DEEP NETWORKS WITH STOCHASTIC DEPTH

Gao Huang*, Yu Sun*, Zhuang Liu, Daniel Sedra, Kilian Weinberger

*Authors contribute equally

Cornell University
RESIDUAL NETWORKS*

*Deep Residual Learning for Image Recognition, CVPR’16
by Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun*
RESIDUAL NETWORKS

*Deep Residual Learning for Image Recognition, CVPR’16 by Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun
RESIDUAL NETWORKS*

identity connection

*Deep Residual Learning for Image Recognition, CVPR’16 by Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun
STOCHASTIC DEPTH
STOCHASTIC DEPTH
BENEFITS OF STOCHASTIC DEPTH

Train short networks, get deep networks

Implicit ensemble of $2^L$ models

Speedup 25% during training
CIFAR-10 & CIFAR-100 (110-LAYER)

ResNet on CIFAR-10

Test error (%) vs epoch

ResNet 110-layer

6.41%

ResNet on CIFAR-100

Test error (%) vs epoch

ResNet 110-layer

27.88%
CIFAR-10 & CIFAR-100 (110-LAYER)

ResNet on CIFAR-10

- ResNet 110-layer
- ResNet 110-layer w/ Stochastic Depth

18% error reduction

epoch

18% error reduction

ResNet on CIFAR-100

- ResNet 110-layer
- ResNet 110-layer w/ Stochastic Depth

10% error reduction

epoch
ResNet on CIFAR-10

- Red line: ResNet 110-layer
- Blue line: ResNet 110-layer w/ Stochastic Depth

Test error (%): 6.41% for 110-layer, 5.25% for 110-layer with Stochastic Depth
ResNet on CIFAR-10

- ResNet 110-layer
- ResNet 110-layer w/ Stochastic Depth
- ResNet 1202-layer w/ Stochastic Depth

Test error (%)

Epoch

4.91% 6.41%

5.25%
Come to Poster S-3A-08

More details!
More analysis!
More data sets!

Code: github.com/yueatsprograms/Stochastic_Depth
Email: ys646@cornell.edu