|  |
| --- |
| **Trying Multi-Categories for Convolution Neural Network** |
|  |
| **요 약**  3가지 카테고리를 적절히 섞은 이미지를 classification 모델에 학습시키면 cutout 혹은 cutmix 알고리즘을 적용한 모델보다 더 좋은 성능을 보인다. 전처리를 통한 이미지 입력은 classification에서 있어서 더 강력한 일반화 알고리즘을 제공한다. | | |

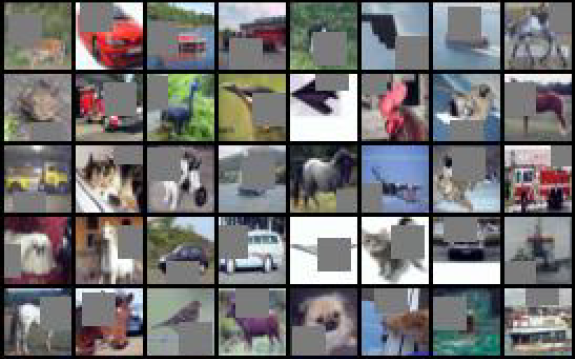
1. **서 론**

최근에 이미지를 분류함에 있어서 전처리를 통해 성능을 향상시키는 연구가 진행되었다. 기존의 연구는 이미지의 일부를 없애고[2], 2개를 합성하는[1] 등의 전처리 알고리즘을 제시하였다. 기존의 연구에서 오마주를 얻어 다른 전처리 알고리즘을 제시한다. 3개 이상의 카테고리에 대해 합성하는 전처리 알고리즘으로 몇 개의 카테고리를 합성하는 것이 최대의 성능을 보이는가에 대한 지표를 제시한다.

관련된 연구로는 이미지의 일부를 잘라내서 모델에 학습시킨 연구[1]와 이미지 2개를 적절히 합성시킨 연구[2] 등이 존재한다.

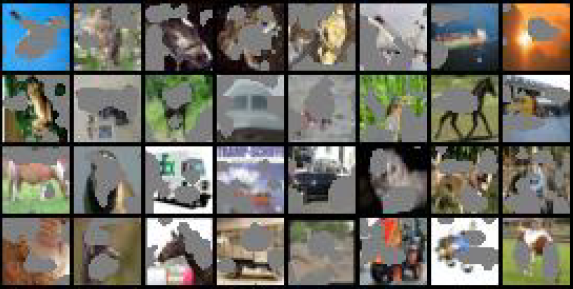
1. **기존 연구**
   1. **Improved Regularization with Convolution Neural Networks with Cutout**

이 논문은 이미지 분류의 트레이닝 데이터에 있어 이미지의 일부를 가려 학습시켰다.

