Machine Learning Engineer Course
Day 27

- ResNet and VGG -

DIVE INTO CODE

Thursday October 21, 2021

DIOP Mouhamed

Agenda

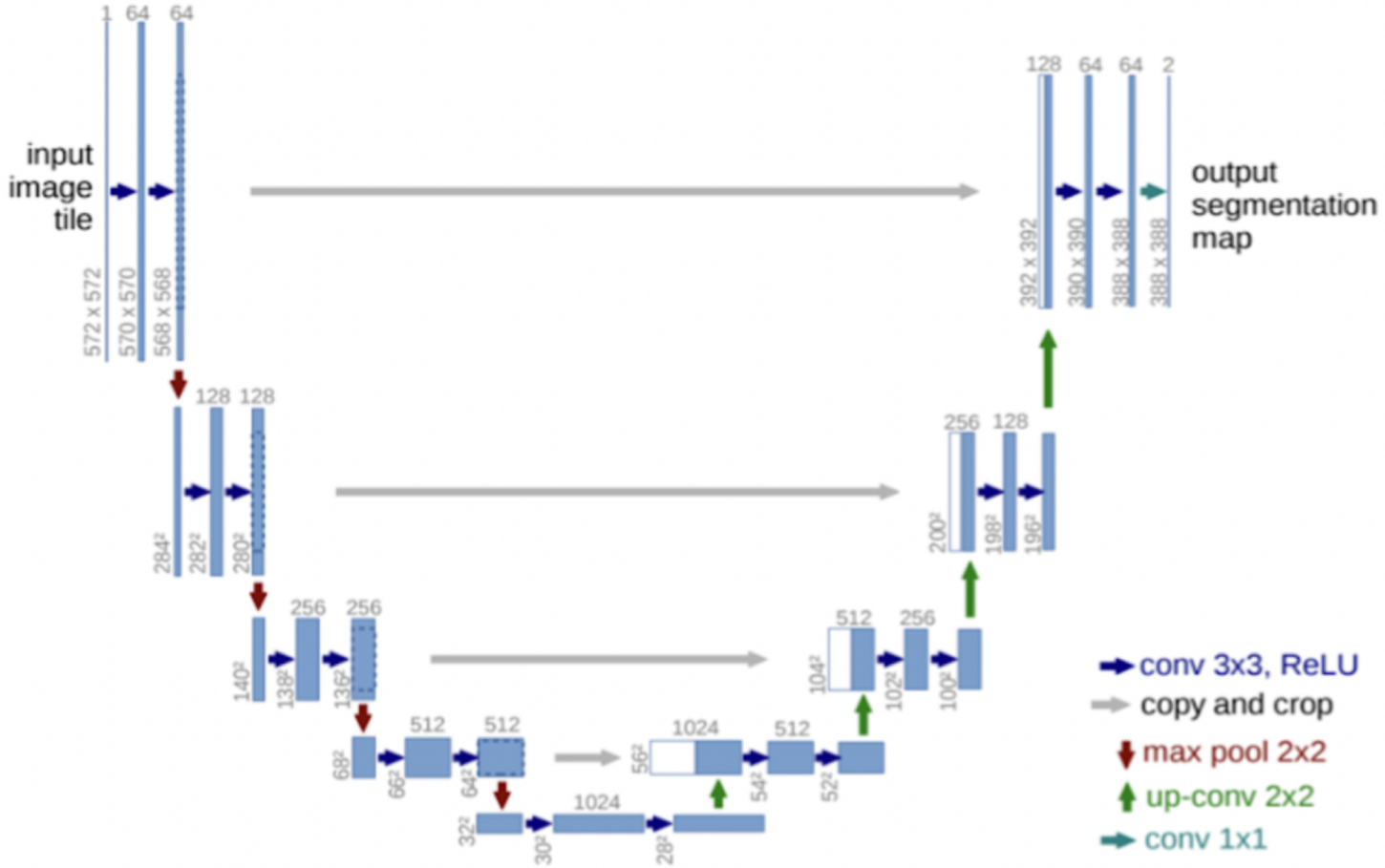
1. Check-in
2. Quick Review
3. ILSVRC
4. ResNet
5. Transfer Learning
6. Sample code
7. To do by next class
8. Check-out

Check-in

3 minutes Please post the following point to Zoom chat.

Q. What did you learn in the previous week?
(Anything is fine.)

Quick Review (U-Net)

