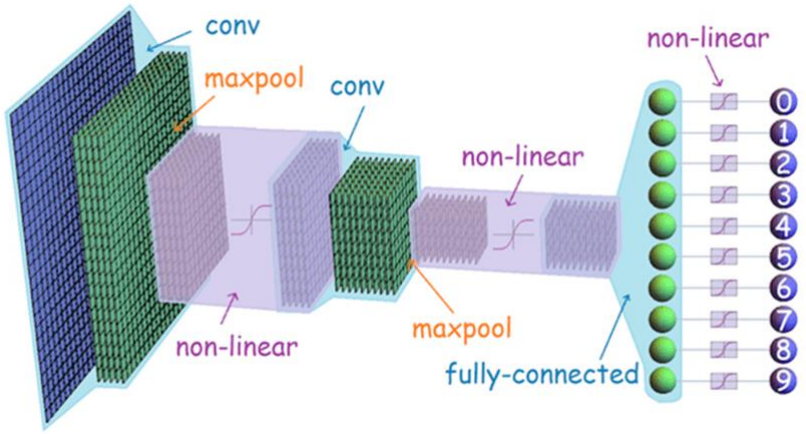
Deep Learning Workshop

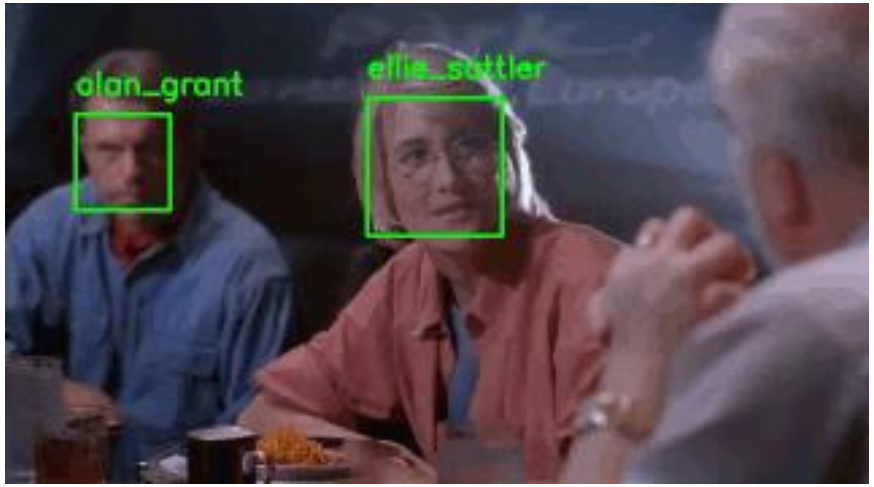
Convolutional Neural Networks

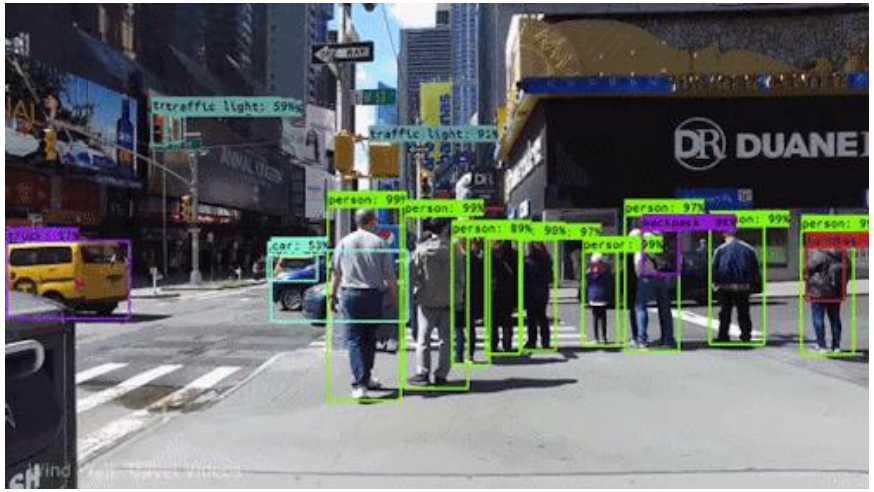


Instructor: Aaron Low

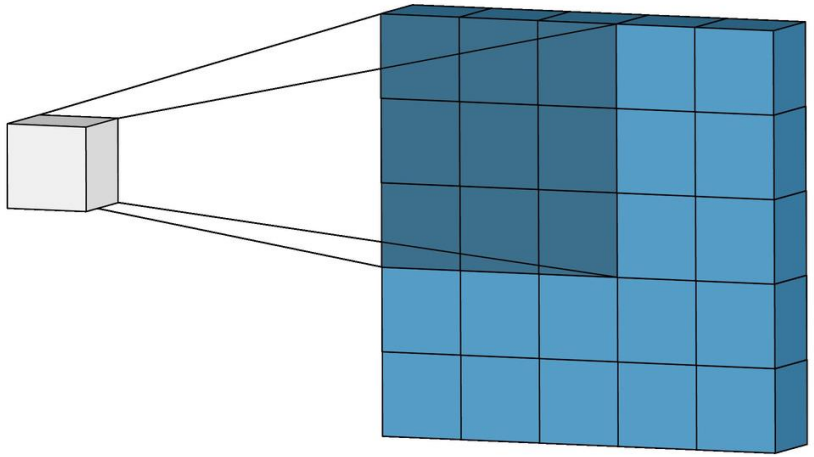
HELP University, Faculty of Computing and Digital Technology