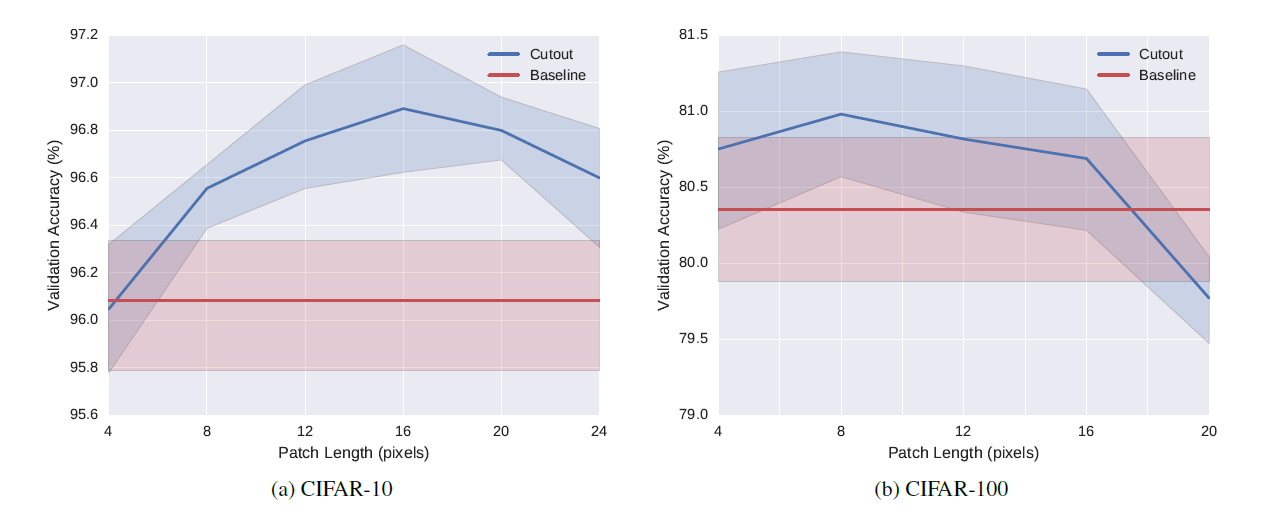


이미지의 일부분을 가림으로써 image augment-ation의 효과를 얻을 수 있다. 왜냐하면 기존의 이미지와 가려진 이미지가 생성되기 때문이다. 또한 이미지의 임의의 feature를 가려도 나머지 feature들로 분류학습이 가능하기 때문에 더 강력한 classification 모델을 제공할 수 있다.

Cutout은 처음에 이미지의 중요 feature를 수동으로 가려냈다.

****

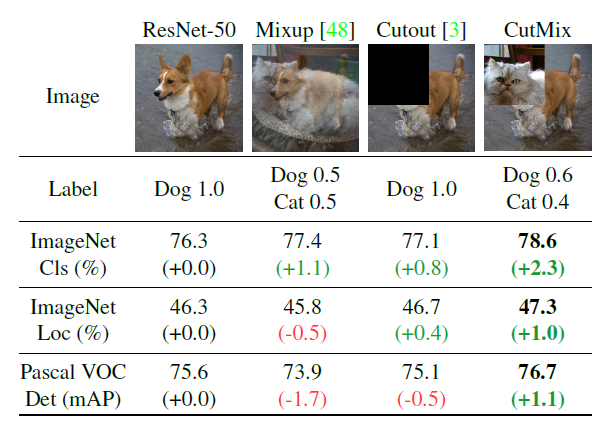
수동으로 가린 사진을 일반화하여 random으로 위친 box를 만들어 내어 가린 이미지를 생성하여 학습시켰다.



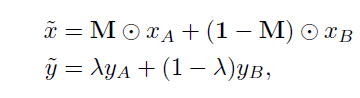
Cifar10, cifar100의 Validation Accuracy의 결과를 보면 Cutout시킨 이미지의 정확성이 더 좋다는 것을 확인할 수 있다.

* 1. **Cutmix : Regularization Strategy to Train Strong Classifiers with Localizable Features**

이 논문은 cutout에서 더 확장된 실험을 했다. Cutout은 이미지의 일부를 지우는 것에서 그쳤지만 cutmixt는 지워진 이미지의 일부에 다른 카테고리를 입력했다.

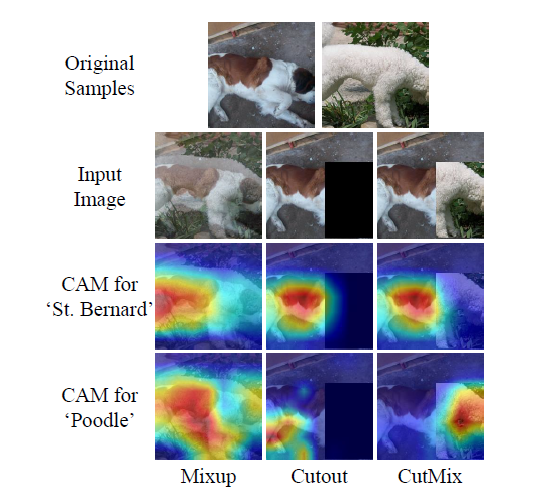


예를 들면 개의 몸에 고양이의 머리를 붙인 합성 이미지를 생성할 수 있다. 다음과 같은 처리를 거치면 2개의 카테고리에 대한 y값을 생성할 수 있다.



random으로 생성된 마스크를 이용하여 input 이미지를 구성한다. 또한 loss는 cross entropy를 이용하여 y 값을 예측할 수 있다.

