Mo Patel
Director, Artificial Intelligence
Think Big Analytics
Laura Froelich
Data Scientist
Think Big Analytics

Training vision models with public transportation datasets

Computer Vision

Convolutional Neural Networks

Object Detection

Image Segmentation

© 2017 Teradata

TERADATA.

Agenda

- Image Segmentation
- Fully Convolutional Networks (FCN)
- Training Vision Models Challenges

TERADATA.

Image Segmentation

- Evolution from Classification → Detection → Segmentation
- Goal: Pixel level identification
- Metrics: Jaccard Index
  - $IoU = \frac{True\ Positive}{True\ Positive + False\ Positive + False\ Negative}$
- Datasets:
  - PASCAL VOC
  - MS COCO
  - Cityscapes
  - ADE20K
  - Imagenet
  - KITTI

TERADATA.

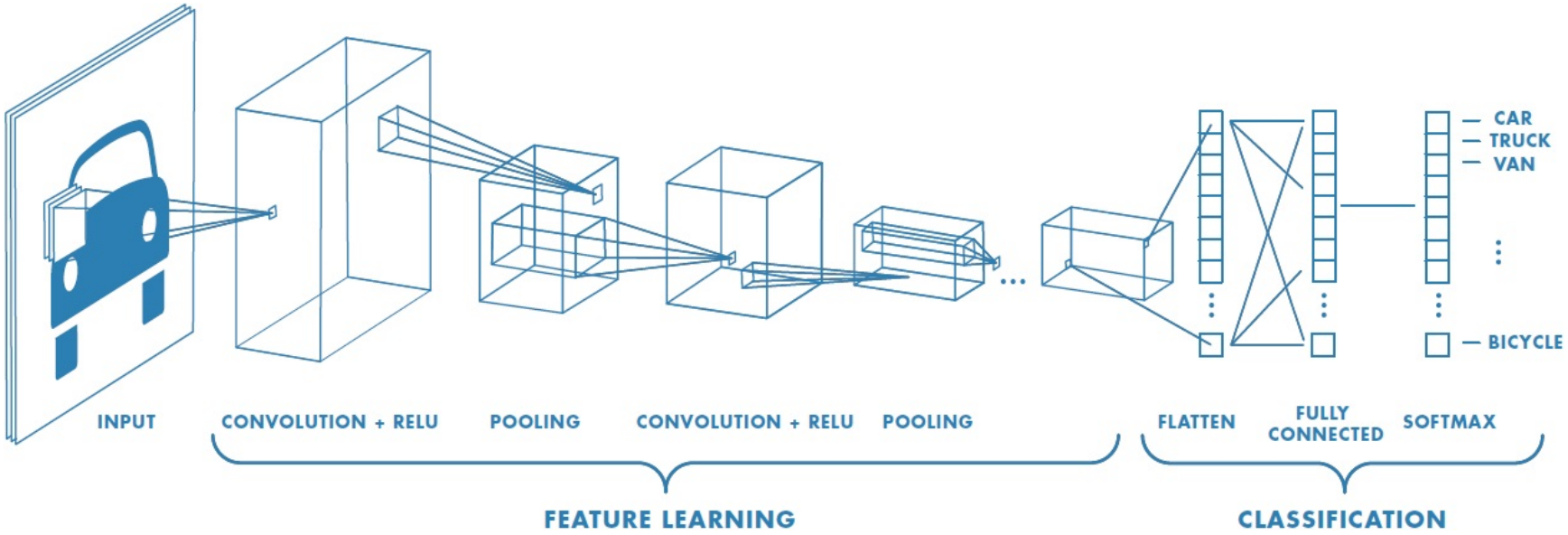
Agenda

- Image Segmentation
- Fully Convolutional Networks (FCN)
- Training Vision Models Challenges

TERADATA.