Computer Vision



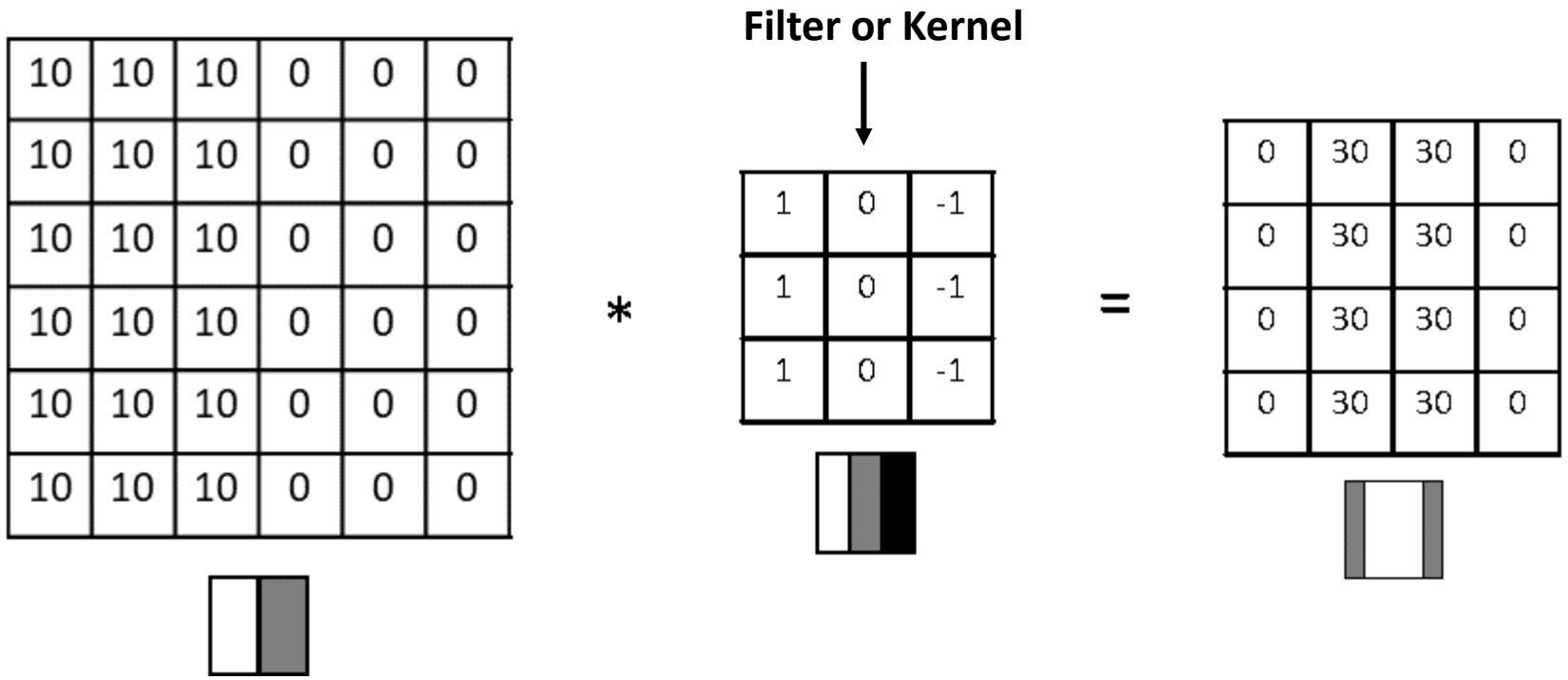


Convolution Operation



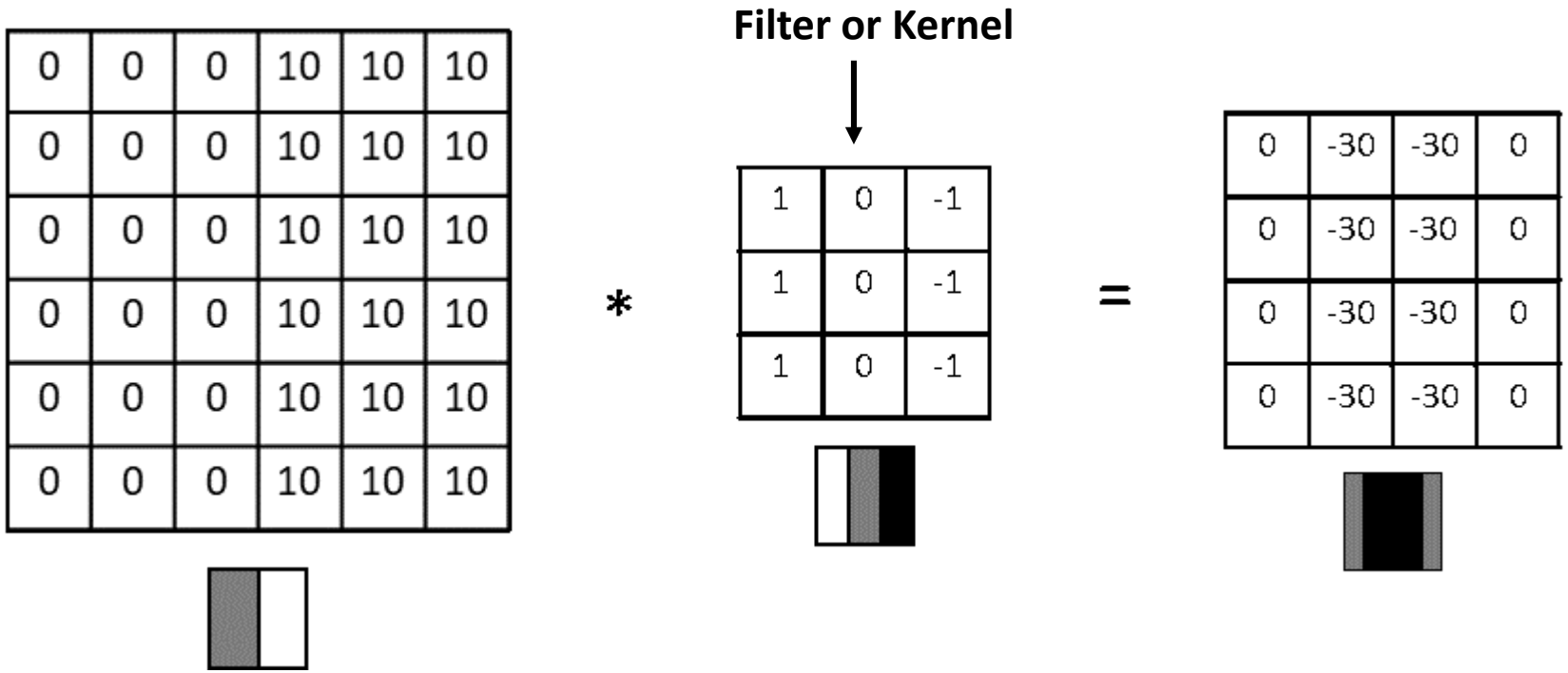
Convolutional Layer GIF from https://blog.usejournal.com/convolutional-neural-networks-why-what-and-how-f8f6dbebb2f9?gi=3faa9b8cfe4c

