**Chapter 9.1 – 9.3, Chollet 2nd ed.**

Talk notes and links are posted on Github folder:

<https://github.com/JennEYoon/deep-learning/tree/main/chollet2/chp7_8_9>

Jennifer E. Yoon, April 11, 2023, DSML Meetup talk

[jenneyoon@gmail.com](mailto:jenneyoon@gmail.com)

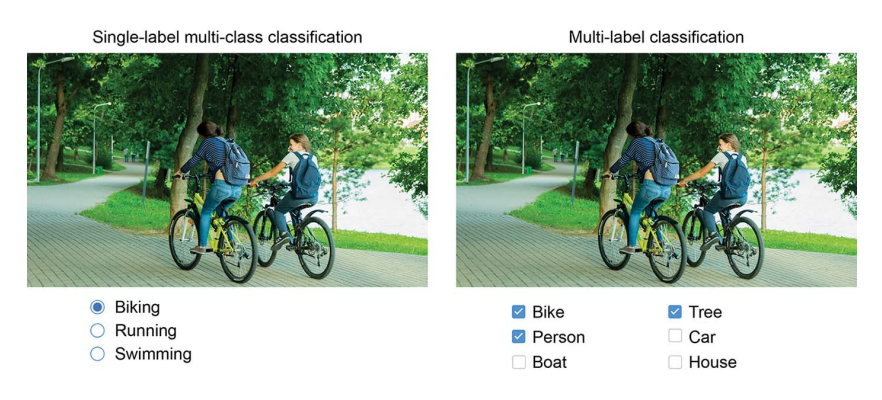
datasciY.com

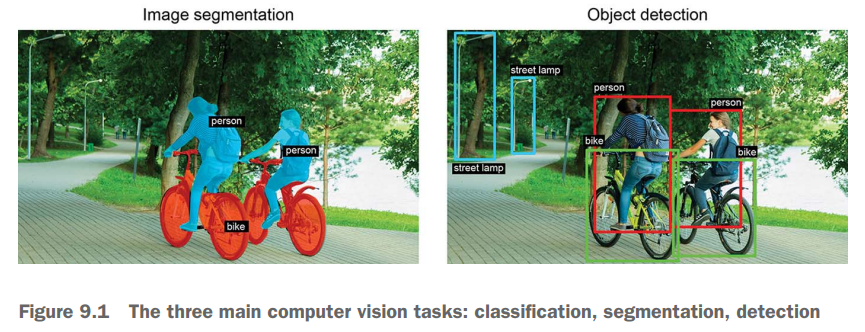
**Overview**

* **Image Classification** 
  + Chapter 8
  + Dog vs Cat, CIFAR10
* **Image Segmentation**
  + This Chapter
  + Pixel grouping by category
* **Object Detection** 
  + Book does not cover
  + See Coursera Class 4, Deep Learning Specialization, links at End page
  + YOLO model, You Only Look Once. Links at End page

