

SENet 구현 및 실험 결과 보고



AutoML 연구실
석사과정 문아성
20.08.18

Introduction

Squeeze-and-Excitation Networks

기존 연구의 학습 메커니즘은 feature간의 공간적인 관계 (spatial correlation)를 찾아
성능을 향상시키는 네트워크 모듈을 만드는데에 집중 [GoogLeNet-Inception Module, ResNet-Skip Connection]

본 논문에서는 공간적인 관계가 아닌 채널간의 관계를 재조정 (Recalibrate) 를 하게 해주는
“Squeeze-and-Excitation Block” 을 제안하여 Feature Recalibration을 수행할 수 있게 해주는 학습 메커니즘을 제안

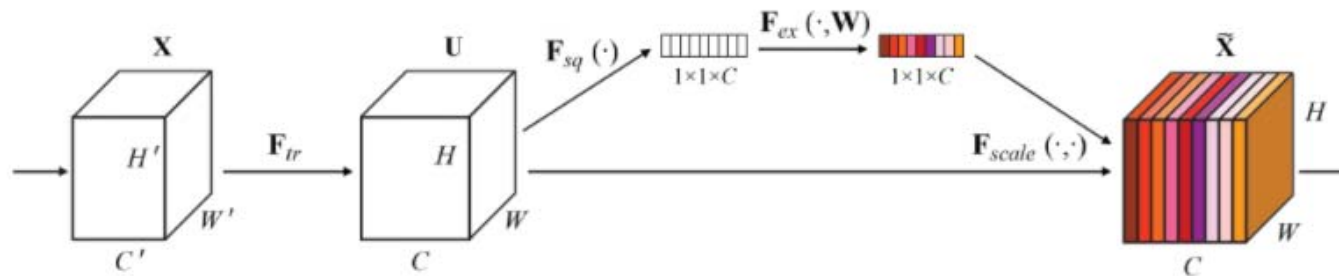


Fig.1

A Squeeze-and-Excitation block.

Squeeze-and-Excitation Networks (1)

Squeeze-and-Excitation Blocks

1. 입력 특징 맵으로부터 Squeeze 연산을 통해, 각 채널 별로 공간 차원($H \times W$) 특징맵을 모아서 표현한다.
2. 모아진 특징맵은 비선형 Excitation 연산을 통해, 채널 단위 특징들 간에 상관 관계를 파악한다.
3. 채널 수가 동일한 입력 특징 맵에 곱함 (Scale) 으로써, 중요한 특징에 대해 강조한다.

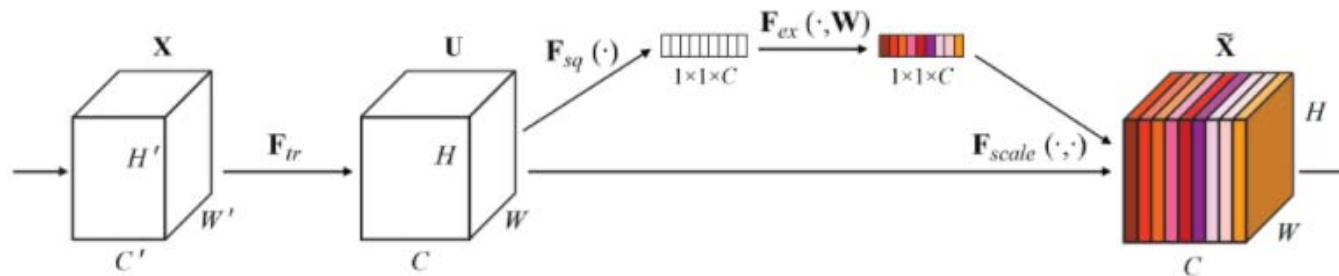


Fig.1

A Squeeze-and-Excitation block.

Squeeze-and-Excitation Networks (2)

Squeeze: Global Information Embedding

각각의 채널단위로 전체적인 정보를 반영하기 위해 Channel Descriptor 를 만들어 표현하여 중요한 정보를 압축

본 논문에서는 Channel Descriptor Function 을 Global Average Pooling (GAP) 를 사용

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j)$$

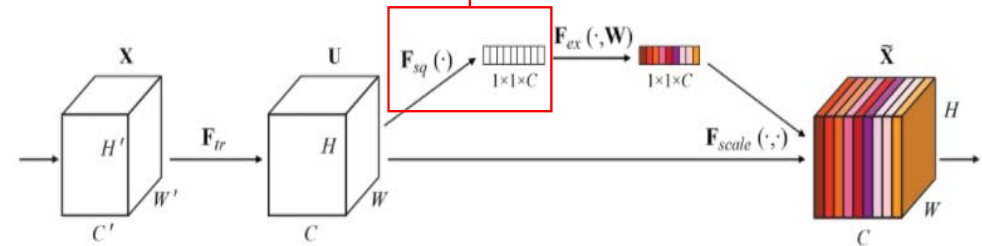


Fig.1

A Squeeze-and-Excitation block.