Fully Convolutional Networks

- Image Classification Objective: Predict Image Label

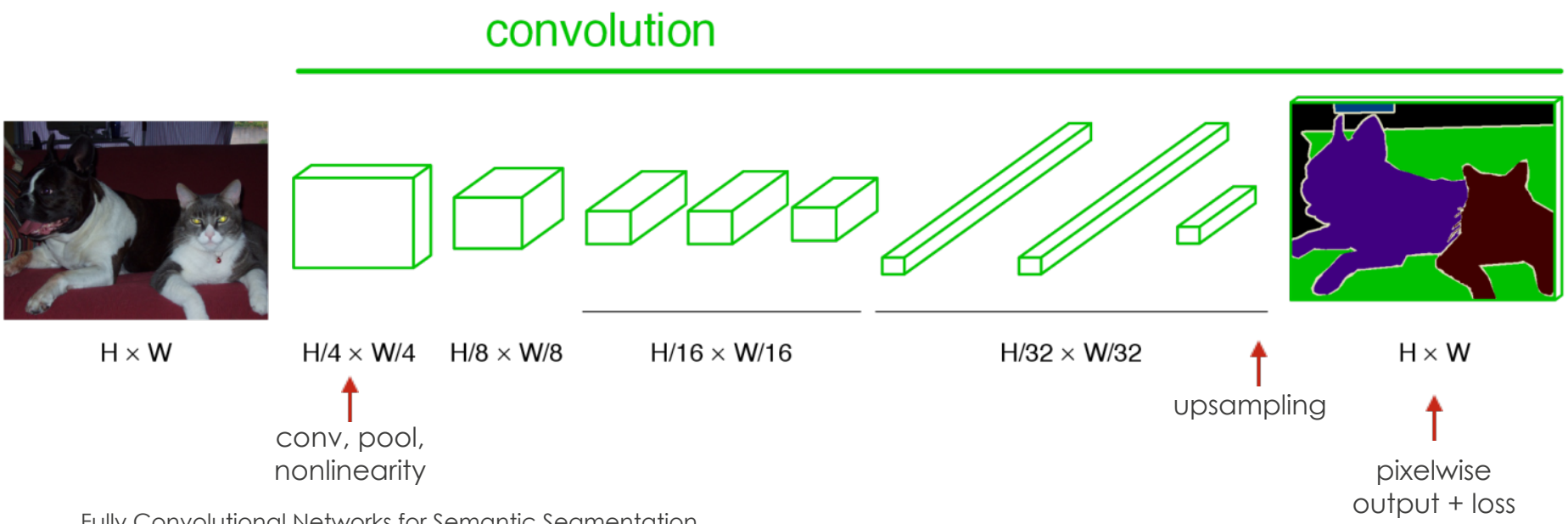


Source: Mathworks

© 2017 Teradata

Fully Convolutional Networks

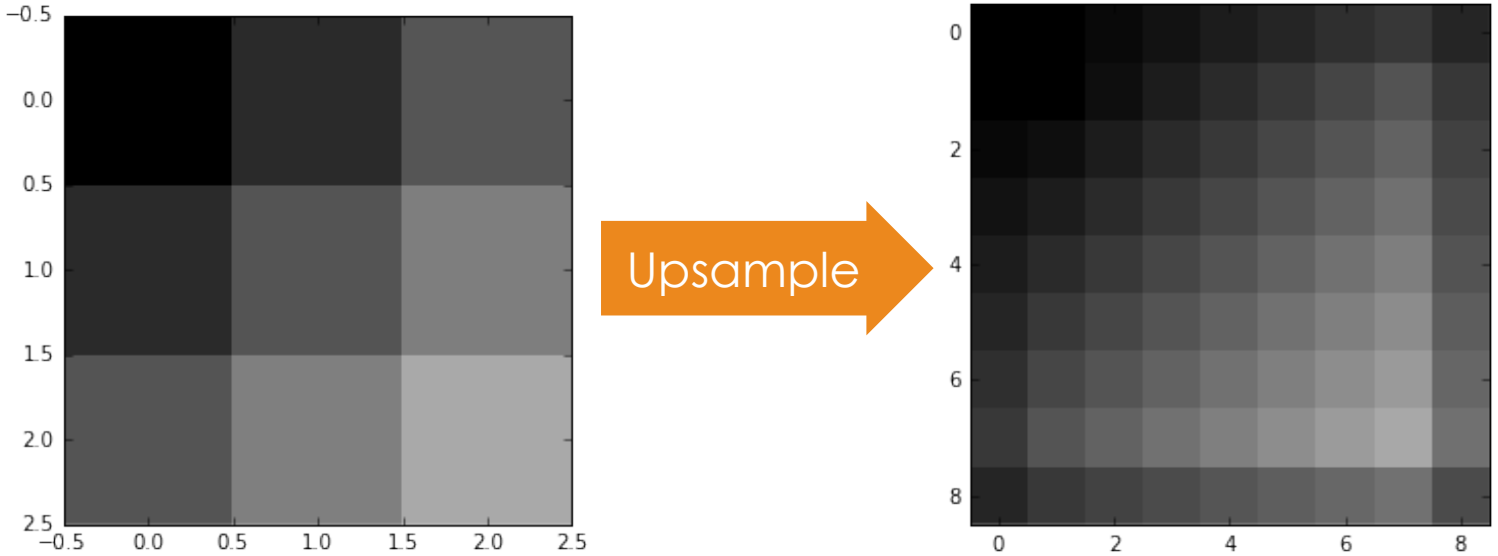
- Segmentation Objective: Predict Pixel Label