Convolution Operation: Edge Detection



Convolution image from http://datahacker.rs/edge-detection-extended/

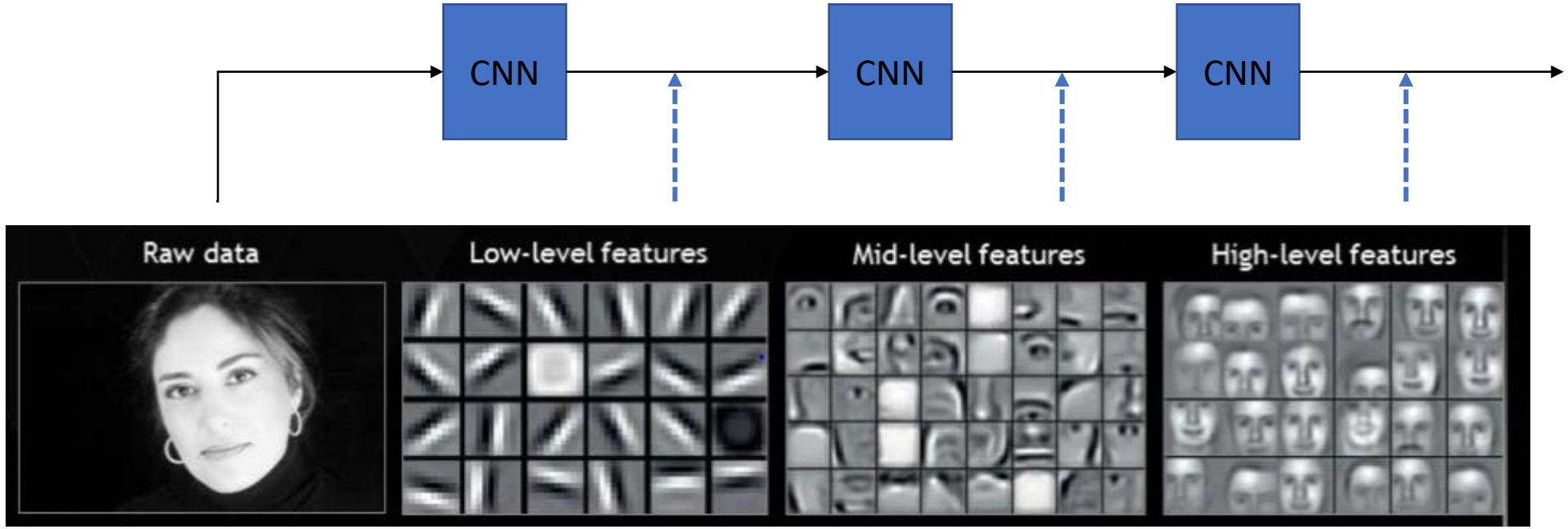
Convolution Operation: Edge Detection



Convolution image from http://datahacker.rs/edge-detection-extended/

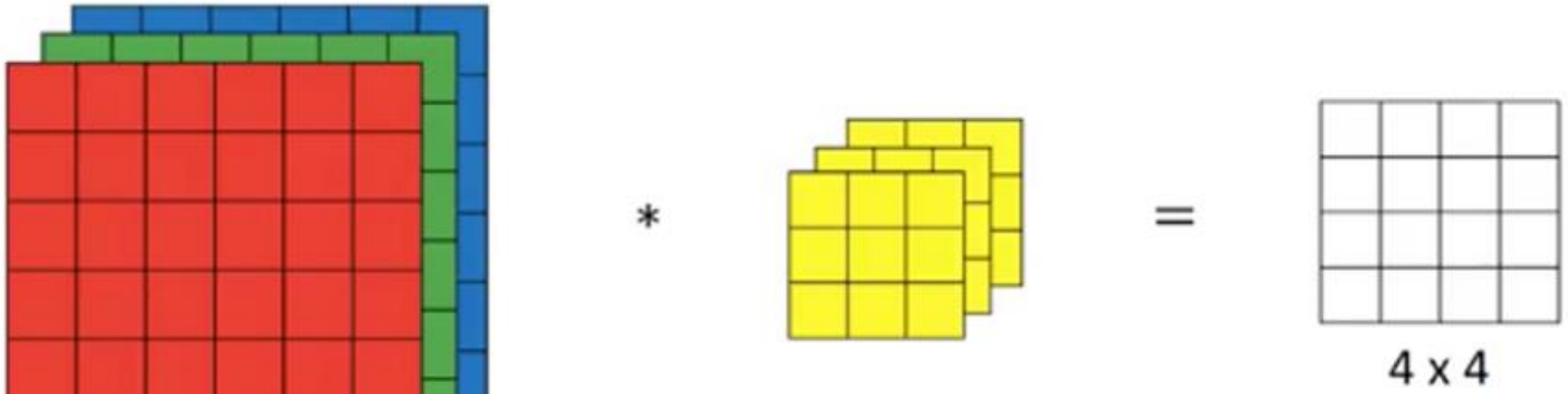
Feature Learning

- Deep Learning is **Representation/Feature Learning**

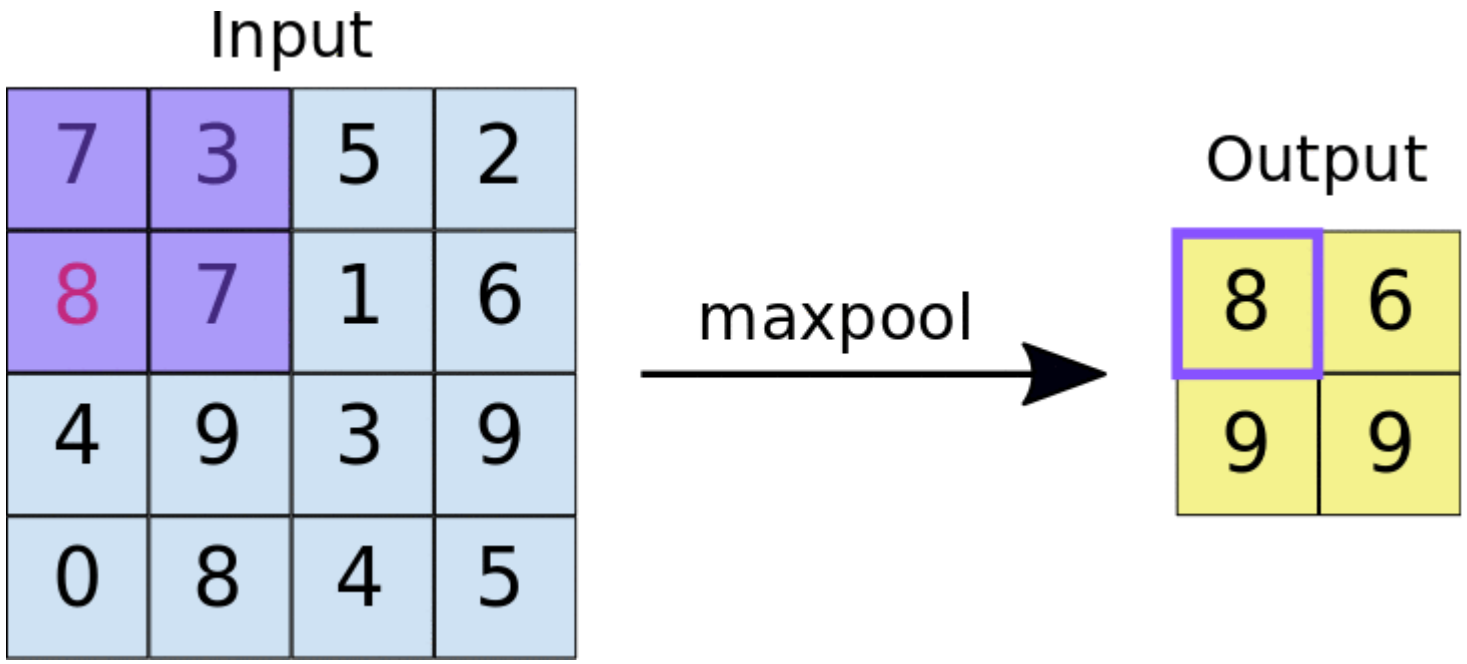


Convolutional Neural Network: Convolutional Layer

- No need to hand choose filters just **learn the correct weights**

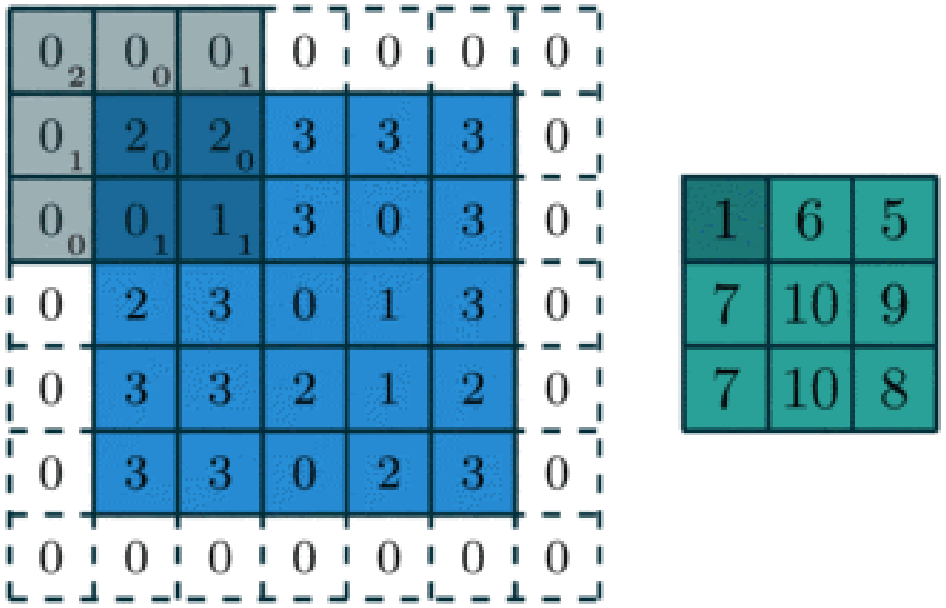


Convolutional Neural Networks: Max Pool



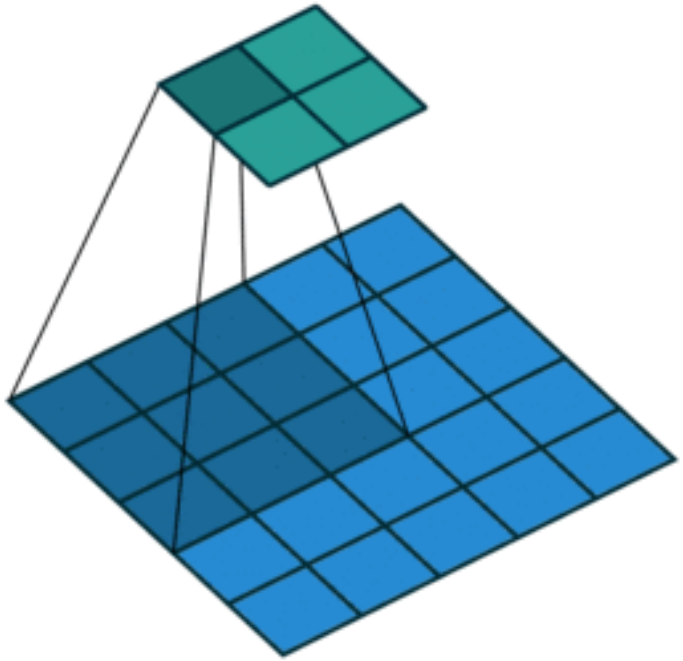
Max Pool layer from https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks

Convolutional Neural Networks: Padding



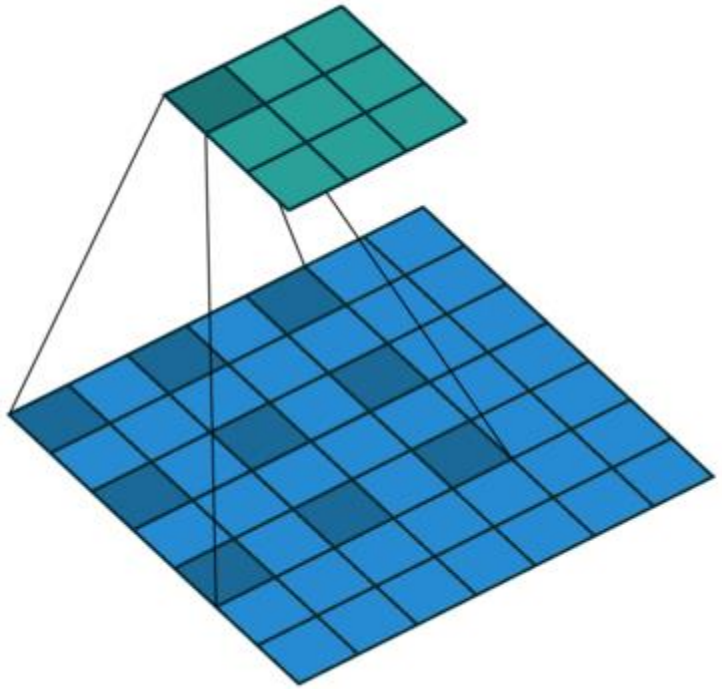
Padding gif from http://deeplearning.net/software/theano/tutorial/conv_arithmetic.html

Convolutional Neural Networks: Strided Convolution



Strided convolution gif from http://deeplearning.net/software/theano/tutorial/conv_arithmetic.html

Convolutional Neural Networks: Dilated Convolution



Dilated convolution gif from http://deeplearning.net/software/theano/tutorial/conv_arithmetic.html

Convolution Output

$$H_{out} = \frac{H_{in} + 2 \text{ x padding}[0] - \text{kernel\_size}[0]}{stride[0]} + 1$$