Squeeze-and-Excitation Networks (3)

Excitation: Adaptive Recalibration

중요한 정보들을 압축 (Squeeze)한 후, 채널 간 의존성 (Channel-wise dependencies)을 계산하여 Feature Recalibration 을 수행

FC Layer 와 비선형 활성화 함수를 사용하면서 중요한 특징 부분을 강조하면서 재조정

$$s = F_{ex}(z, w) = \sigma(W_2 \delta(W_1 z))$$

σ : Sigmoid Function
 δ : ReLU Function

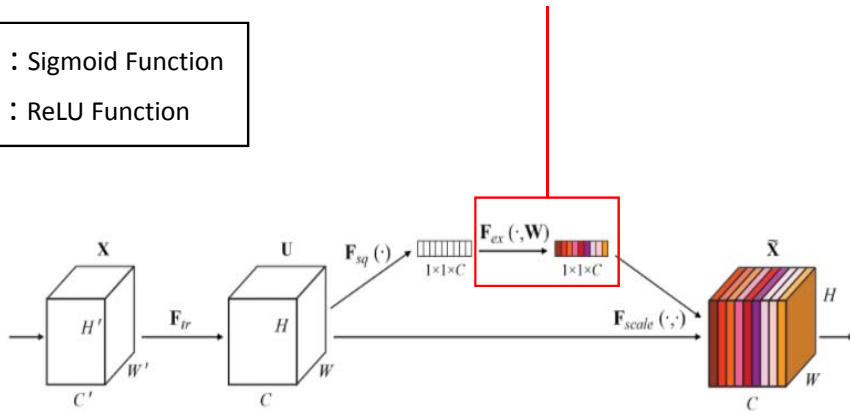
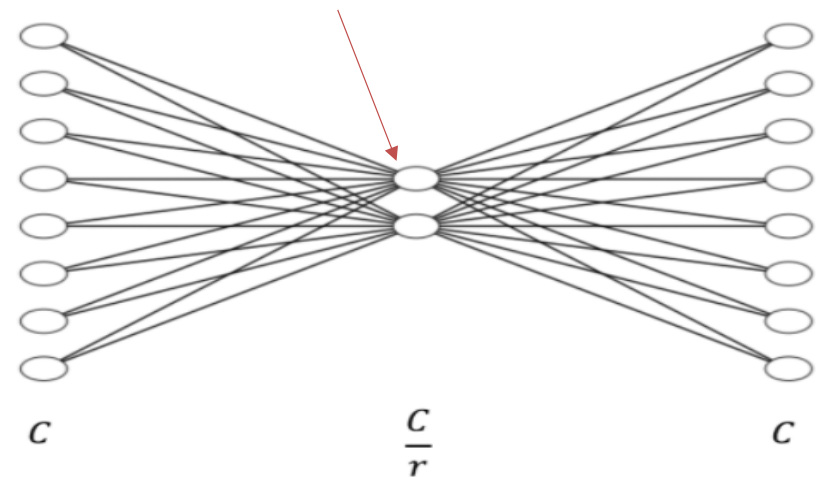


Fig.1
A Squeeze-and-Excitation block.

$$W_1 \in \mathbb{R}^{\frac{C}{r} \times C} \quad W_2 \in \mathbb{R}^{C \times \frac{C}{r}} \quad r : \text{reduction ratio}$$

FC layer 파라메타 수를 줄이기 위해



Squeeze-and-Excitation Networks (4)

Model Flexibility: SE-Inception and SE-ResNet

SE block을 보다 복잡한 아키텍처 (Inception, ResNet)에 통합하여 SENet 을 구성

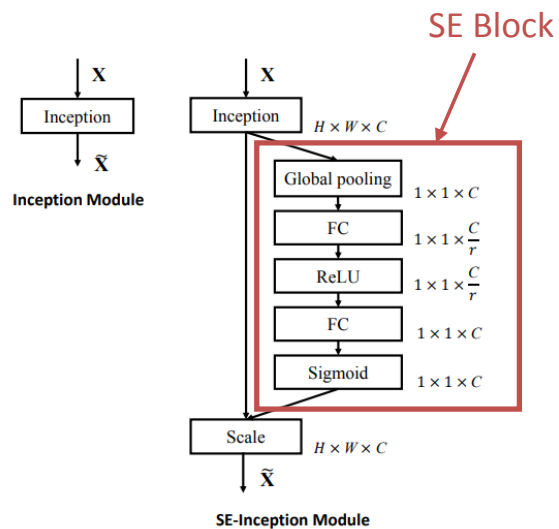


Figure 2: The schema of the original Inception module (left) and the SE-Inception module (right).

