Chapter 01. Image Classification 모델 구현 및 성능 개선하기

Image Classification

Copyright **FASTCAMPUS** Corp. All Rights Reserved

For Better Performance(Single Model)

- Better Network Architecture
  - Use better model – ex) ImageNet SOTA
  - More layers
  - More channels
  - Bigger resolutions
  - Use better activation function – ex) swish, etc.
  - Use additional architecture – ex) skip connection, SE-module, etc.
- Training Tricks!

Copyright FASTCAMPUS Corp. All Rights Reserved

Copyright FASTCAMPUS Corp. All Rights Reserved

Bag of Tricks for Image Classification

Bag of Tricks for Image Classification with Convolutional Neural Networks

Tong He Zhi Zhang Hang Zhang Zhongyue Zhang Junyuan Xie Mu Li

Amazon Web Services
{htong,zhiz,hzaws,zhongyue,junyuanx,mli}@amazon.com

2 [cs.CV] 5 Dec 2018

Abstract

Much of the recent progress made in image classification research can be credited to training procedure refinements, such as changes in data augmentations and optimization methods. In the literature, however, most refinements are either briefly mentioned as implementation details or only visible in source code. In this paper, we will examine a collection of such refinements and empirically evaluate their impact on the final model accuracy through ablation study. We will show that, by combining these refinements together, we are able to improve various CNN models significantly. For example, we raise ResNet-50's top-1 validation accuracy

Model	FLOPs	top-1	top-5
ResNet-50 [9]	3.9 G	75.3	92.2
ResNeXt-50 [27]	4.2 G	77.8	-
SE-ResNet-50 [12]	3.9 G	76.71	93.38
SE-ResNeXt-50 [12]	4.3 G	78.90	94.51
DenseNet-201 [13]	4.3 G	77.42	93.66
ResNet-50 + tricks (ours)	4.3 G	**79.29**	**94.63**

Table 1: **Computational costs and validation accuracy of various models.** ResNet, trained with our "tricks", is able to outperform newer and improved architectures trained with standard pipeline.

Copyright FASTCAMPUS Corp. All Rights Reserved

fast campus

Copyright FASTCAMPUS Corp. All Rights Reserved

Bag of Tricks for Image Classification

- Efficient Training
  - Linear scaling learning rate
  - Learning rate warm up
  - Zero gamma in batchnorm
  - No bias decay
  - Low-precision training

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1...m}\}$;
Parameters to be learned: γ, β
Output: $\{y_i = \mathrm{BN}_{\gamma,\beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{// mini-batch mean}$$

$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_{\mathcal{B}})^2 \qquad \text{// mini-batch variance}$$

$$\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}} \qquad \text{// normalize}$$

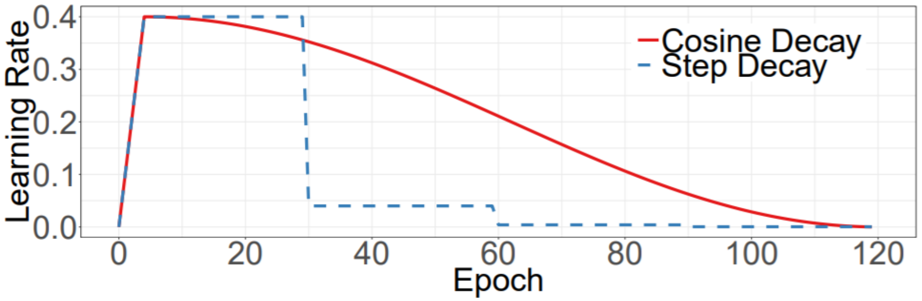
$$y_i \leftarrow \gamma\widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i) \qquad \text{// scale and shift}$$

Copyright FASTCAMPUS Corp. All Rights Reserved

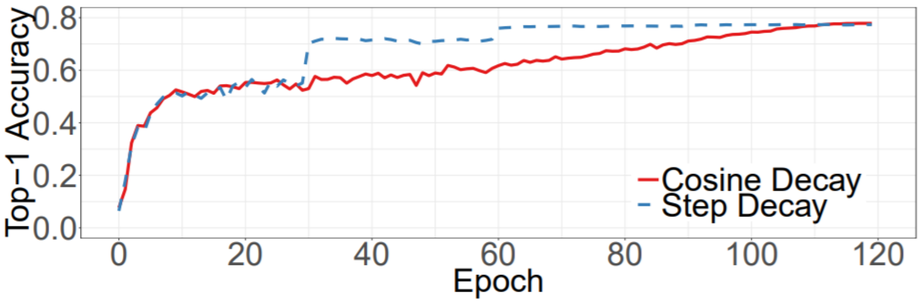
Copyright FASTCAMPUS Corp. All Rights Reserved