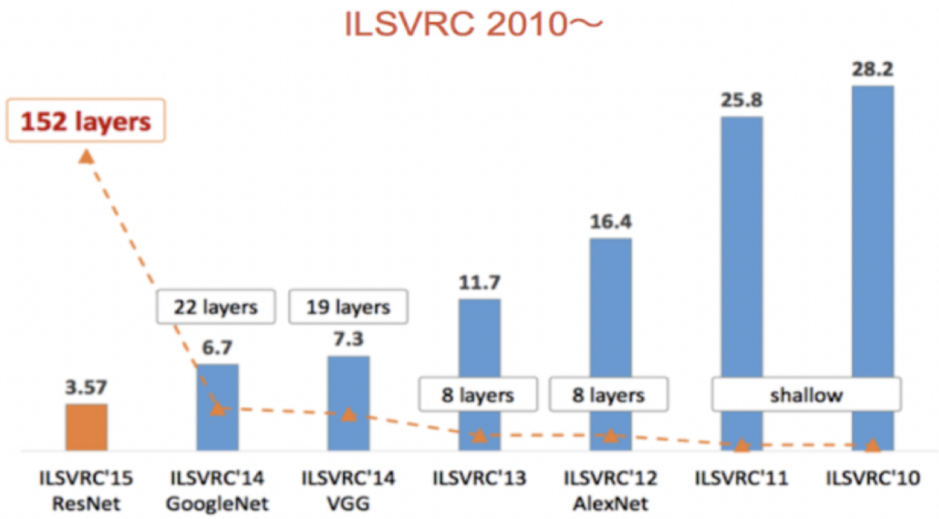


Large-scale image recognition competition

Image and Video Recognition Competition using Benchmark Datasets

When submitting a research paper on an image recognition or machine learning method, you are required to compare it with other methods on a standard benchmark dataset and show its quantitative superiority.

ILSVRC (Large-scale image recognition using ImageNet)
http://image-net.org/challenges/LSVRC/2017/
COCO challenge (object region segmentation and image description generation using Microsoft COCO)
http://cocodataset.org/#home
Places chalenge (Large-scale scene recognition using MIT Places)
http://places2.csail.mit.edu/demo.html
VQA challenge (Video Question and Answer)
http://vqa.cloudcv.org/
The First LVIS Challenge
(Instance segmentation using LVIS)
http://cocodataset.org/workshop/coco-mapillary-iccv-2019.html#lvis-challenge



ILSVRC

ImageNet Large-Scale Visual Recognition Challenge（ILSVRC）

What is the ILSVRC?

This is a large-scale image recognition competition that ran from 2010 to 2017, using some of the data from **ImageNet**. It was co-organized by several universities in the US, including Stanford University.

In 2012, Professor Hinton and his group at the University of Toronto made a huge impact on researchers around the world when they overwhelmingly won the 1,000-class discrimination test using an 8-layer convolutional neural network (CNN) with an error rate of more than 10% over the second-place team.

In 2015, Microsoft Research Asia achieved an error rate of 3.57% in 1,000-class identification using a 152-layer CNN. With a structure called **residual network (ResNet)**, which has a coupling that bypasses the convolutional layer, it is now possible to train even deeper models than before.

networks trained on ILSVRC data are now shared on open source, and **transfer learning**, which performs **fine-tuning** on other datasets, is now established as one of the cornerstones in the field. since 2017, ILSVRC has been taken over by kaggle.

Slide by Mr. Li Fei-Fei, creator of IImageNet
http://www.image-net.org/papers/ImageNet_2010.pdf

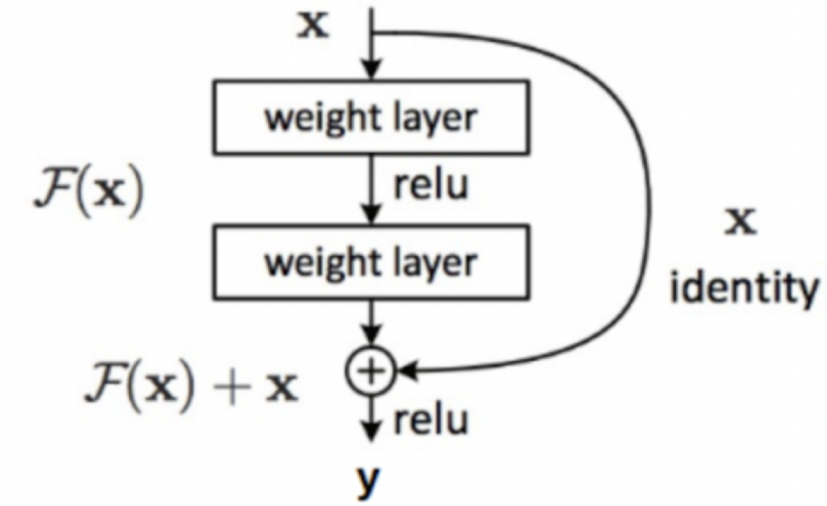
ResNet

Deep Residual Learning for Image Recognition (ResNet)（2015）

What is ResNet?

https://arxiv.org/pdf/1512.03385.pdf

At ILSVRC 2015, Kaiming He et al. proposed an architecture for training 152-layer NNs, which introduces "skip connections" and features a large amount of batch normalization.



$$y = x + F(x)$$

$$\frac{\delta E}{\delta x} = \frac{\delta E}{\delta y} * \frac{\delta y}{\delta x} = \frac{\delta E}{\delta y} * (1 + F'(x))$$

$$= \frac{\delta E}{\delta y} + \frac{\delta E}{\delta y} * F'(x)$$

Transfer learning

Transfer Learning

What is transfer learning?

https://medium.com/@14prakash/transfer-learning-using-keras-d804b2e04ef8

One of the research challenges in machine learning is to focus on knowledge that solves one problem and applies it to different but similar problems.