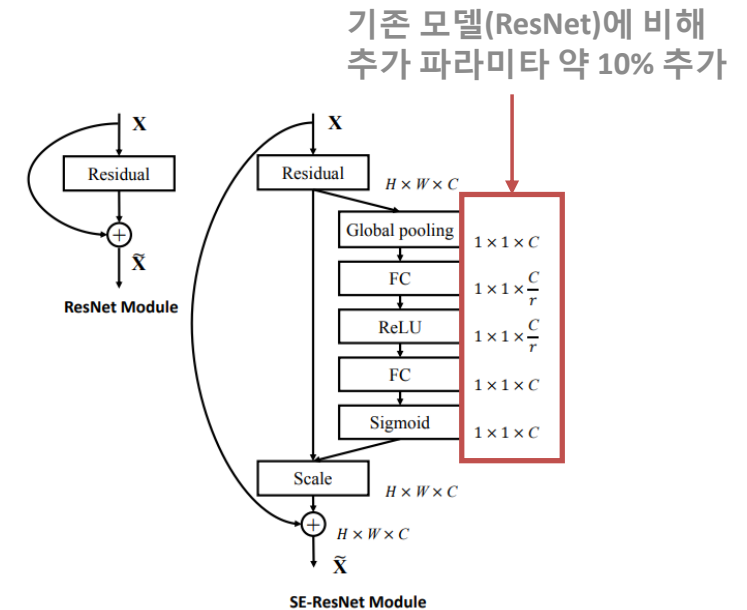


Figure 3: The schema of the original Residual module (left) and the SE-ResNet module (right).

DATASET CIFAR-10

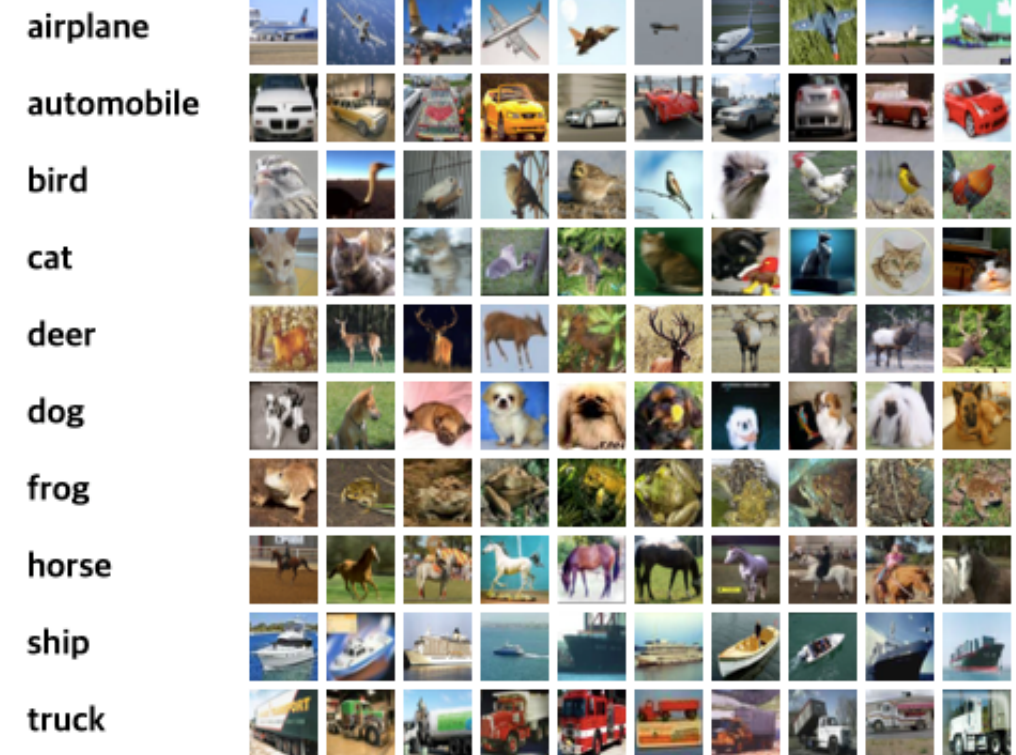
32x32사이즈의 RGB 이미지와
10가지의 레이블로 구성 (32, 32, 3), (10)

INPUT DATA : 50000

TEST DATA : 10000

* ResNet-20는 K-fold Cross Validation 사용

Here are the classes in the dataset, as well as 10 random images from each:



실험 결과

정밀도 (Precision): $\frac{TP}{TP+FP}$

재현율 (Recall): $\frac{TP}{TP+FN}$

정확도 (Accuracy): $\frac{TP+TN}{TP+FP+FN+TN}$

Name	# params	precision	recall	accuracy
ResNet-20	0.27M	91.69%	91.62%	91.63%
ResNet-32	0.46M	92.21%	92.18%	92.19%
ResNet-44	0.66M	91.57%	91.52%	91.53%
ResNet-56	0.85M	92.46%	92.48%	92.48%
ResNet-110	1.7M	93.19%	93.17%	93.17%
SE-ResNet32	0.48M	93.12%	93.12%	93.12%
SE-ResNet110	1.78M	93.76%	93.77%	93.77%