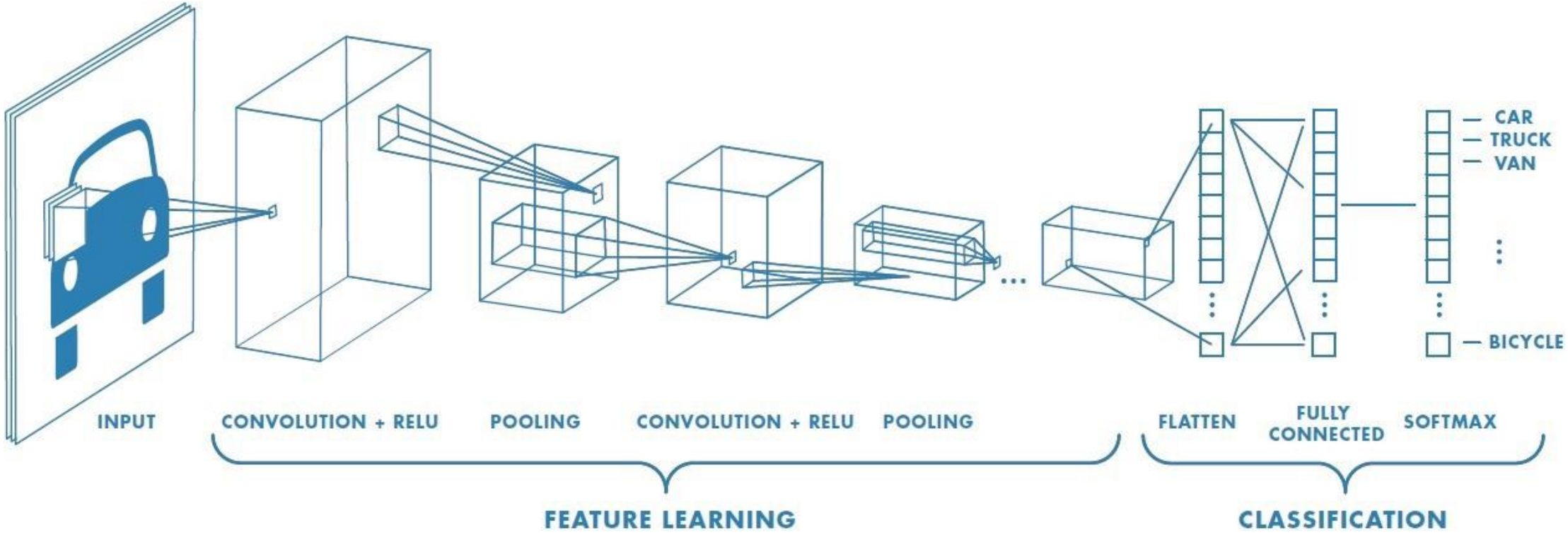
$$W_{out} = \frac{W_{in} + 2 \text{ x padding}[1] - \text{kernel\_size}[1]}{stride[1]} + 1$$

Convolution Output

$$H_{out} = \frac{H_{in} + 2 \text{ x padding}[0] - \text{kernel\_size}[0]}{stride[0]} + 1$$

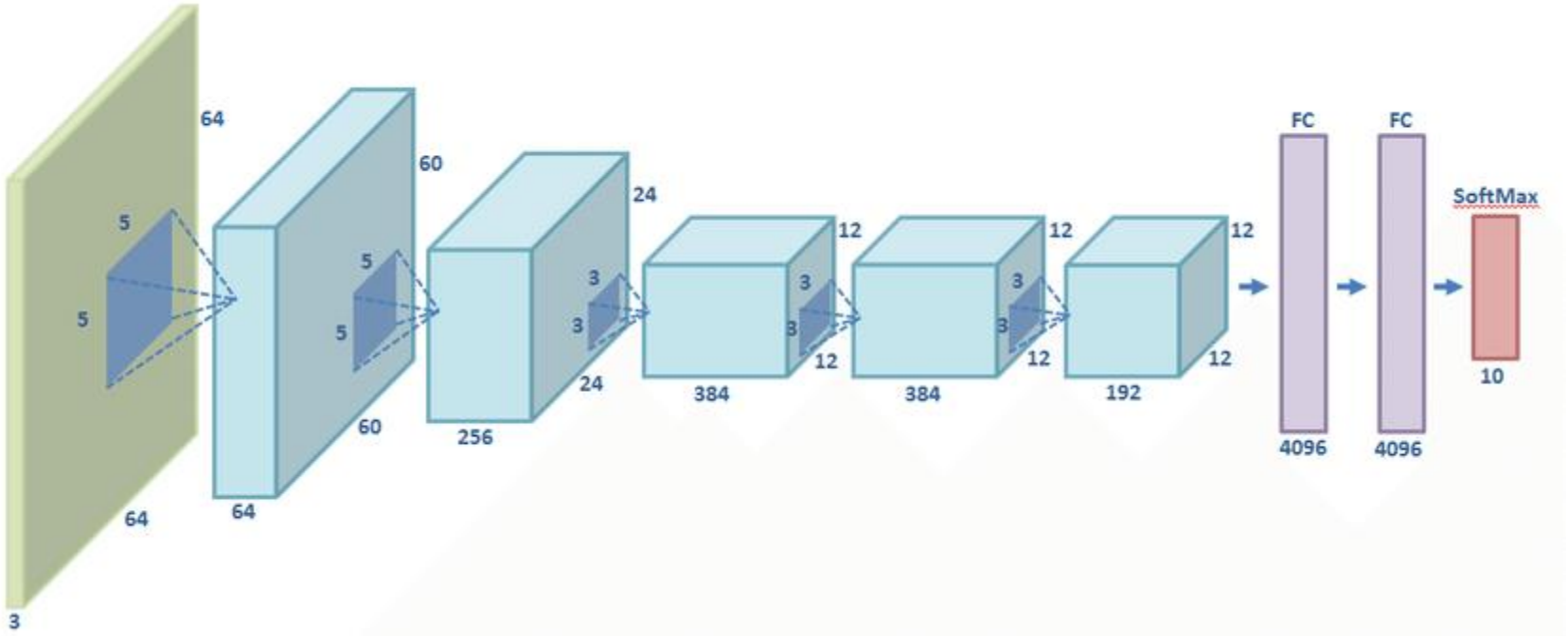
$$W_{out} = \frac{W_{in} + 2 \text{ x padding}[1] - \text{kernel\_size}[1]}{stride[1]} + 1$$

Convolutional Neural Networks: Putting it all together



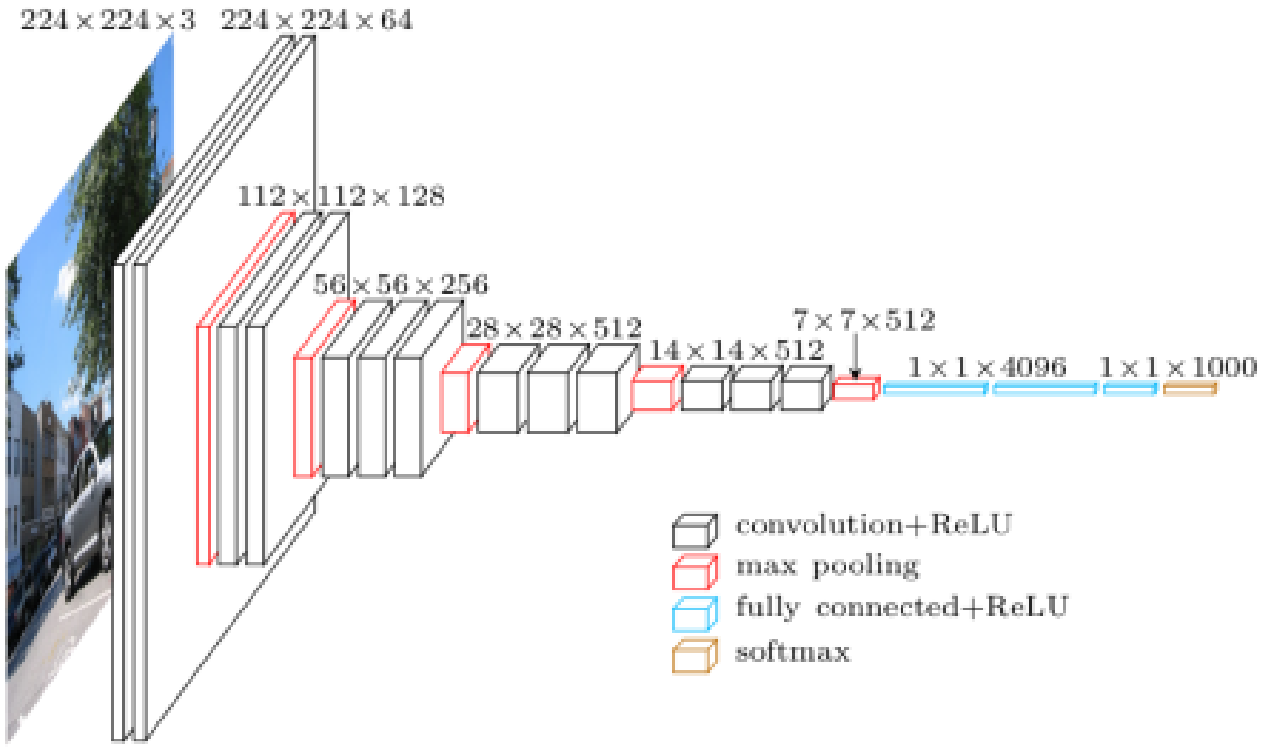
Convolutional Network from https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

