Road Segmentation with Neural Networks

Andy Lau

University of Minnesota, Morris

UMM CSci Senior Seminar Conference, April 2018

What is road segmentation?

- Crucial component for enabling fully autonomous driving
- Determining road surfaces from a series of images

Outline

Cameras

- Cameras are mounted around the vehicle to create a 360° view of the surroundings
- Limitations include
  - Adverse weather condition
  - Shadows and reflections

LIDAR

- Stands for Light Detection and Ranging
- Have three major components
  - Transmitter
  - Receiver
  - Optical analyzing system
- Collecting the distance and time it takes for the emitted signal to return

Neural Networks

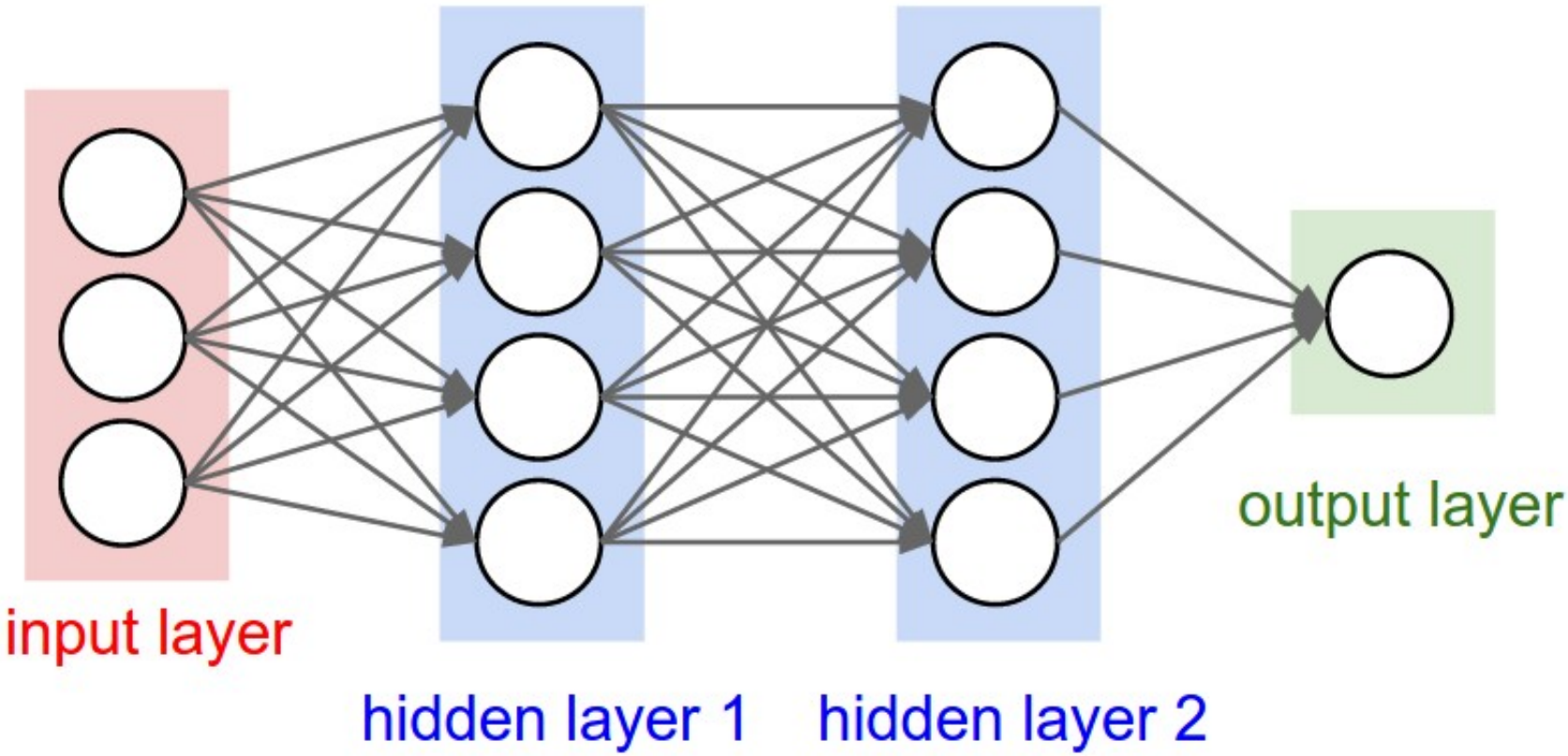


Figure: A simple neural network diagram with two hidden layers

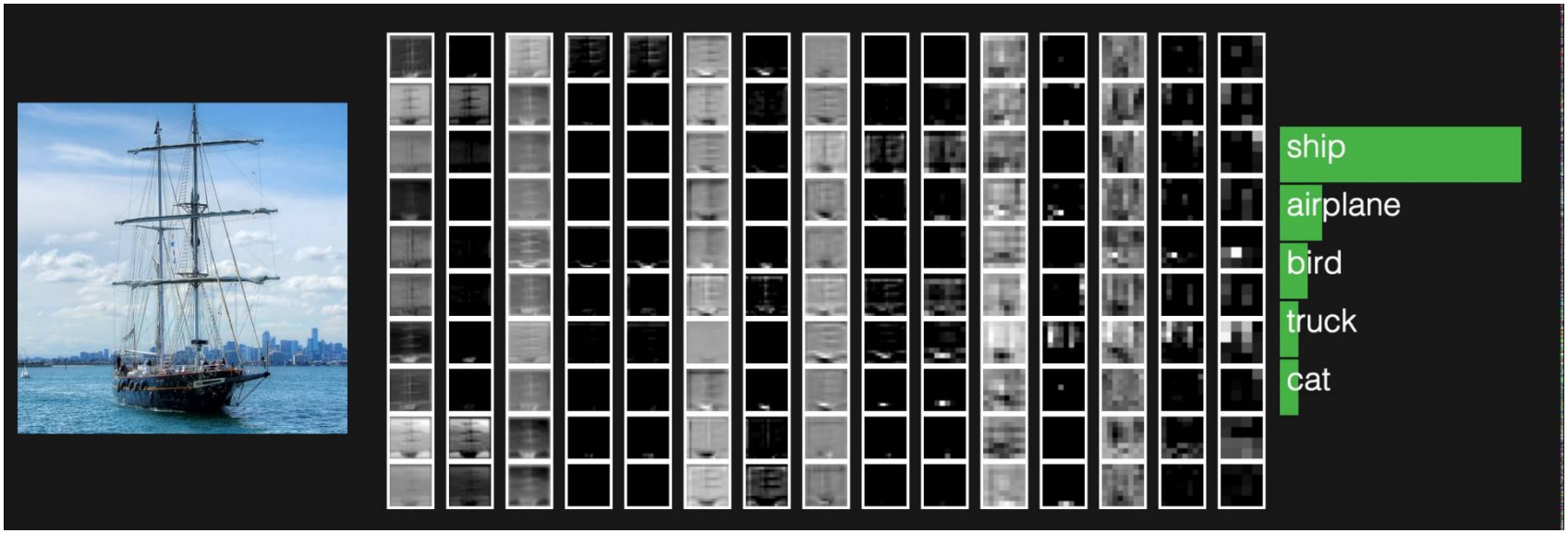
Training

- Loss (object or cost) function, represents the amount of inconsistencies between the output of the network and supposedly correct output
- Goal is to minimize loss

Training

- Initialize all weights
- Process of training a network
  - Forward propagate with a training set
  - Calculate the loss
  - Propagate backward
  - Update and revise weights
- Gradient Descent

Image Classification Problem



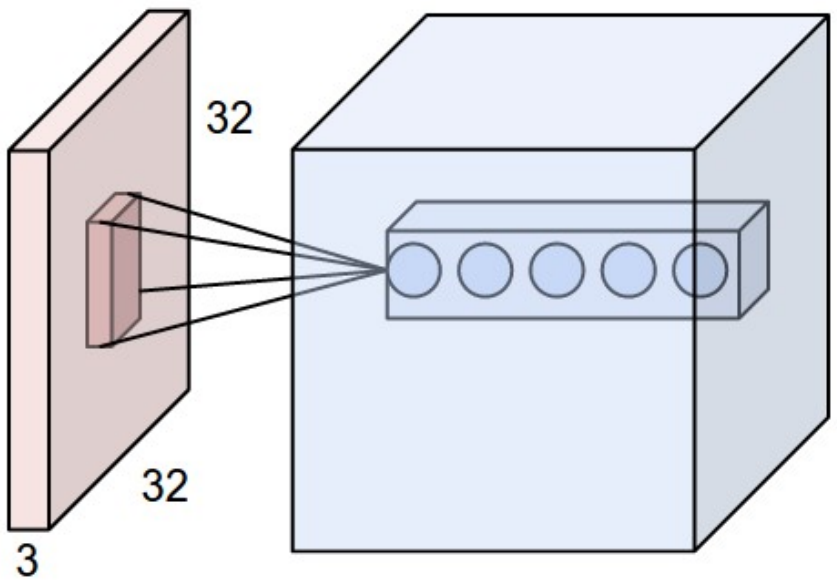
Feed an image into the network and output a list of most likely objects