Specifically, the weights of a pre-trained network are used as initial values for the weights of another network.

A fixed feature extractor can be useful in solving many problems.

Transfer learning

Why do we do transfer learning?

Very deep networks are expensive to train.
A complex model requires hundreds of **expensive** GPU-equipped machines to train for weeks.

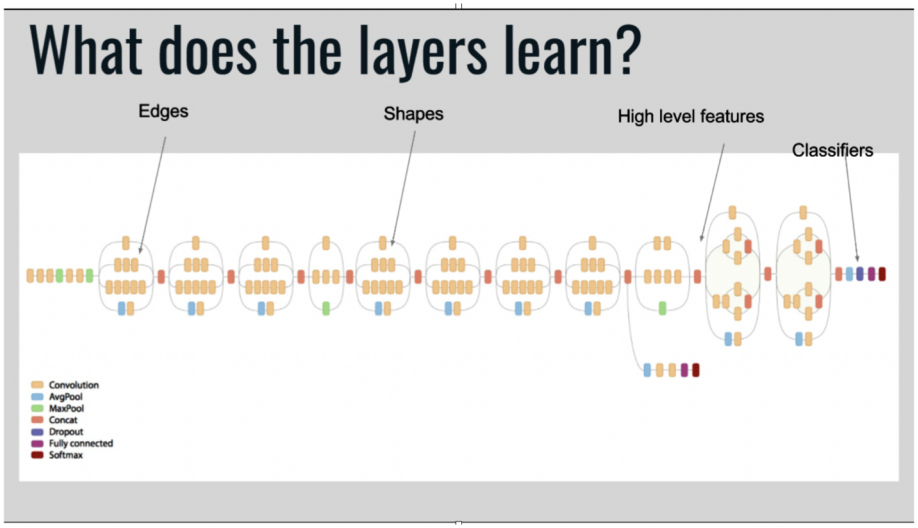
Does ResNet take several weeks to learn?

https://www.slideshare.net/iwiwi/nips17-86470238
https://blog.nnabla.org/ja/news/imagenet-in-224s/

The role of weights in transfer learning

What is the role of weights in transfer learning?

The **first layer** of the network tries to detect edges, the **middle layer** tries to detect shapes, and the **second layer** tries to detect advanced features.
In other words, the forward layer is more general and the backward layer is more specific to the original data set.
Trained networks with such detectors are generally useful for solving other computer vision problems.



Learned weights

Where are the pre-trained networks?

In the case of keras

Documents

https://keras.io/ja/applications/

GitHub：

https://github.com/keras-team/keras-applications/tree/master/keras_applications

Pre-trained classification models for Image Net images provided in Keras Applications.

- Xception
- VGG16
- VGG19
- ResNet50
- InceptionV3
- InceptionResNetV2
- MobileNet
- DenseNet
- NASNet
- MobileNetV2

Transfer learning

How do you handle cases where the original data and the new data are different?

The following cases are possible.

http://cs231n.github.io/transfer-learning/

①The new data set is small and similar to the original data set

```
# Train only the last classifier (FC layers)

for layer in model.layers:
  layer.trainable = False
```

②The new data set is large and similar to the original data set

```
# Train all networks. Also True by default.

for layer in model.layers:
  layer.trainable = True

# If you want the first half weight (5 layers) to detect the edge, you can write

for layer in model.layers[:5]:
  layer.trainable = False
```

Transfer learning

How do you handle cases where the original data and the new data are different?

The following cases are possible.

http://cs231n.github.io/transfer-learning/

③The new data set is small but very different from the original data set

Patterns that are difficult to transfer-learn.
We want to train only the upper layer to prevent over-learning because of the small amount of data, but since we are learning using data that is not similar, we think that we cannot learn well using the features in the second half.
If we connect the classifier to a layer somewhere in the middle instead of the second half, it might work.

④The new data set is large and differs significantly from the original data set

Use a random initial value or use the weights of the pretrained model as initial values. In general, the latter is better.

Sample code

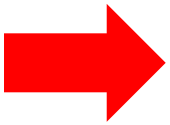
How to solve problems
”ResNet & VGG"

[Problem 1] Code review
[Problem 2] Code rewriting
[Problem 3] Learning and estimation

Sprint 20 – ResNet & VGG

Explanation about this Sprint is given but please try it on your own first.

Sprint 20 – ResNet & VGG



Please work on your own after class and submit your assignments on DIVER.

Sprint 20 – ResNet & VGG

A Sample Code of this Sprint is given but please try it on your own.

Sprint 20 – ResNet & VGG

Please work on your own after class and submit your assignments on DIVER.

ToDo by next class

Next class will be Zoom : Thursday October 28, 2021 19:30 ~ 20:30

🚨ToDo: Natural Language Processing
https://diver.diveintocode.jp/curriculums/2010

Check-out

3 minutes Please post the following point to Zoom chat.

Q. Current feelings and reflections
(joy, anger, sorrow, anticipation, nervousness, etc.)

Thank You For Your Attention