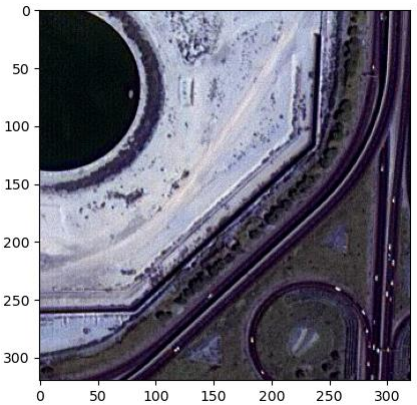
Architectures

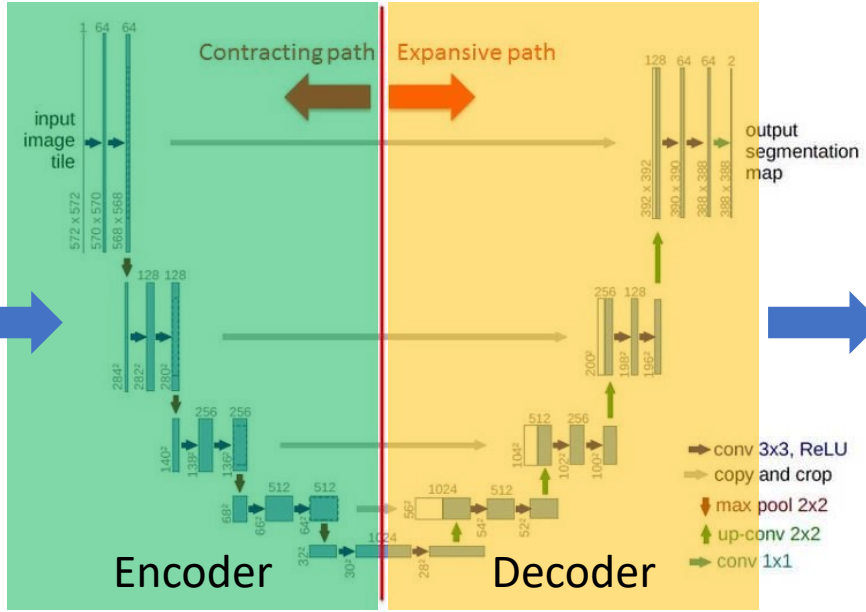
Semantic Segmentation

General Introduction

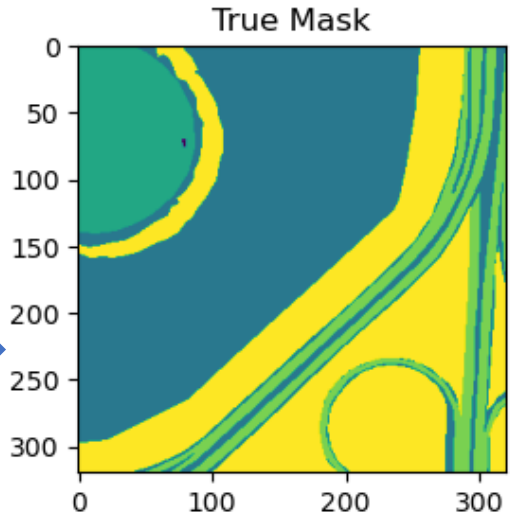
Training Loop



Input Image
from train dataset



Source: http://deeplearning.net/tutorial/unet.html

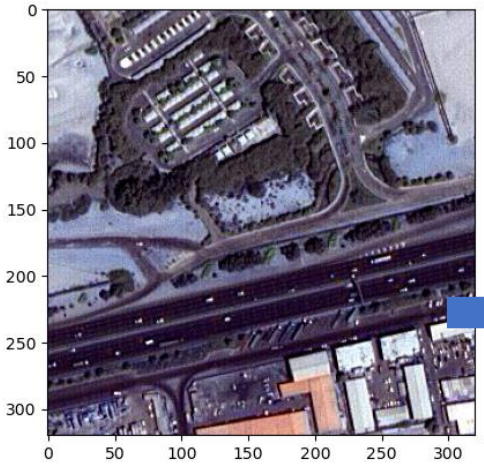


Output Image
from train dataset
(True Mask)