**9.1 Three Computer Vision Tasks**

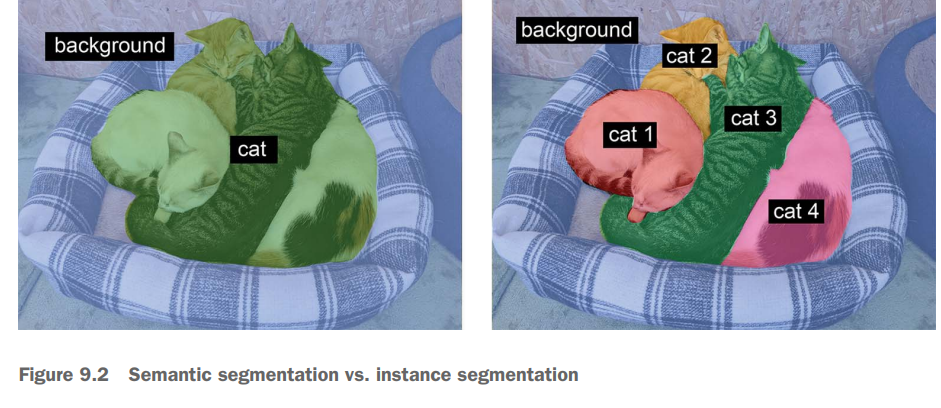
* Image classification—Where the goal is to assign one or more labels to an image. It may be either single-label classification (an image can only be in one category, excluding the others), or multi-label classification (tagging all categories that an image belongs to, as seen in figure 9.1). For example, when you search for a keyword on the Google Photos app, behind the scenes you’re querying a very large multilabel classification model—one with over 20,000 different classes, trained on millions of images.
* Image segmentation—Where the goal is to “segment” or “partition” an image into different areas, with each area usually representing a category (as seen in figure 9.1). For instance, when Zoom or Google Meet diplays a custom background behind you in a video call, it’s using an image segmentation model to tell your face apart from what’s behind it, at pixel precision.
* Object detection—Where the goal is to draw rectangles (called bounding boxes) around objects of interest in an image, and associate each rectangle with a class. A self-driving car could use an object-detection model to monitor cars, pedestrians, and signs in view of its cameras, for instance



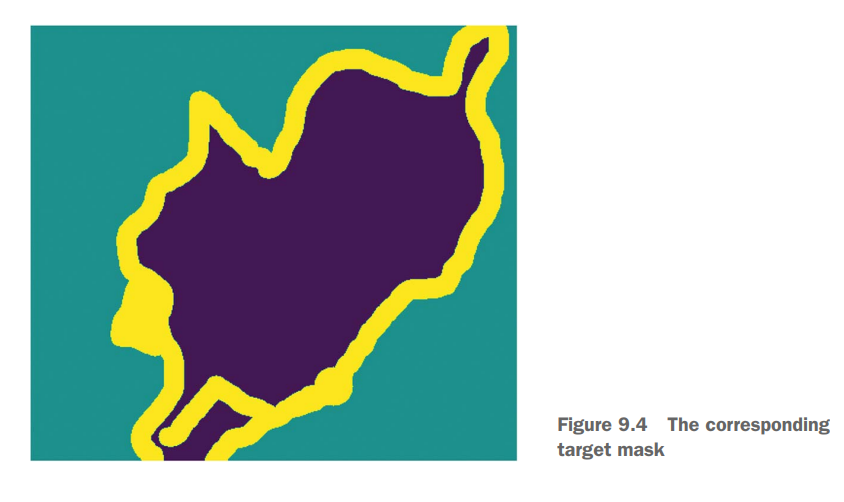


**9.2 Image Segmentation**

* 1 (foreground)
* 2 (background)
* 3 (contour)