Popular Conv Net: AlexNet



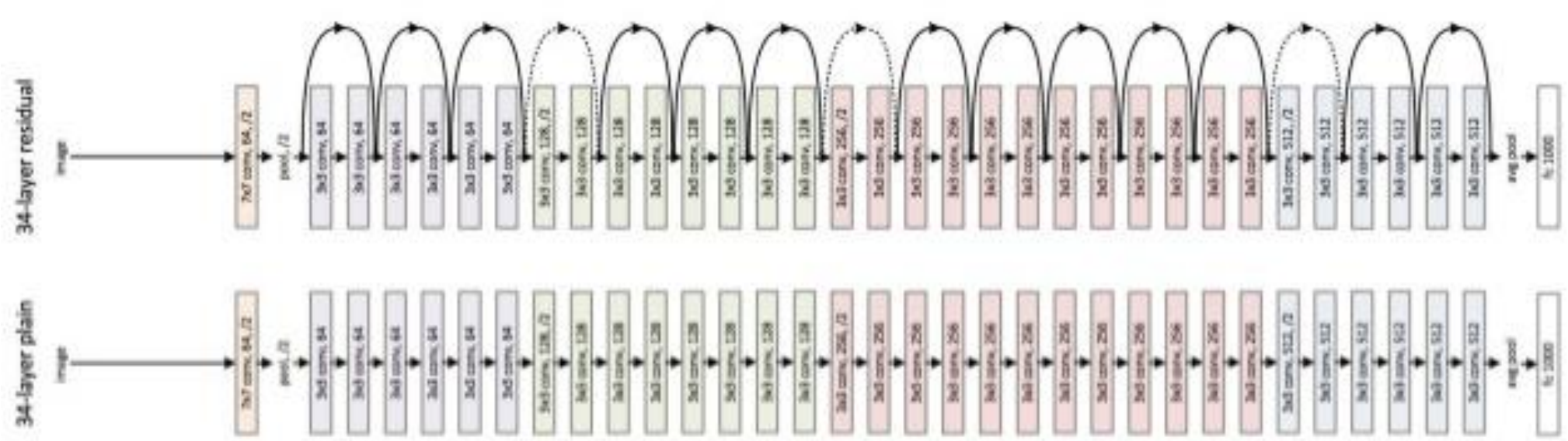
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

Popular Conv Net: VGG



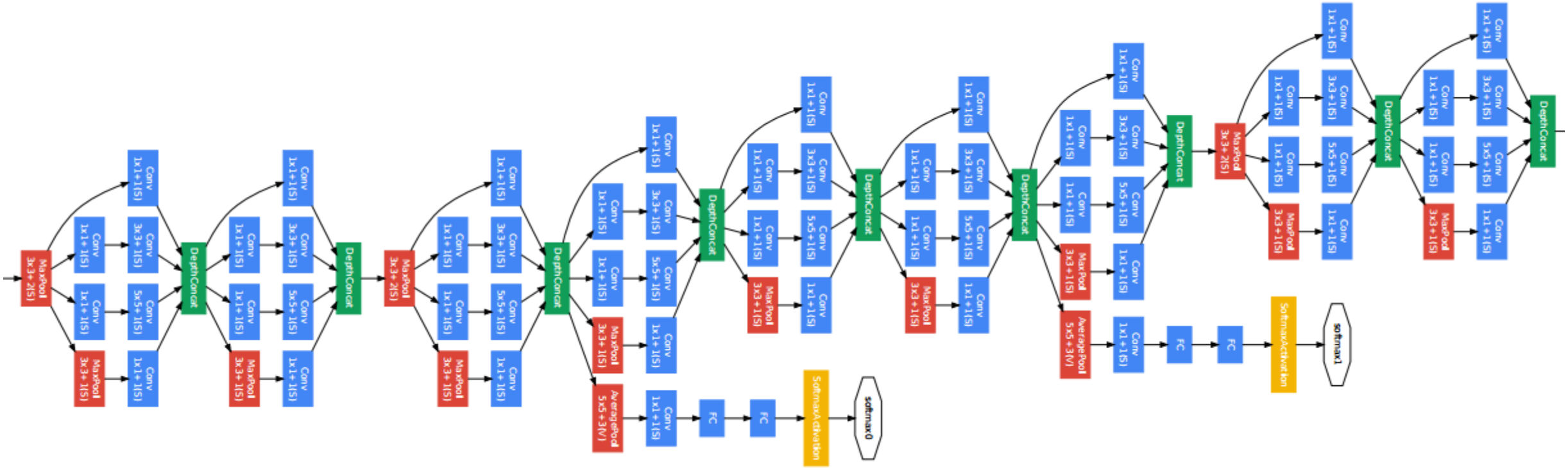
Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

Popular Conv Net: ResNet



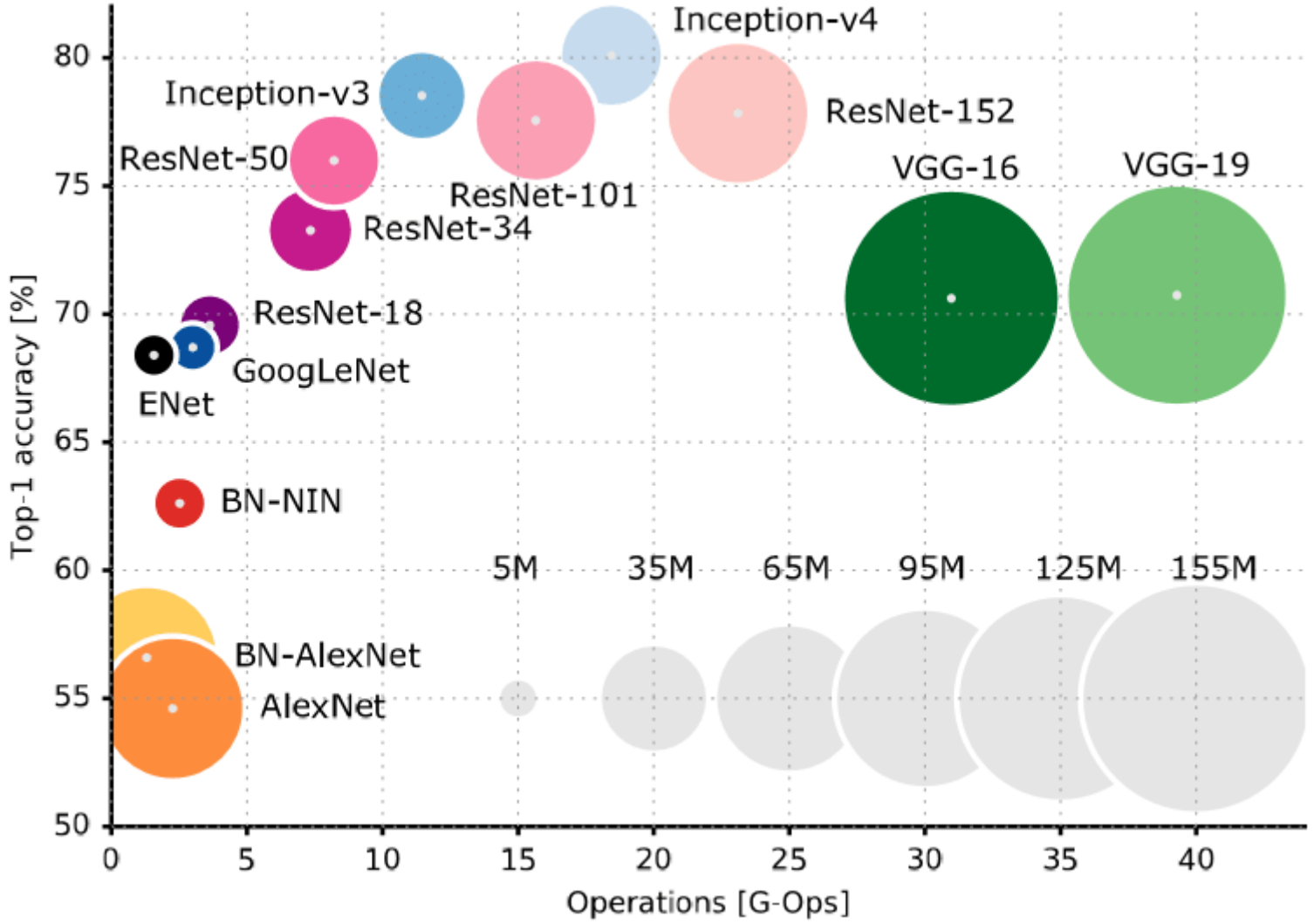
He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