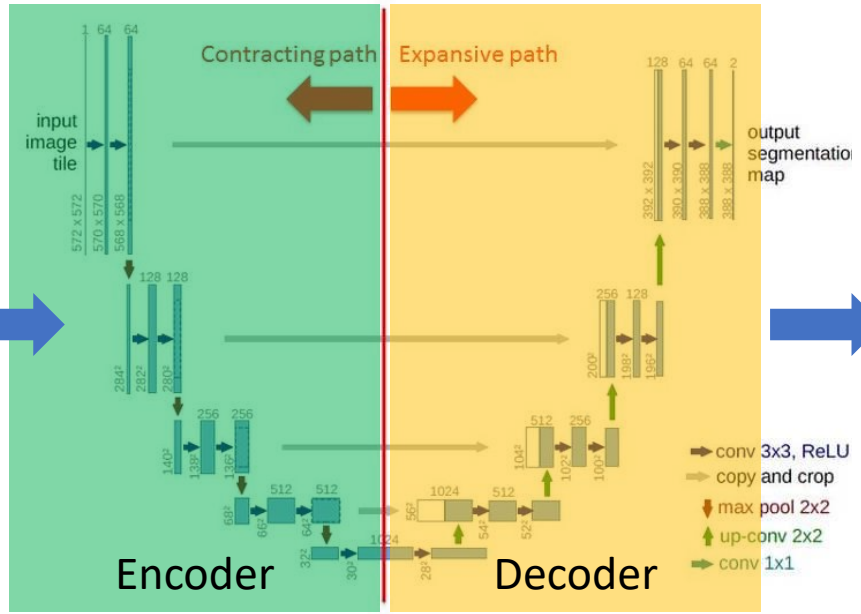
Semantic Segmentation

General Introduction

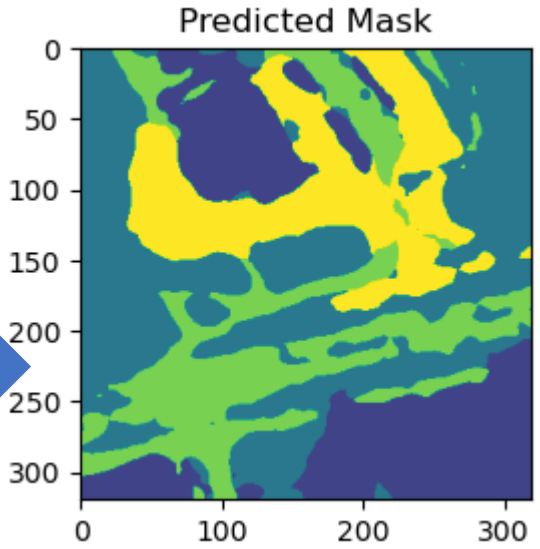
Inference



Input Image
from test dataset
(unseen)



Source: http://deeplearning.net/tutorial/unet.html

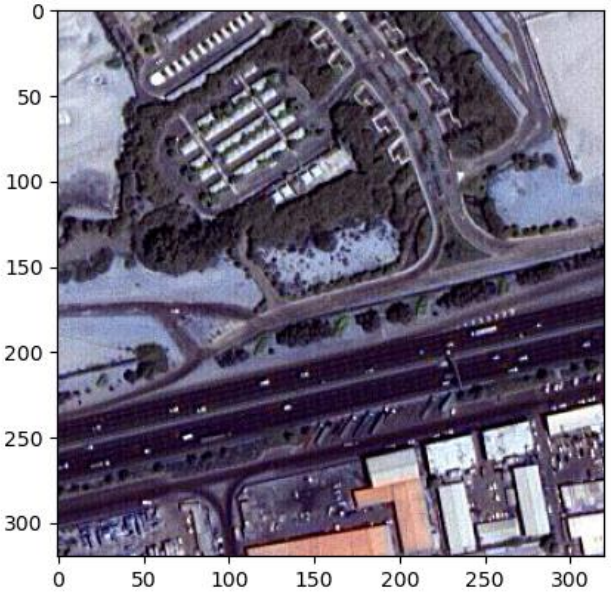


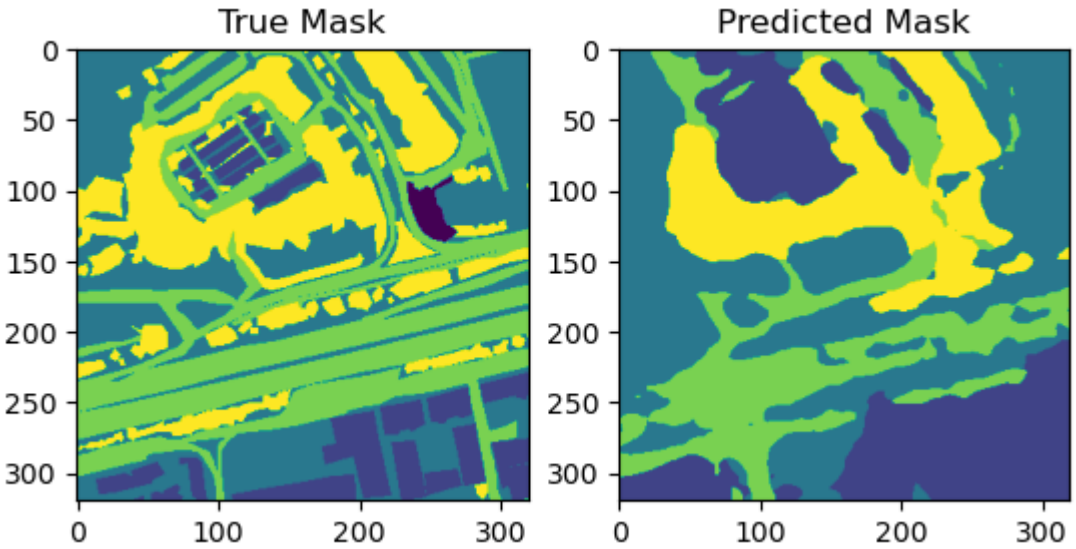
Output Image
(Predicted Mask)

