

ResNet 정확도 향상을 위한 깊이별 Residual Connection 비율 조절 방법 제안

Depth-Specific Variational Scaling Method to Improve Accuracy of ResNet

송영재¹ · 장승우² · 김광수^{3†}

Youngjae Song, Seungwoo Jang and Kwangsu Kim[†]

¹성균관대학교 기계공학과

E-mail: thddudwo1313@gmail.com

²성균관대학교 전자전기컴퓨터공학과

E-mail: se93wo@gmail.com

³성균관대학교 소프트웨어학과

E-mail: kim.kwangsu@skku.edu

요 약

이미지 분야에서 우수한 결과를 내고 있는 ResNet은 깊이별 residual block의 중요도 차이를 고려하지 않고 동일한 비율로 residual connection이 이루어진다. 본 논문에서는 residual block 중요도에 따라 residual connection 비율을 달리하는 variational scaling 방법을 제안한다. 실험 결과 CIFAR-10, CIFAR-100 데이터셋에서 기존 ResNet과 비교하여 연산량 증가 없이 정확도 향상을 이루었다.

키워드 : ResNet, Residual Connection, Residual Block, Variational Scaling

1. 서 론

심층 합성곱 신경망은 컴퓨터 비전 문제에서 전통적인 기계학습 방법론을 제치고 압도적인 결과를 내고 있다. He et al.[1]은 2015년 ILSVRC(ImageNet Large Scale Visual Recognition Challenge) 이미지 인식 대회에서 우승 모델인 ResNet을 설계했다. ResNet은 깊은 신경망에서 발생하는 학습의 어려움을 해소하고자 residual block을 깊이 방향으로 쌓아 만든 인공신경망이다. Residual block은 합성곱 신경층 2개에 residual connection을 적용한 형태다. 이 때, residual connection은 합성곱층을 건너뛰어 residual block 입력값을 출력에 더해주는 기법이며 일반적으로 항등함수를 적용한다.

He et al. [2]은 항등함수를 사용하지 않는 ResNet을 실험하기 위해 모든 residual connection에 0.5의 비율을 곱해 학습한 후 기존 ResNet과 정확도를 비교하였다. 그 결과 기존 항등함수를 적용한 ResNet보다 낮은 정확도를 보였다. 하지만 이 방법은 깊이별 residual block이 이미지 분류 정확도에 기여하는 정도가 다를 수 있으므로 활용하지 않는다. 따라서 본 논문에서는 이 문제를 해결하기 위해 깊이별 residual connection의 비율을 달리하는 ‘variational scaling’ 기법을 제안한다. Variational scaling 기법은 학습 파라미터 개수 증가 없이 ResNet의 정확도를 향상시킨다.

2. Variational Scaling 기법 제안

항등함수를 사용하는 기존의 ResNet은 일반적으로 아래 수식으로 표현한다.

$$y_l = h(x_l) + F(x_l, W_l), \quad (1)$$

$$h(x_l) = x_l, \quad (2)$$

$$x_{l+1} = f(y_l). \quad (3)$$

x_l 와 x_{l+1} 는 각각 l 번째 residual block의 입력과 출력이며 $h(x_l)$ 는 residual connection 함수이다. 여기서 $h(x_l)$ 는 입력을 그대로 반환하는 항등함수이다. F 는 x_l 와 가중치 행렬인 W_l 로 이루어진 함수이며 여기서는 합성곱층을 의미한다. f 는 활성화 함수이며 여기서는 ReLU 함수를 사용한다. 기존의 ResNet은 깊이와 무관하게 항등함수를 사용하기 때문에 모든 residual connection이 일정한 비율을 갖는다. 반면에 variational scaling은 아래와 같이 표현한다.

$$h(x_l) = V(d) \times x_l, \quad (4)$$

$$d = l \times \left(\frac{15}{L}\right). \quad (5)$$

$V(d)$ 는 variational scaling 함수이며 residual connection의 위치(l)에 따라 다른 비율값을 갖는 사용자의 함수이다. L 은 residual block의 총 개수이고 $V(d)$ 는 $d \in [1, 15]$ 에서 정의된다. d 는 실험에 사용하는 $V(d)$ 를 고정된 범위 내에서 재사용하기 위해 정의

감사의 글 : 이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00990, 5G-IoT 환경에서 이기종·비정형·대용량 데이터의 고성능·저지연 처리를 위한 플랫폼 개발 및 실증)

된 매개변수다. ResNet-32에서 실험을 진행하는 경우에는 d 와 l 이 구분되지 않는다. 하지만 L 이 15가 아닌 경우에도 d 를 활용하면 $V(d)$ 의 가로축 범위가 변하지 않기 때문에 이전 케이스에서 사용한 $V(d)$ 을 그대로 재사용할 수 있다. 따라서 실험의 편의를 위해 설정한 변수이며 사용자에게 따라 d 의 범위를 마음대로 조정할 수 있다. 혹은 d 를 사용하지 않고 $l \in [1, L]$ 내에서 $V(l)$ 를 정의할 수도 있지만 L 값이 바뀐 경우에 세부 조정을 새로 해야 하는 번거로움이 있다.

본 논문에서는 residual connection이 깊이에 따라 다른 비율값을 갖도록 $V(d)$ 를 다양하게 정의하고 정확도를 서로 비교한다. 이후 최종 $V(d)$ 이 될 만한 모델에 세부 조정을 거쳐 ResNet의 정확도를 향상시키는 최적 variational scaling 모델 선정을 목표로 한다.