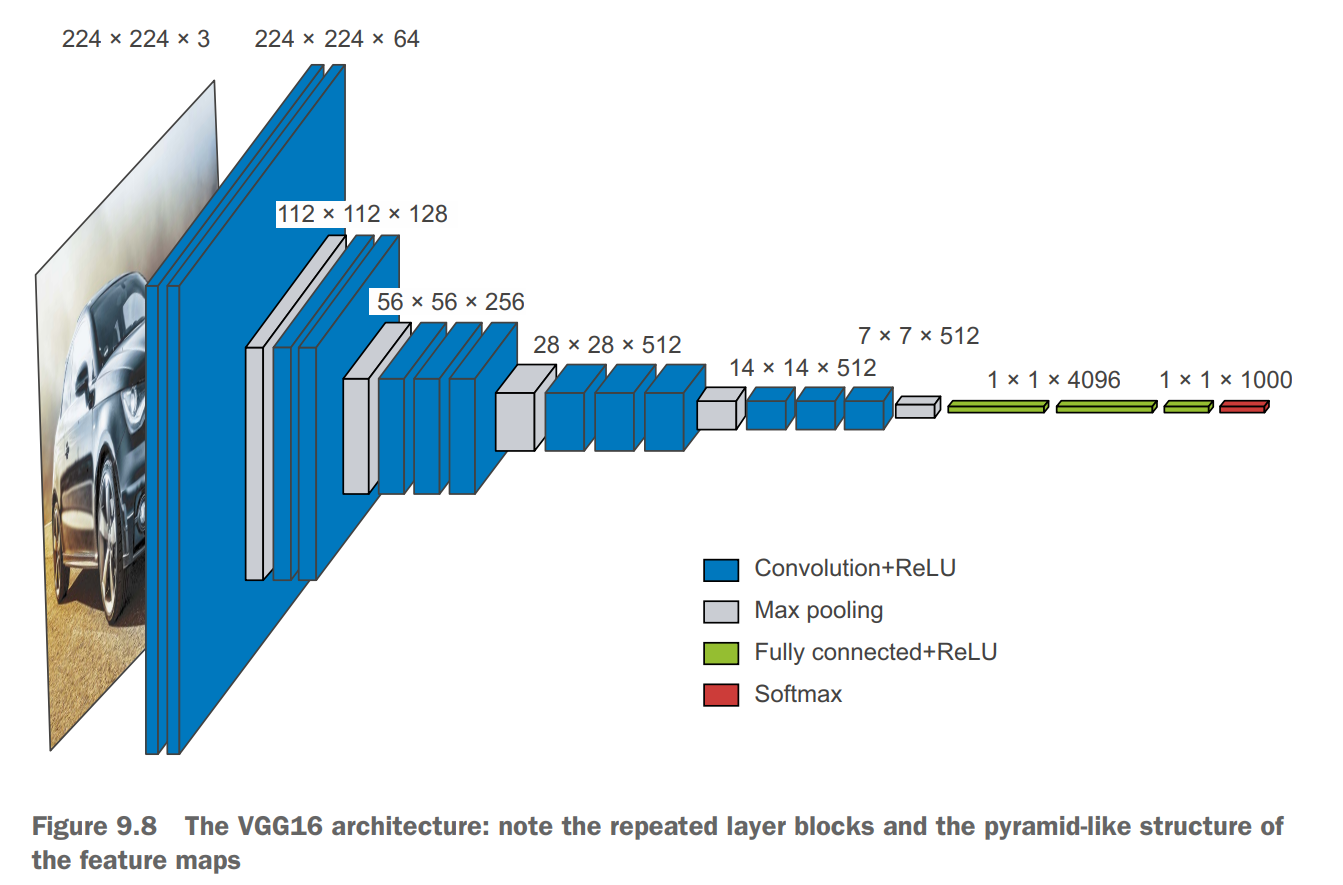
****

****

****

**9.3.0 Model Blocks Reuse (Lego Blocks)**

* Architecture image: figure 9.8, VGG16

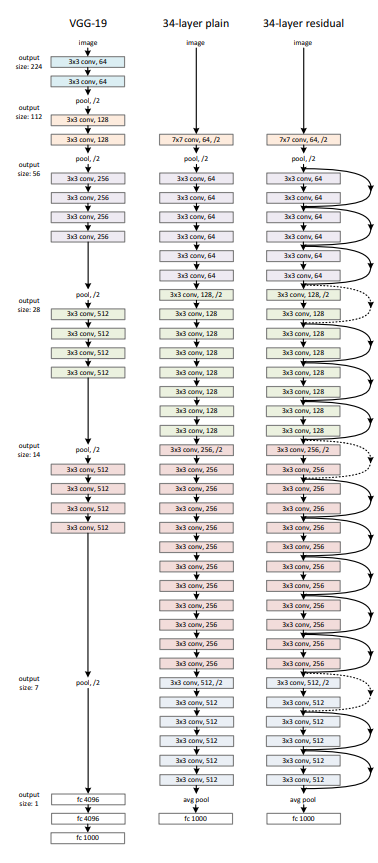


* **Importance of Ablation Studies**

If you become a deep learning researcher, cut through the noise in the research process: do ablation studies for your models. Always ask, “Could there be a simpler explanation? Is this added complexity really necessary? Why?”