Semantic Segmentation

General Introduction

- used architecture has a huge impact on final result

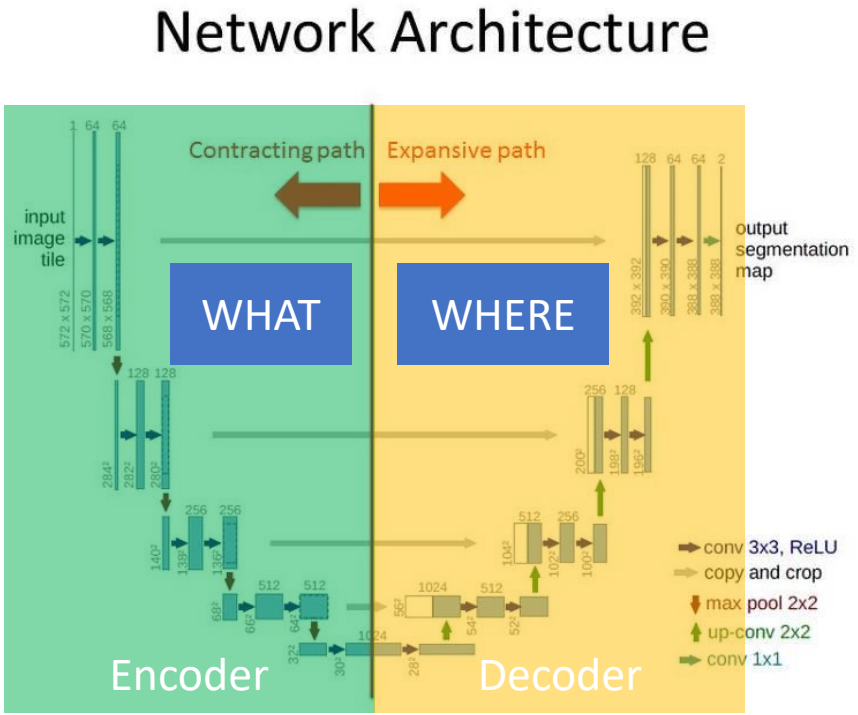




Semantic Segmentation

UNET

- Symmetrical topology
- Three parts:
  - Contracting path
  - Bottleneck path
  - Expansive path
- Keeps local information from contracting path
- Contextual information in expansive path
- skip connections
- Images of different size can be used as inputs

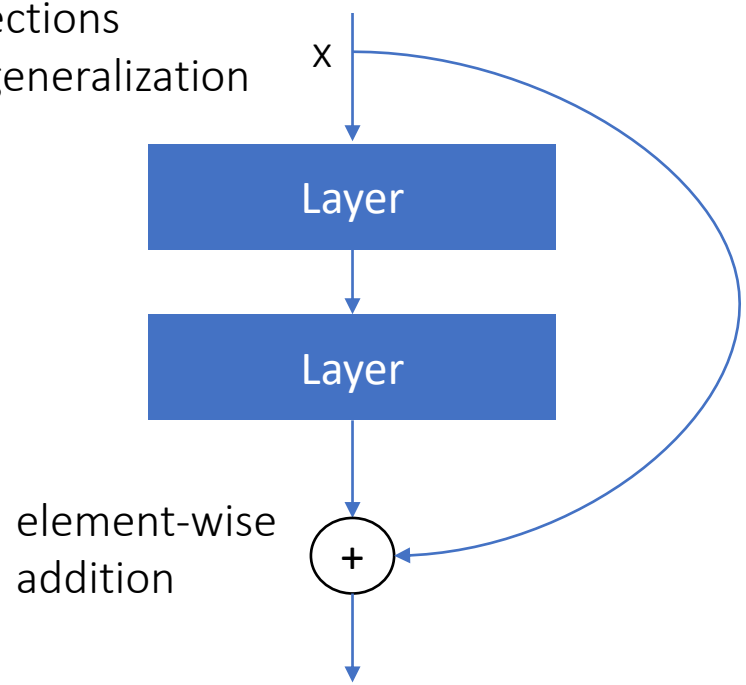


Source: http://deeplearning.net/tutorial/unet.html

Semantic Segmentation

Skip Connections

- also called Shortcut Connections or residual connections
- useful technique for improving performance and generalization
- allow gradient signal to bypass one or more layers
- often used in networks with deep architectures
- purpose:
  - overcome vanishing gradient problem
  - overcome overfitting
  - help network learn more easily
  - improve performance
- useful in different applications
  - image classification
  - language translation
  - speech recognition



Semantic Segmentation

FPN

- Feature Pyramid Network
- uses network like Faster R-CNN for object detection
- acts as feature extractor
- image input processed to feature maps at multiple levels
- improves accuracy
- bottom-up pathway
  - normal CNN
  - different backends, e.g. Resnet
- top-down pathway
  - deeper features merged with lower features

Source: Tsung-Yi Lin, et.al. „Feature Pyramid Networks for Object Detection”