Bag of Tricks for Image Classification

- Training Refinements
  - Cosine Learning Rate Decay



(a) Learning Rate Schedule



(b) Validation Accuracy

Copyright FASTCAMPUS Corp. All Rights Reserved

Copyright **FASTCAMPUS** Corp. All Rights Reserved

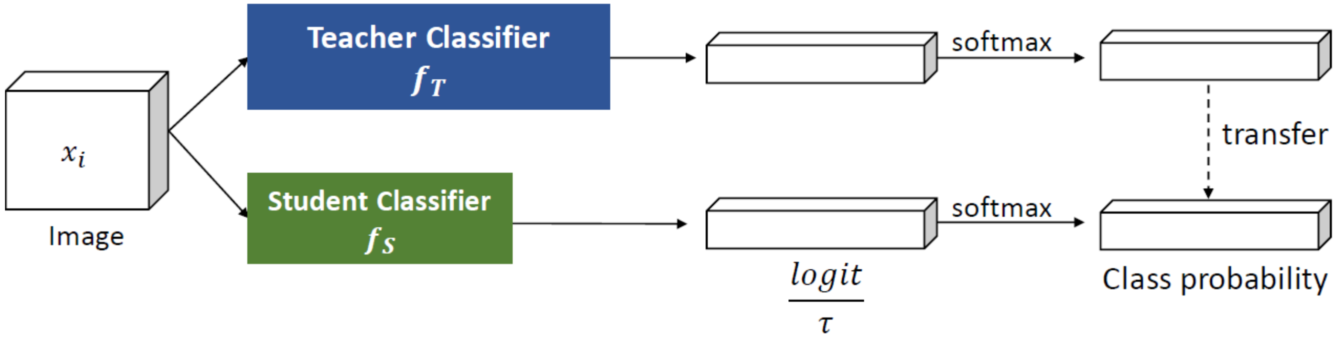
Bag of Tricks for Image Classification

- Training Refinements
  - Label Smoothing

$$q_i = \begin{cases} 1-\varepsilon & \text{if } i = y, \\ \varepsilon/(K-1) & \text{otherwise,} \end{cases}$$

  - Knowledge Distillation

Objective: $\sum_{x_i \in \mathcal{X}} \mathrm{KL}\Big(\mathrm{softmax}\big(\frac{f_T(x_i)}{\tau}\big), \mathrm{softmax}\big(\frac{f_S(x_i)}{\tau}\big)\Big)$



Copyright FASTCAMPUS Corp. All Rights Reserved

Copyright FASTCAMPUS Corp. All Rights Reserved

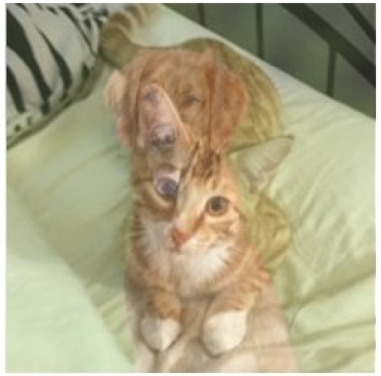
Bag of Tricks for Image Classification

- Training Refinements
  - Mixup Training

$$
\begin{aligned}
\hat{x} &= \lambda x_i + (1-\lambda)x_j, \\
\hat{y} &= \lambda y_i + (1-\lambda)y_j,
\end{aligned}
$$

where $\lambda \in [0,1]$ is a random number

Image



Label

[1.0, 0.0]
cat dog

[0.0, 1.0]
cat dog

[0.7, 0.3]
cat dog

Copyright FASTCAMPUS Corp. All Rights Reserved

Copyright FASTCAMPUS Corp. All Rights Reserved

Bag of Tricks for Image Classification

- Cutmix

	ResNet-50	Mixup [48]	Cutout [3]	CutMix
Image				
Label	Dog 1.0	Dog 0.5 Cat 0.5	Dog 1.0	Dog 0.6 Cat 0.4
ImageNet Cls (%)	76.3 (+0.0)	77.4 (+1.1)	77.1 (+0.8)	**78.6** **(+2.3)**
ImageNet Loc (%)	46.3 (+0.0)	45.8 (-0.5)	46.7 (+0.4)	**47.3** **(+1.0)**
Pascal VOC Det (mAP)	75.6 (+0.0)	73.9 (-1.7)	75.1 (-0.5)	**76.7** **(+1.1)**

$$\tilde{x} = \mathbf{M} \odot x_A + (\mathbf{1} - \mathbf{M}) \odot x_B$$
$$\tilde{y} = \lambda y_A + (1 - \lambda) y_B,$$

$$r_x \sim \text{Unif}\ (0, W),\quad r_w = W\sqrt{1-\lambda},$$
$$r_y \sim \text{Unif}\ (0, H),\quad r_h = H\sqrt{1-\lambda}$$

Copyright FASTCAMPUS Corp. All Rights Reserved

Copyright FASTCAMPUS Corp. All Rights Reserved