* Resnet50 Image

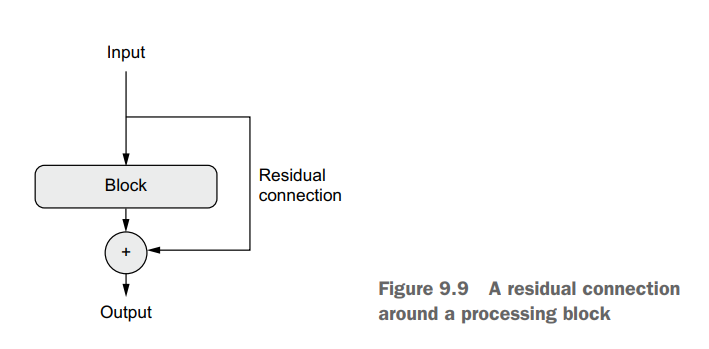
Deep Residual Learning for Image Recognition paper: https://arxiv.org/abs/1512.03385



* Other important models:
  + Transformers (Large Language Model, Vision Transformer)
  + Recurrent Neural Network (Alpha-Go)

**9.3.2 Residual Connection, Skip Connection**

* Figure 9.9 residual connection



* Adding weights with residual connection (skip connection)

X = add( [ x, residual ])

Ex: add(0.01, 0.45) 🡪 0.46

Ex: add(0.45, 0.45) 🡪 0.90

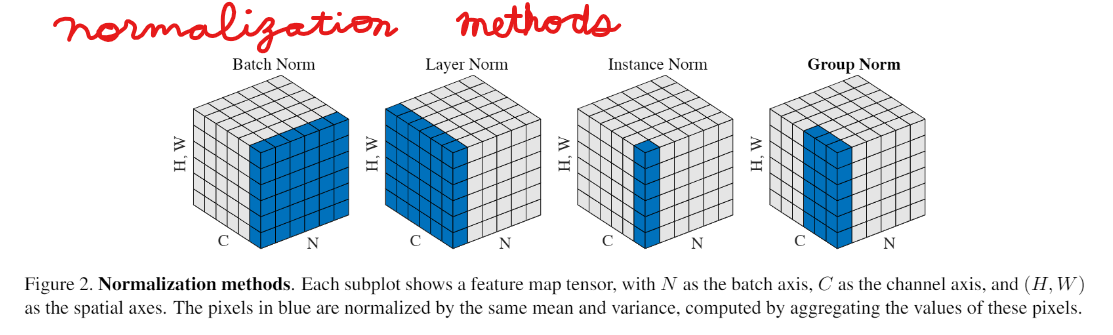
Weight results cannot be worse than not adding this block to the flow.

* ResNet Models (34, 50) frequently used as initial model.