Popular Conv Net: ResNet



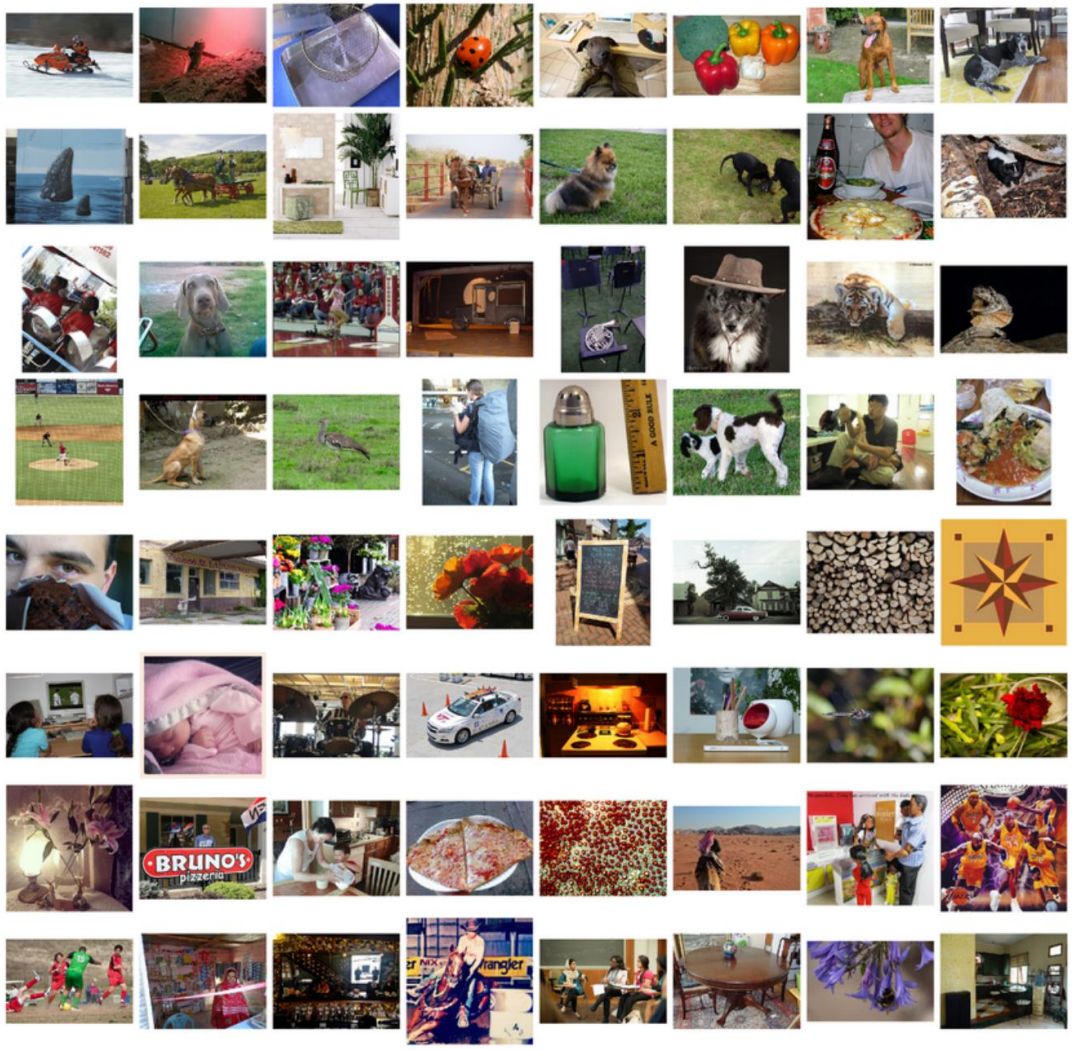
Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

Comparison of Networks on ImageNet Classification



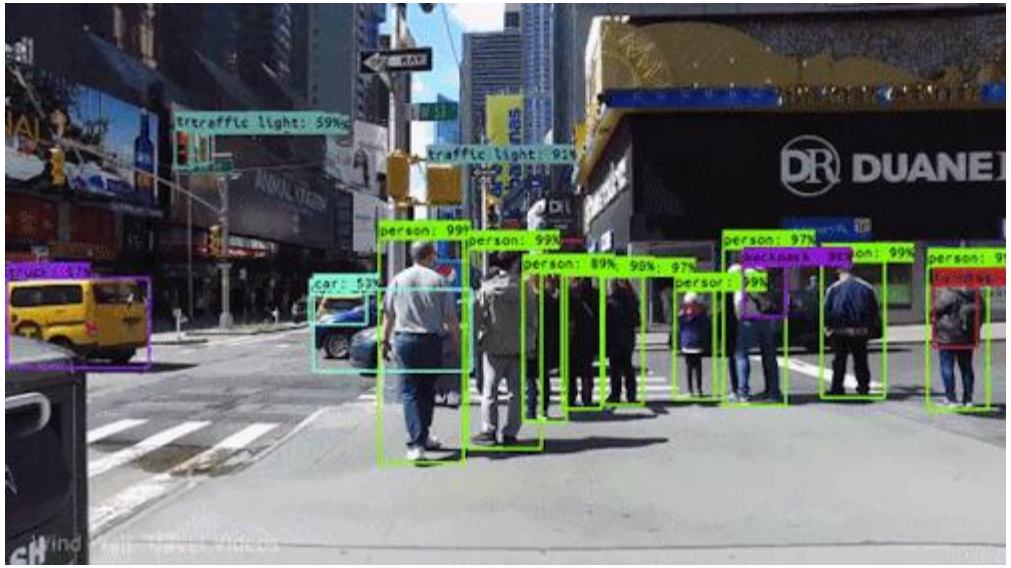
Computer Vision Tasks: Image Classification

- ImageNet is a popular dataset
- 1.2 million training images
- 100 thousand testing images
- 1000 classes



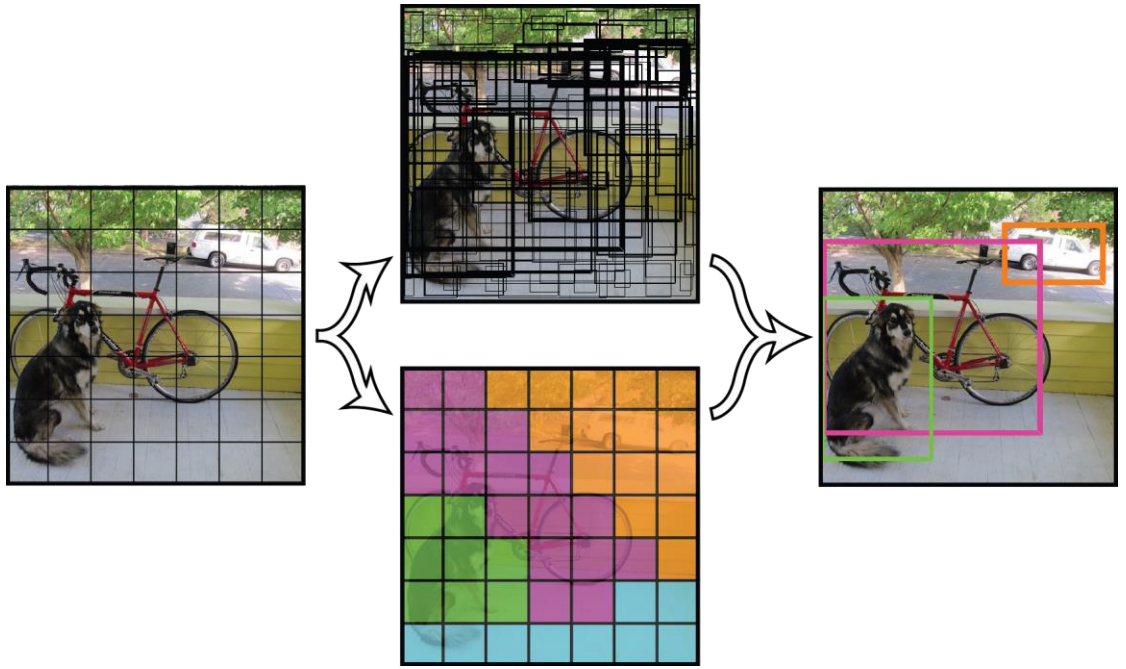
Computer Vision Tasks: Object Detection

- Detect one or multiple object bounding boxes in an image
- Typically multi-tasked with image classification



Object Detection Methods: YOLO (You Only Look Once)

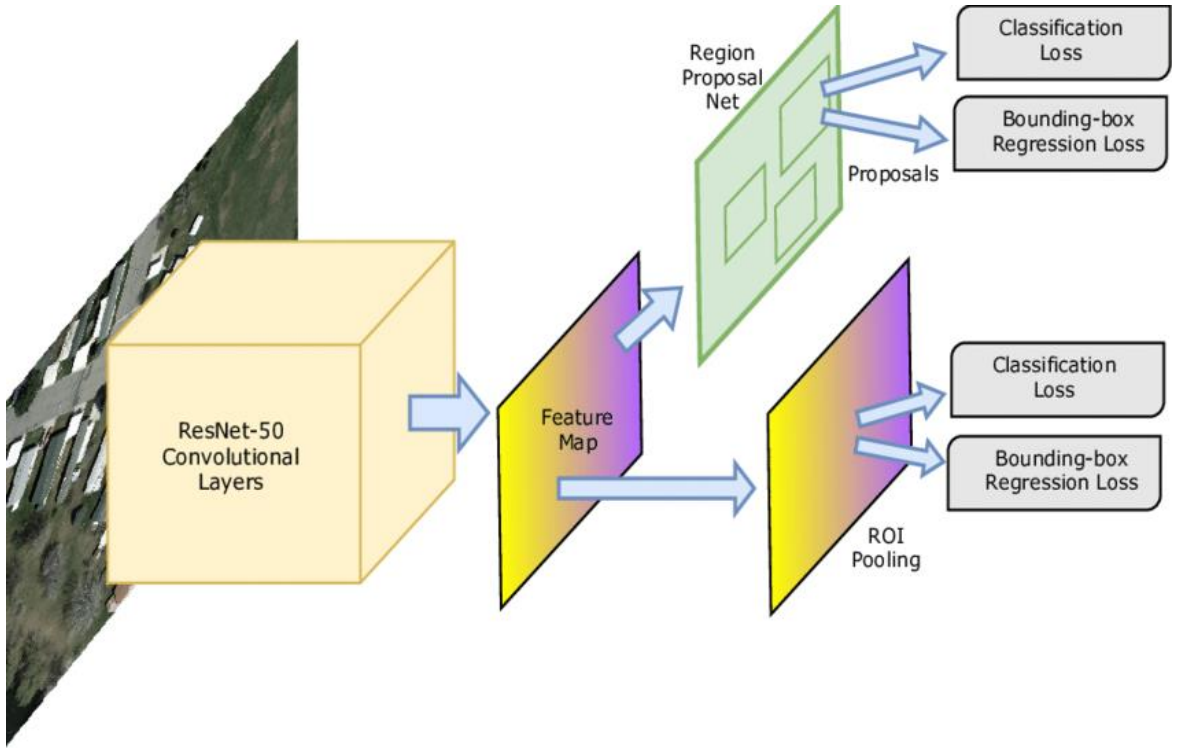
- Single-stage detector



Redmon, Joseph, and Ali Farhadi. "YOLO9000: better, faster, stronger." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

