Transfer Learning

Use Pre-trained Model

Ming-Jia Huang, ICL Div. X, ITRI

Overview of Transfer Learning

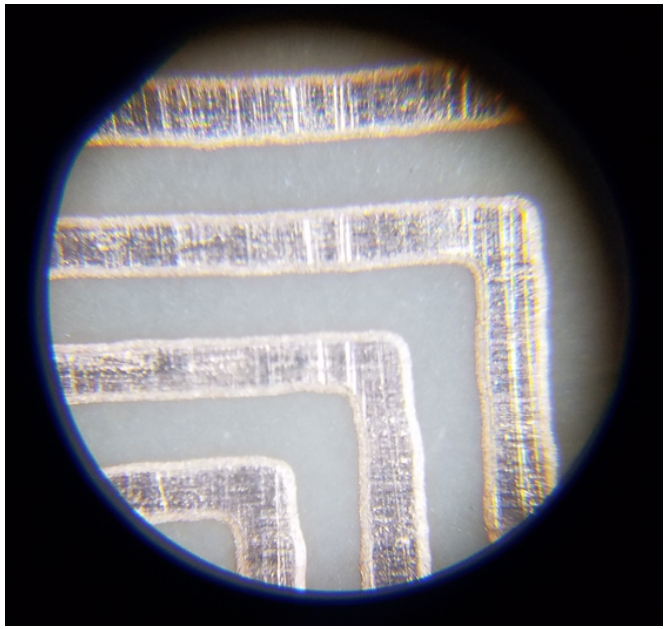
- Very deep network are expensive to train, in practice a very few people train a convolution network from scratch

- When we use it
  - You don't have enough labeled dataset
  - Problem similar and already existing pre-trained model

- Advantage
  - Get higher accuracy
  - Reduce the training time
  - Good starting point before training

The Use Case in Real World

- PCB defect recognition
- 5 Classes: 正常, 皮屑, 髒污, 殘膠, 斷路短路
- Use pre-trained model which trained from ImageNet
- 2K images as training set
- Accuracy
  - From scratch: 91.4%
  - Transfer learning: 96.6%



10x10 PCB, Corner Runs

LAB03-2

- Dataset is sampling from Food-11/training
- Testing dataset still use from Food-11/evaluation
- ResNet-18 as network
- Use pre-trained model & don't use pre-trained model
- Hyper parameters are free to setup, but its must be consistence between with & without pre-trained model
  - Except the learning rate, you may consider decay the learning rate when use pre-trained model
- Show the summary report with what things be observed during the experiment
  - Validation accuracy?

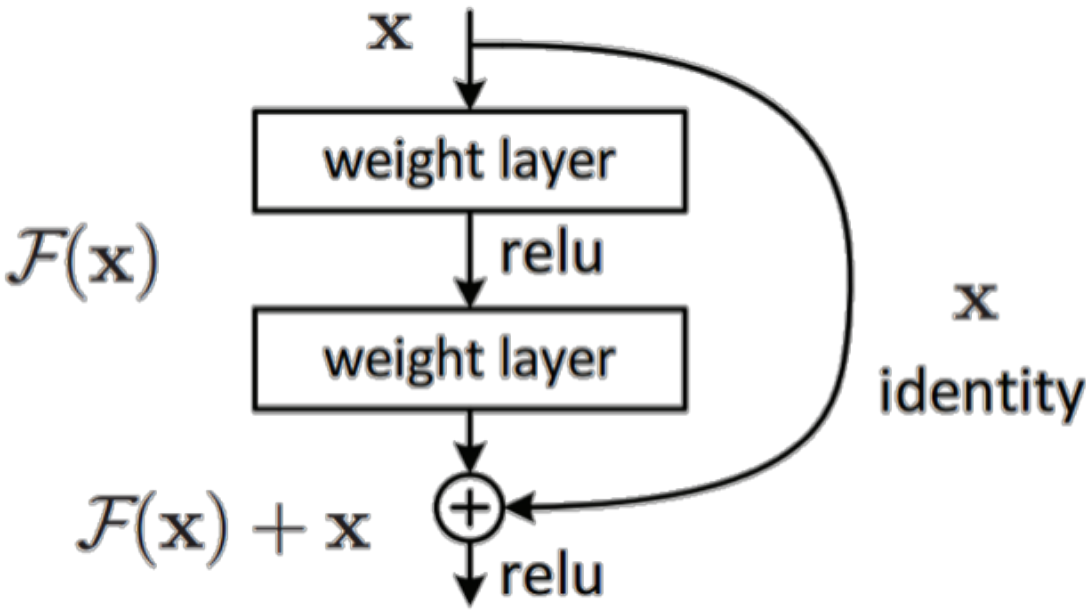
ResNet-18

- Paper
  - Deep Residual Learning for Image Recognition

Layer Name	Output Size	ResNet-18
conv1	$112 \times 112 \times 64$	7×7, 64, stride 2
conv2_x	$56 \times 56 \times 64$	3×3 max pool, stride 2
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
conv3_x	$28 \times 28 \times 128$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
conv4_x	$14 \times 14 \times 256$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
conv5_x	$7 \times 7 \times 512$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
average pool	$1 \times 1 \times 512$	7×7 average pool
fully connected	1000	512×1000 fully connections
softmax	1000	