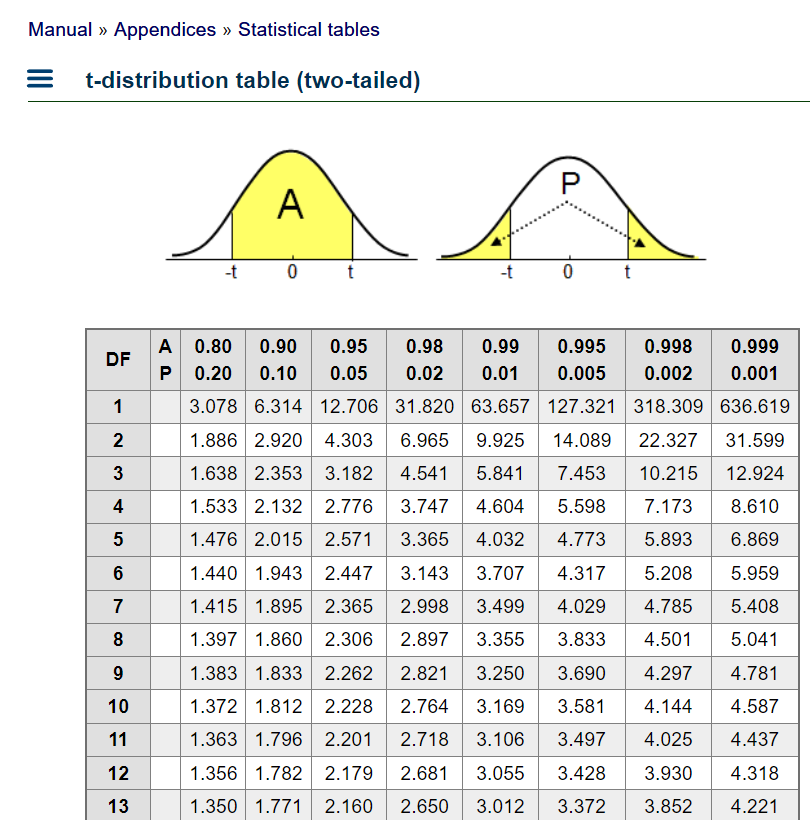
**Batch Norm**

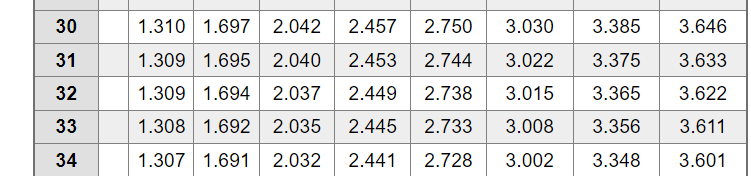
* Normalization Methods figure, from Group Nomalization, Yuxin Xu, Kaiming He, Facebook AI Research, 11 Jun 2018, arXiv: 1803.08494v3

<https://github.com/JennEYoon/deep-learning/blob/main/chollet2/chp7_8_9/1803.08494.pdf>

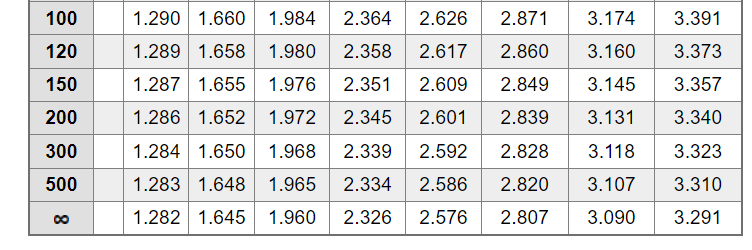


* Batch size need to be 34 or more. Maybe related to degrees of freedom, where 30+ observation samples are needed to reach a good approximation of a probability distribution. See Central Limit Theorem.









* **On batch normalization and fine-tuning**

Batch normalization has many quirks. One of the main ones relates to fine-tuning: when fine-tuning a model that includes BatchNormalization layers, I recommend leaving these layers frozen (set their trainable attribute to False). Otherwise they will keep updating their internal mean and variance, which can interfere with the very small updates applied to the surrounding Conv2D layers

**END Links:**

My github folder for today:

<https://github.com/JennEYoon/deep-learning/tree/main/chollet2/chp7_8_9>

Normalization methods paper: <https://github.com/JennEYoon/deep-learning/blob/main/chollet2/chp7_8_9/1803.08494.pdf>

Coursera Class 4:

Object Detection Youtube: <https://www.youtube.com/watch?v=SWzC6NrKfrM&list=PLkRLdi-c79HKEWoi4oryj-Cx-e47y_NcM&index=12>

YOLO paper: <https://github.com/JennEYoon/Coursera-DLAI/blob/main/source/Papers/YOLO.pdf>

Coursera notebook: <https://github.com/JennEYoon/Coursera-DLAI/blob/main/dlai-class4-CNN/week12/Autonomous_driving_application_Car_detection_v3a.pdf>