Object Detection Methods: Faster R-CNN

- Multi-stage detector



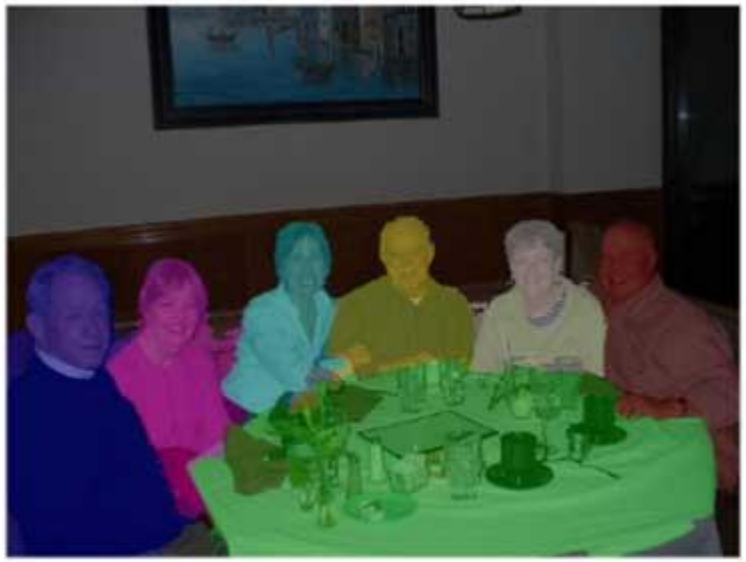
Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.

Computer Vision Tasks: Segmentation

Computer Vision Tasks: Segmentation

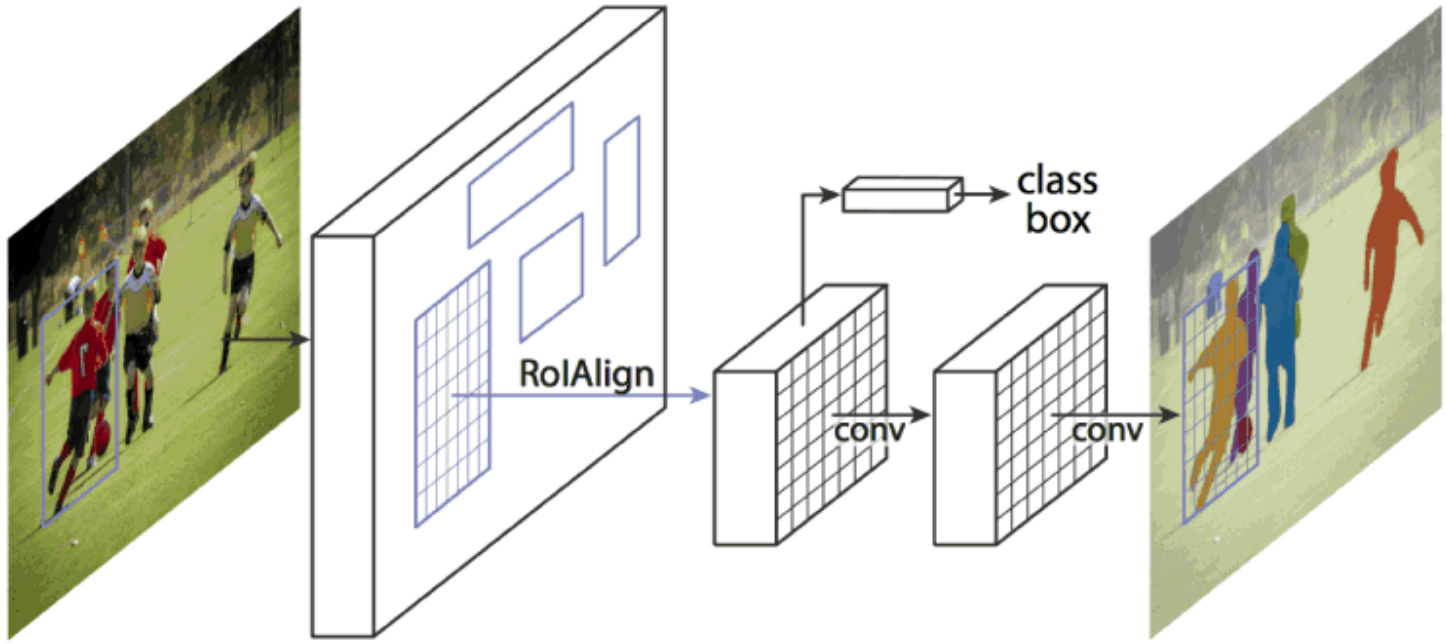
- **Semantic Segmentation:** Detects each pixel and overall object category
- **Instance Segmentation:** Detects each pixel and identifies individual instances of the object

Semantic Segmentation



Instance Segmentation

Instance Segmentation Methods: Mask R-CNN

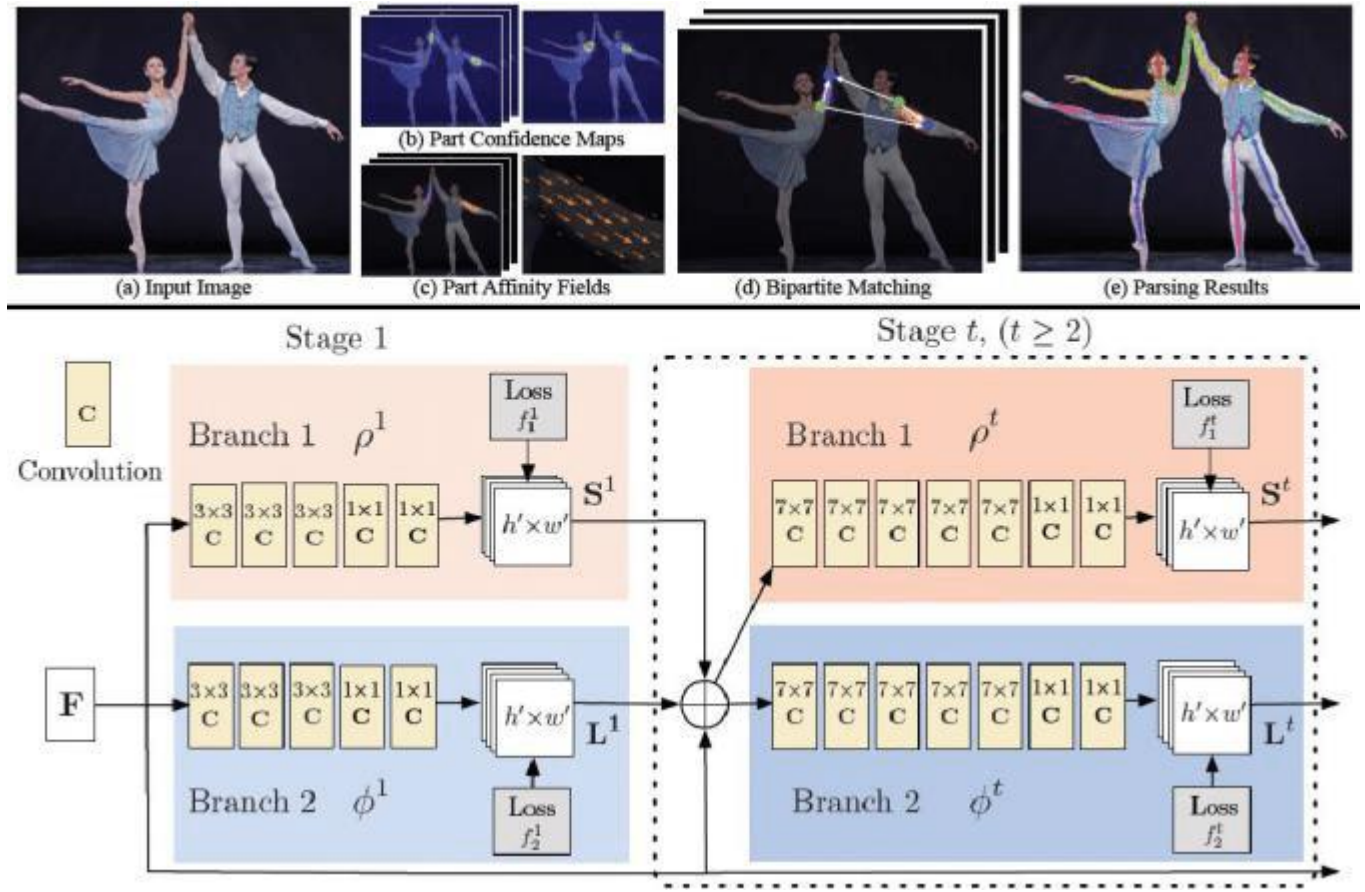


The Mask R-CNN framework for instance segmentation

He, Kaiming, et al. "Mask r-cnn." Proceedings of the IEEE international conference on computer vision. 2017.