Fully Convolutional Networks for Semantic Segmentation
Evan Shelhamer Jonathan Long Trevor Darrell - UC Berkeley in CVPR'15, PAMI'16

© 2017 Teradata

FCN: Upsampling



Upsampling and Image Segmentation with Tensorflow and TF-Slim
Source: http://warmspringwinds.github.io/tensorflow/tf-slim/2016/11/22/upsampling-and-image-segmentation-with-tensorflow-and-tf-slim/

© 2017 Teradata

Fully Convolutional Networks: Implementations

- Future easy to use– Coming to Keras: Dense Prediction API Design, Including Segmentation and Fully Convolutional Networks
  - https://github.com/fchollet/keras/issues/6538#issuecomment-301342345
- Original Caffe: https://github.com/shelhamer/fcn.berkeleyvision.org
- TensorFlow: https://github.com/MarvinTeichmann/tensorflow-fcn

© 2017 Teradata

TERADATA.

Agenda

- Image Segmentation
- Fully Convolutional Networks (FCN)
- Training Vision Models Challenges

TERADATA.

Training Vision Models Challenges: Time

Batch Size	epochs	Top-1 Accuracy	hardware	cost ($)	time
256	100	58.7%	8-core CPU + K20 GPU	3,000	144h
512	100	58.8%	1 DGX station	129,000	6h 10m
4096	100	58.4%	1 DGX station	129,000	2h 19m
32K	100	58.5%	512 KNLs	1.2 million	24m

Table 7: The speed and hardware cost for training AlexNet.
For batch size=32K, we changed local response norm in AlexNet to batch norm.

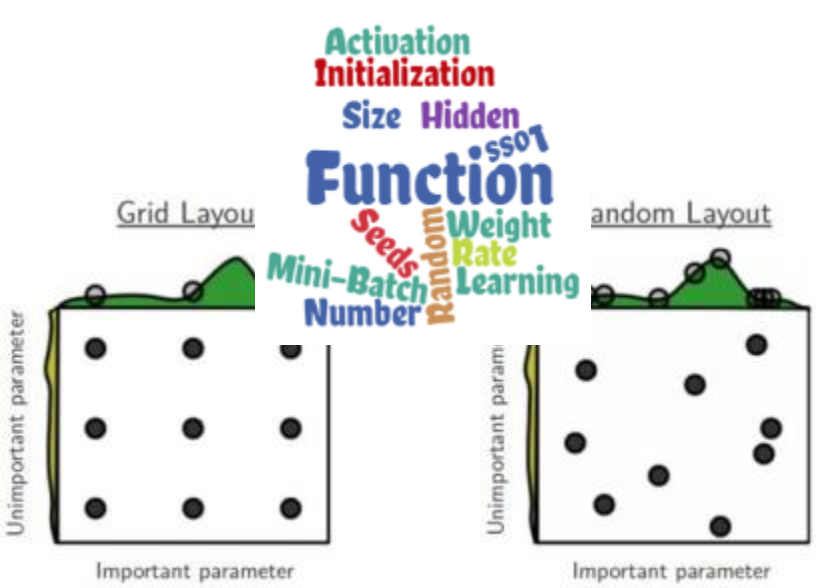
Batch Size	epochs	Top-1 Accuracy	hardware	cost ($)	time
256	90	73.0%	1 DGX station	129,000	21h
8192	90	72.7%	1 DGX station	129,000	21h
8192	90	72.7%	32 DGX stations	4.1 million	1h
32K	90	72.4%	512 KNLs	1.2 million	1h

