였다.

### 참 고 문 헌

- [1] Zhang, Hongyi, et al. "mixup: Beyond empirical risk minimization." arXiv preprint arXiv:1710.09412 (2017).
- [2] Yun, Sangdoo, et al. "Cutmix: Regularization strategy to train strong classifiers with localizable features." Proceedings of the IEEE/CVF international conference on computer vision. 2019.
- [3] Kim, Jang-Hyun, Wonho Choo, and Hyun Oh Song. "Puzzle mix: Exploiting saliency and local statistics for optimal mixup." International Conference on Machine Learning. PMLR, 2020.
- [4] Uddin, A. F. M., et al. "Saliencymix: A saliency guided data augmentation strategy for better regularization." arXiv preprint arXiv:2006.01791 (2020).
- [5] Qin, Jie, et al. "Resizemix: Mixing data with preserved object information and true labels." arXiv preprint arXiv:2012.11101 (2020).