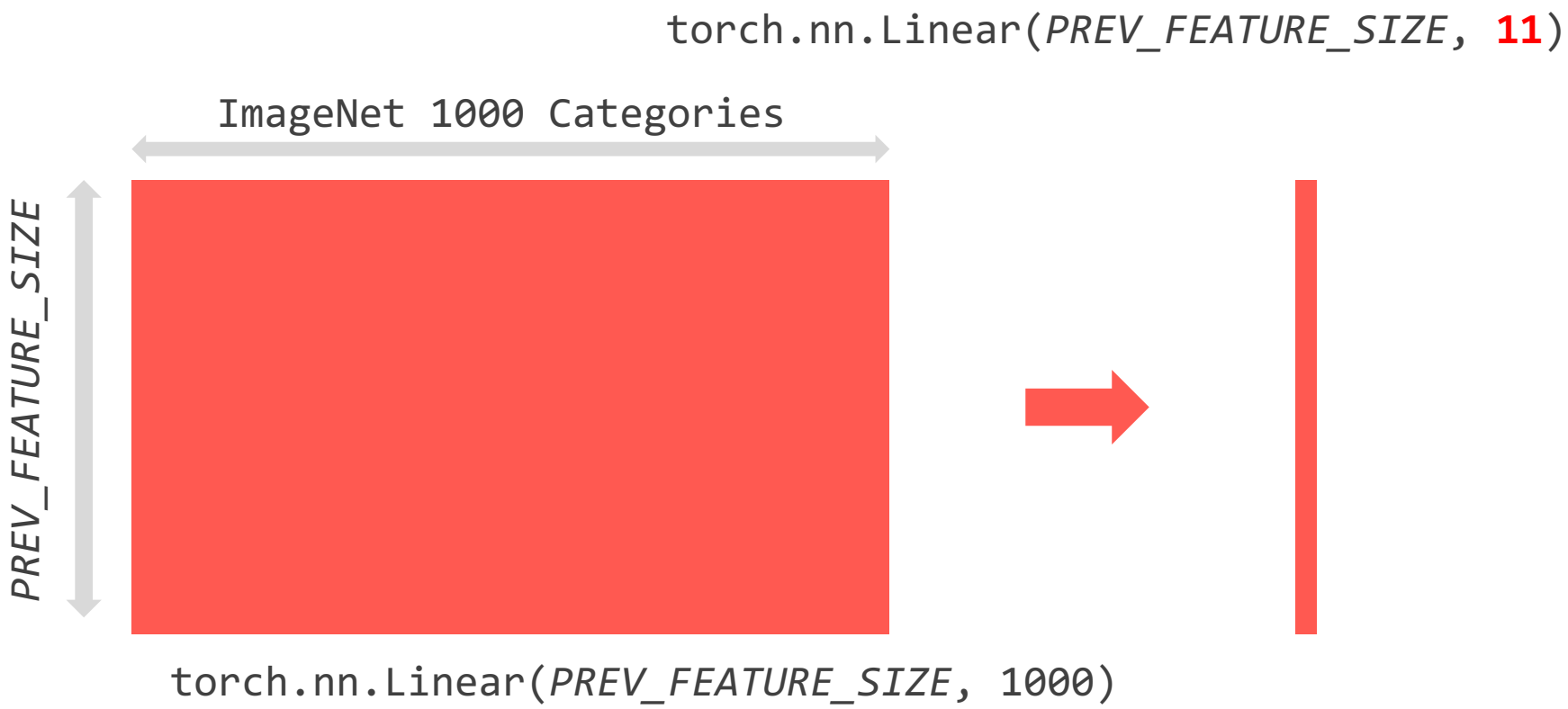
Building Block in ResNet

- A residual block—the fundamental building block of residual networks



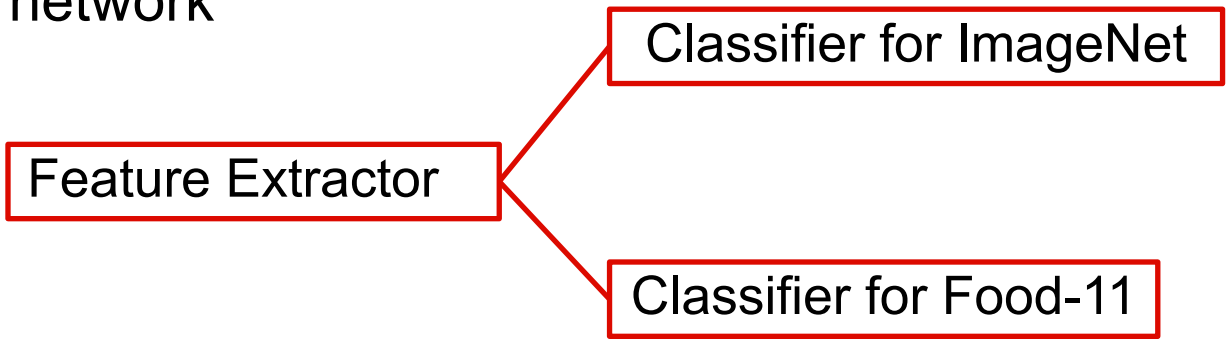
Note

- Reset final fully-connected layer
  - Original model was designed for ImageNet (1,000 categories)
  - Our dataset Food-11 which only has 11 categories



BOOOOOONUS

- It's skippable
- A CNN is composed of two basic parts of feature extraction and classification
- Try to freeze it if you want to keep the original weights
- ResNet-50 as network



The Goal

- Just for fun
- Try to get higher accuracy in the same network:
  - ImageNet testing dataset (provided by this LAB)
  - Food-11 evaluation dataset