Computer Vision Tasks: Pose Estimation

Pose Estimation Methods: OpenPose



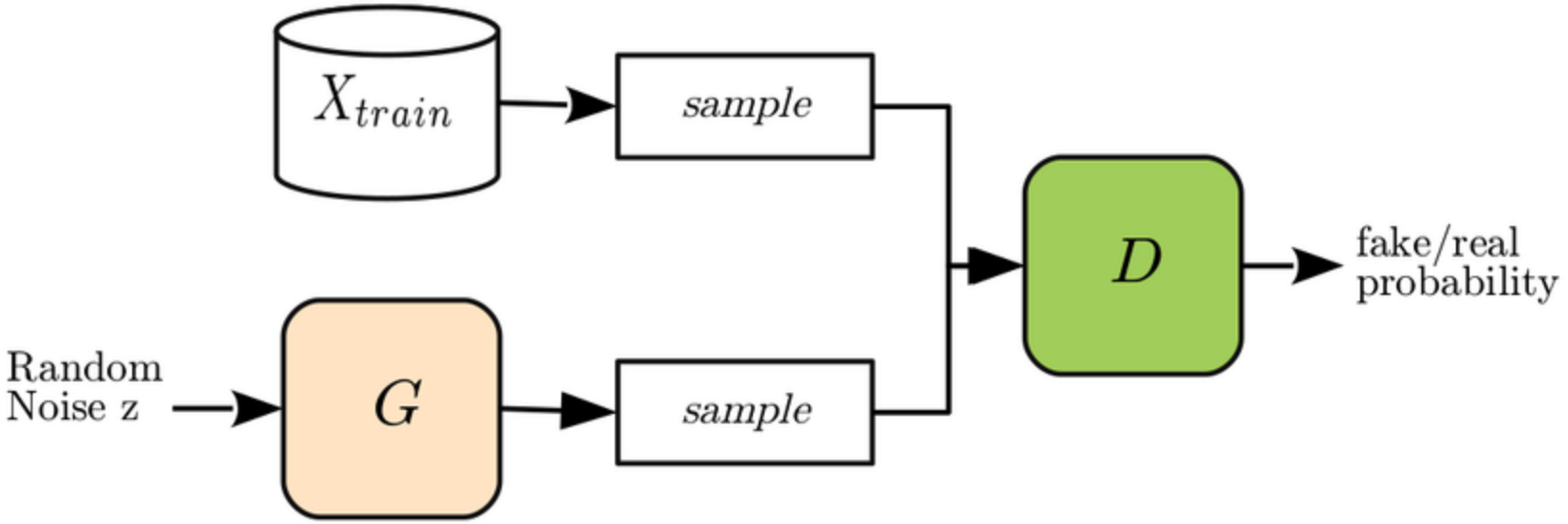
Cao, Zhe, et al. "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields." arXiv preprint arXiv:1812.08008 (2018).

Computer Vision Tasks: Style Transfer

Computer Vision Tasks: Image Generation

Popular Networks: Generative Adversarial Networks

Discriminator tries to distinguish real from fake

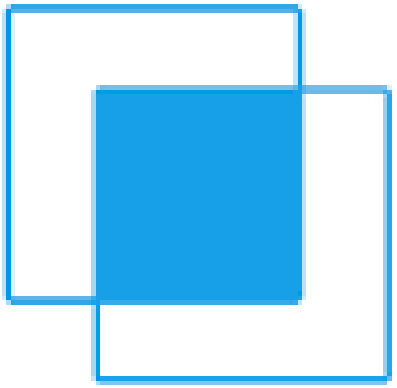


Generator tries to fool the **Discriminator**

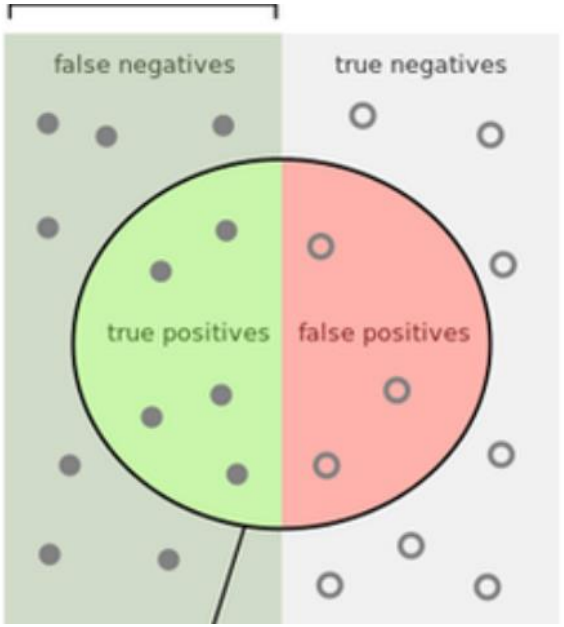
1: Goodfellow, Ian, et al. "Generative adversarial nets." Advances in neural information processing systems. 2014.

Evaluation Metrics: Intersection over Union

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Evaluation Metrics: Precision and Recall



How many selected items are relevant?

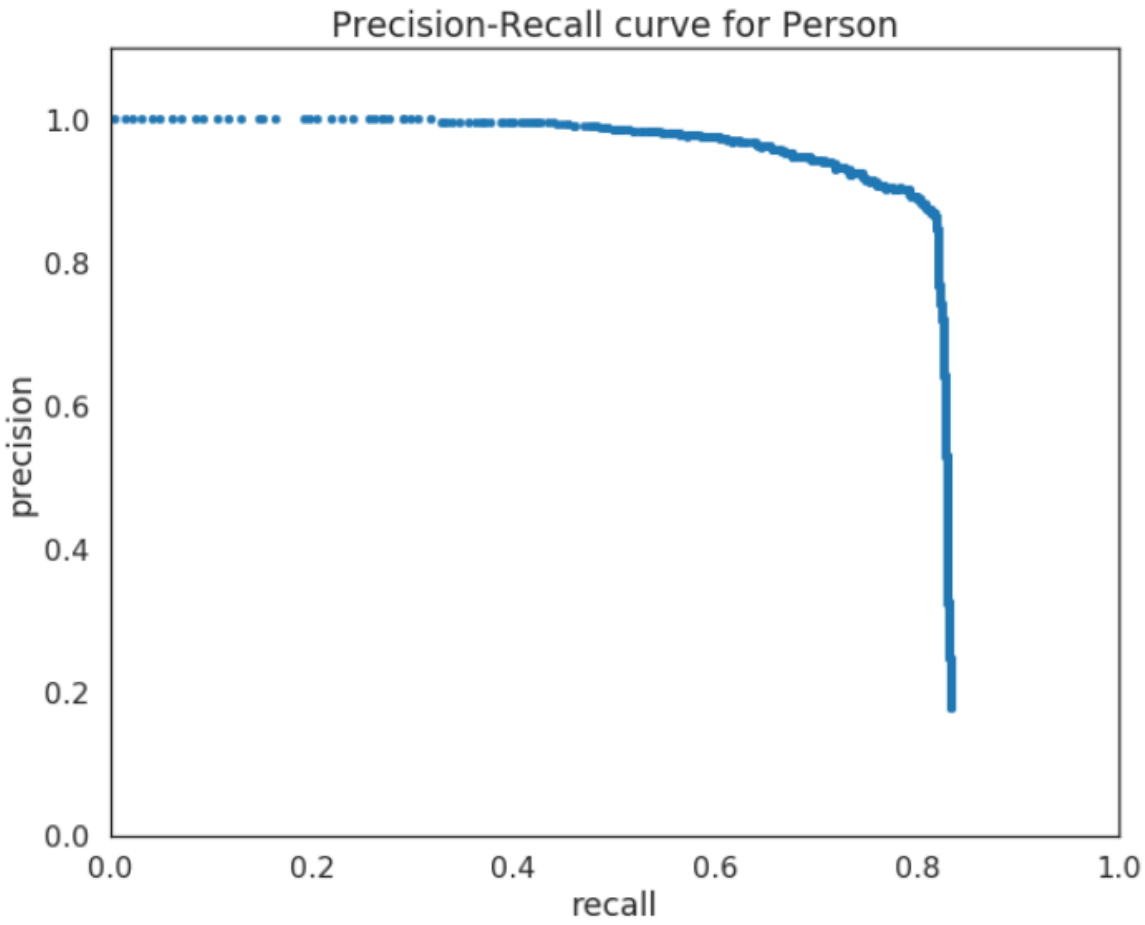
$$\text{Precision} = \frac{\text{true positives}}{\text{true positives} + \text{false positives}}$$

How many relevant items are selected?

$$\text{Recall} = \frac{\text{true positives}}{\text{true positives} + \text{false negatives}}$$

Evaluation Metrics: Precision and Recall

- Plot for varying positive prediction thresholds



Evaluation Metrics: F1 Score

- Single metric considering both precision and recall
- Good for single metric evaluation
- Gives equal importance to precision and recall (this may cause issues when you regard precision more over recall or vice versa)

$$F1 = 2\left(\frac{precision\ x\ recall}{precision + recall}\right)$$

Questions?