Cutmix를 통하여 얻을 수 있는 효과는 하나의 인풋에 두개의 activation을 얻을 수 있다는 것이다.



기존에 진행됐던 연구들과 비교해보면 더 좋은 activation을 가진다는 것을 알 수 있다.

1. **시스템 모델**
   1. **기존 연구와 차이점 및 해결방안**

기존의 cutmix 연구에서는 2개의 카테고리에 대해서만 이미지 합성을 진행했다. 본 연구에서는 2개 이상의 카테고리를 합성함으로써 더 증가되는 카테고리의 양에 따른 성능 지표를 제시한다.

3개 이상의 카테고리에 대한 처리를 위해 beta 분포가 아닌 Dirichlet 분포를 이용한다.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Cutout | Cutmix | Ours |
| Usage of full image region | **X** | **O** | **O** |
| Regional dropout | **O** | **O** | **O** |
| Use more than two categories | **X** | **X** | **O** |

* 1. **알고리즘**

**(1)**

**(2)**

1. 에서 는 element-wise multiplication이다. M은 각각의 카테고리에 대해 random으로 생성된 마스크를 나타낸다. 카테고리를 3개 이용함으로 마스크끼리 겹치는 부분을 해결하기 위해 중복되는 부분은 빼준다. X는 해당 카테고리에 대한 이미지이다.
2. 에서 theta는 y값에 대한 비율이다. 전체 이미지에 대해서 비율을 계산하여 Cross Entropy를 계산한다.

위의 그림은 하나의 예시이다. Dog, Cat, Horse 개체의 일부분을 합친 것이다. 전체 이미지에서 해당 카테고리의 비율로 theta를 계산할 수 있다.



1. **프로젝트 내용**

현재 cifar 100에 대한 데이터로 진행해봤으며 3개의 카테고리를 적용했을 때 기존의 모델과의 성능 차이를 명시한다. 성능은 1순위 오차와 5순위 오차를 기준으로 한다.

* **Pyramid Net**

Hyper parameter information : depth =100, batch size = 64, alpha = 240, learning rate = 0.05, dataset = cifar100, epochs = 300

위와 같은 공통 변수를 가지고 트레인 시켰을 때 결과를 표로 나타낸 것이다.

|  |  |  |
| --- | --- | --- |
| 알고리즘  적용 비율 | Cutmix | Ours |
| 100% | Top-1 err: 18.72  Top-5 err : 4.69 | Top-1 err: 19.58  Top-5 err : 4.38 |
| 50% | Top-1 err: 18.74  Top-5 err : 4.72 | Top-1 err : 16.8  Top-5 err : 3.43 |

1. **결론 및 향후 연구**

Pyramidnet에서는 1순위 오차는 더 크지만 5순위 오차가 더 작은 것을 확인할 수 있었다. 또한 알고리즘을 적용하는 비율을 적절히 조절함으로써 똑같은 조건에서 더 좋은 성능을 확인할 수 있었다.

향후 연구는 pyramidnet 뿐만 아니라 다른 net에서도 똑같은 성능향상이 진행되는지를 파악해보고 activation 그래프를 제공함으로써 성능향상에 대한 근거 정보를 제시한다.

**참 고 문 헌**

**[1]**

Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, Youngjoon Yoo. 1. Clova AI Research, NAVER Corp. 2. Clova AI Research, LINE Plus Corp. 3. Yonsei University CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features arXiv:1905.04899v2, 2019

**[2]** Terrance DeVries and Graham W. Taylor. 1. University of Guelph, 2. Canadian Institute for Advanced Research and Vector Institute. Improved Regularization of Convolutional Neural Networks with Cutout arXiv:1708.04552v2, 2017

**[3]** Hongyi Zhang, Moustapha Cisse, Yann N Dauphin, and David Lopez-Paz. mixup: Beyond empirical risk minimization. arXiv:1710.09412, 2017.