Table 8: The speed and hardware cost for training ResNet50. We did not use data augmentation.

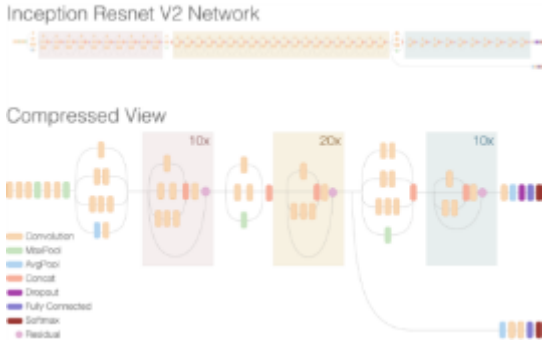
ImageNet Training in 24 Minutes
https://arxiv.org/abs/1709.05011

© 2017 Teradata

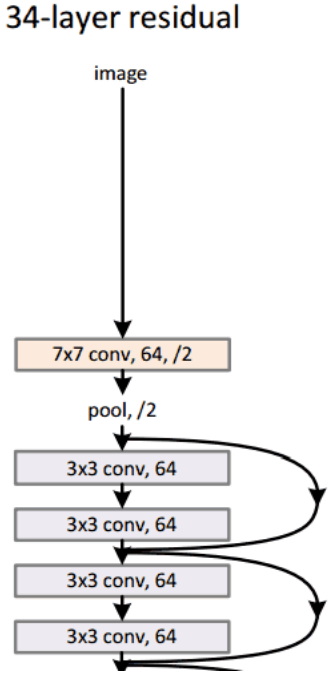
Architectures & Hyper-Parameters



Hyper-Parameter Optimization



Google Inception v3



Microsoft Resnet

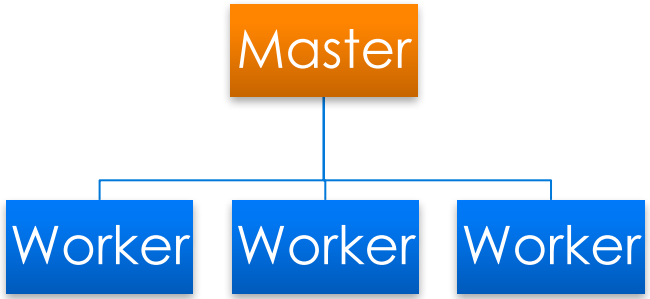
Architecture GIFs: https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html
More about hyper parameteric optimization http://colinraffel.com/wiki/neural_network_hyperparameters

TERADATA.

Distributed Training

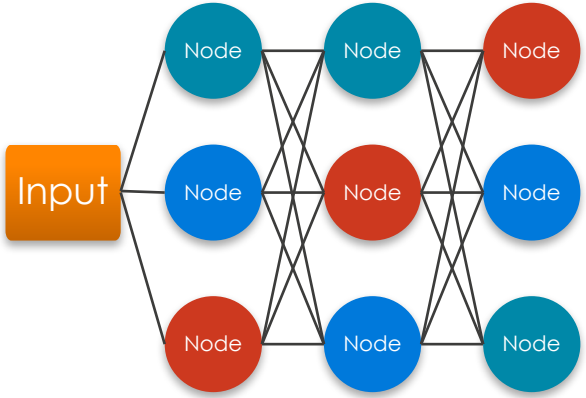
Data Parallel

- Distribute Data
  - Synchronous
  - Asynchronous



Model Parallel

- Distribute Model Operations



© 2017 Teradata