Studying the Effect of Skip Connection

Yiran Wu Sihao Ying

Shanghai Jiao Tong University

December 25, 2017

Our Experiment Last Time

- Dataset: EMNIST (28 by 28 grayscale handwritten letters)
- input -> fc-128 -> fc-256 (-> fc-256) -> fc-26

	train(3-layer)	val(3-layer)	train(4-layer)	val(4-layer)
skip	91.0	89.0	91.6	89.1
vanilla	90.4	88.6	92.0	90.0
skip*	91.2	89.4	92.5	90.9

Table: Experimental results

Exp 1

- Dataset: CIFAR-10 (32 by 32 RGB images, 10 classes)
- input -> fc-512 -> fc-256 (-> fc-256) -> fc-10

	train(3-layer)	val(3-layer)	train(4-layer)	val(4-layer)
skip	64.9	64.1	67.8	66.6
vanilla	64.0	63.2	66.5	65.8
skip*	64.9	64.2	68.0	66.7

Table: Experimental results

Exp 2

- Dataset: CIFAR-10 (32 by 32 RGB images, 10 classes)
- input
  -> conv(32 * [3, 3], ReLU) -> conv(32 * [3, 3], ReLU)
  -> maxpool([2, 2])
  -> conv(64 * [3, 3], ReLU) -> conv(64 * [3, 3], ReLU)
  -> maxpool([2, 2])
  -> conv(128 * [3, 3], ReLU) -> conv(128 * [3, 3], ReLU)
  -> maxpool([2, 2])
  -> fc-10

Exp 2

	train	val
skip	84.2	83.5
vanilla	81.7	80.9

Table: Experimental results

- Unfairness?

	train	val
vanilla*	84.9	84.1

Table: Experimental results

Thanks!