3. 실험 및 결과

실험에 사용한 데이터셋은 CIFAR-10과 CIFAR-100이고 실험 모델은 ResNet-32와 ResNet-56이다. 각 모델의 학습 파라미터는 He et al. [1]가 제시한 값을 기반으로 설정했다. AMD Ryzen 7 3700X 8-Core Processor와 NVIDIA GeForce RTX 2060 SUPER 환경에서 Tensorflow[3]를 활용해 학습을 진행했다. 실험 결과 ResNet-32은 가운데 층에 더 큰 scale을 부여한 경우 성능이 향상되며, ResNet-56은 얇은 층에 더 큰 비율을 부여한 경우 성능이 향상됨을 확인했다. 이 실험 결과를 바탕으로 $V(d)$ 를 세부 조정하여 정확도 향상이 가장 큰 최적의 variational scaling을 제안한다.

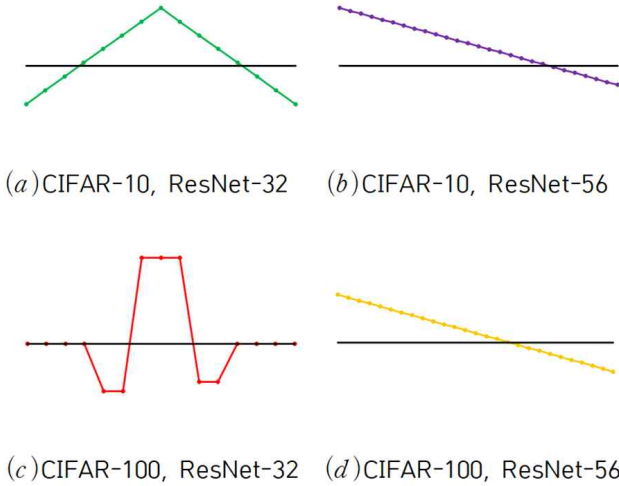


그림 1. 각 데이터별, ResNet별 최종 variational scaling 제안 모델 시각화

Fig 1. Visualization of the final proposed variational scaling model for each dataset and ResNet

그림 1은 데이터별, 실험 ResNet별로 선정된 최종 variational scaling 모델을 시각화한 것이다. 그림 1에서 가로축은 d 이고 세로축은 $V(d)$ 값을 의미한다. 검은 수평선은 값이 1로 일정한 baseline을 의미한다. 그림 1에서 (a)은 가운데가 1.3으로 최고점, 양 끝이 0.8로 최저점을 이루는 $V(d)$ 이며 CIFAR-10에서 ResNet-32의 정확도를 **1.06%** 향상시킨다. (b)은 최고점이 1.3, 최저점이 0.9인 우하향 직선꼴의 $V(d)$ 이고, CIFAR-10에서

ResNet-56의 정확도를 **0.98%** 향상시킨다. (c)는 [1, 1, 1, 1, 1, 0.75, 0.75, 1.45, 1.45, 1.45, 0.8, 0.8, 1, 1, 1]에 속하는 $V(d)$ 를 의미하고 CIFAR-100에서 ResNet-32의 정확도를 **1.02%** 향상시킨다. (d)은 최고점이 1.25, 최저점이 0.85인 우하향 $V(d)$ 이고, CIFAR-100에서 ResNet-56의 정확도를 **1.08%** 향상시킨다.

두 가지 다른 데이터셋과 다른 깊이의 ResNet에서 제안된 variational scaling을 적용하는 경우 정확도가 향상됨을 확인했다. ResNet의 구조를 전혀 바꾸지 않고 residual connection의 비율만 조정했으므로 학습 파라미터 개수 증가가 없다는 장점이 있다.

표 1. 제안 variational scaling 적용 후 모델 정확도
Table 1. Model accuracy after the proposed variational scaling is applied

Datasets	Variational Scaling	ResNet-32 Accuracy[%]	ResNet-56 Accuracy[%]
CIFAR-10	Baseline	87.81	88.39
	Proposal	88.87	89.37
	Improvement	1.06	0.98
CIFAR-100	Baseline	62.19	63.56
	Proposal	63.21	64.64
	Improvement	1.02	1.08

4. 결론 및 향후 연구

기존 ResNet은 residual connection 비율을 깊이와 관계없이 일정하게 설정했다. He et al.[2]는 residual connection의 비율을 절반으로 설정하고 실험을 진행했으나 정확도가 소폭 하락했다. 본 논문에서는 이 문제를 해결하기 위해 variational scaling 기법을 적용했다. 결과적으로 다른 데이터셋, 다른 깊이의 ResNet에서 학습 파라미터 개수 증가 없이 정확도가 향상됨을 보였다. 향후 variational scaling 기법을 일반적으로 활용할 수 있도록 다양한 해상도의 데이터셋과 더 다양한 깊이의 ResNet에서 실험을 진행하여 데이터를 확보하고자 한다.

참 고 문 헌

- [1] He, K., Zhang, X., Ren, S., Sun, J., Deep residual learning for image recognition. In: CVPR (2016)
- [2] K. He, X. Zhang, S. Ren, and J. Sun. Identity mappings in deep residual networks. In ECCV, 2016
- [3] Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M. & others (2016). TensorFlow: A System for Large-Scale Machine Learning.. OSDI (p/pp. 265–283).