Convolutional Neural Network

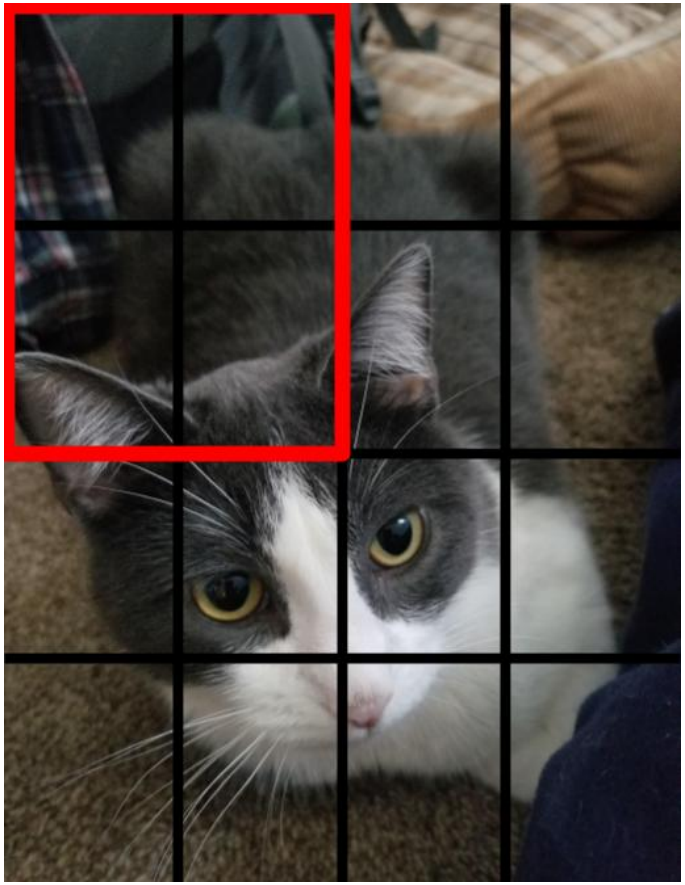
- Convolutional Neural Network (CNN) a more efficient type of neural network that is optimized for image classification
- Works similarly to a regular neural network except for its convolution layer and pooling layer

Convolution Layer

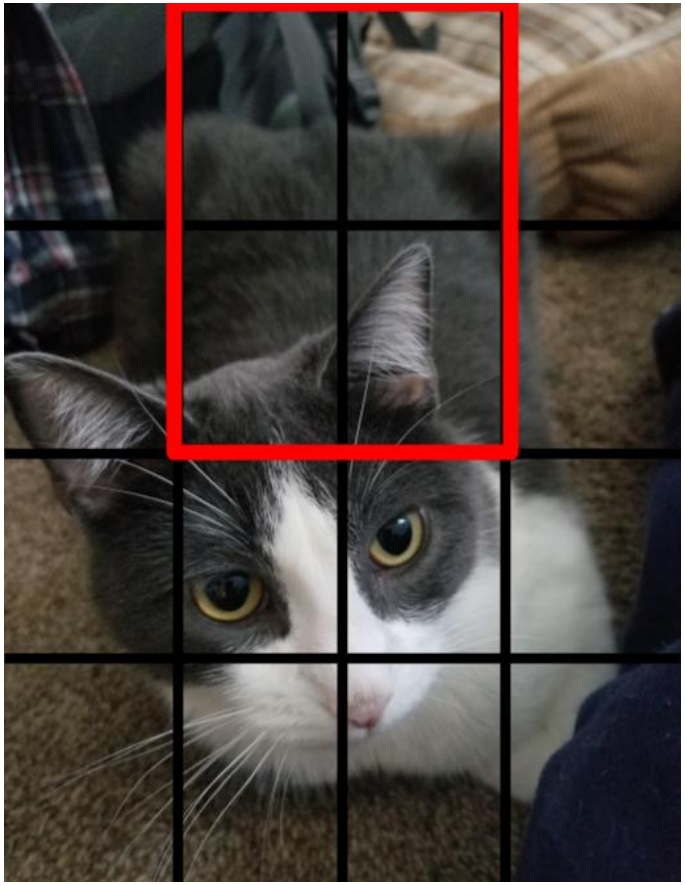
- Receptive fields
- Depth, stride



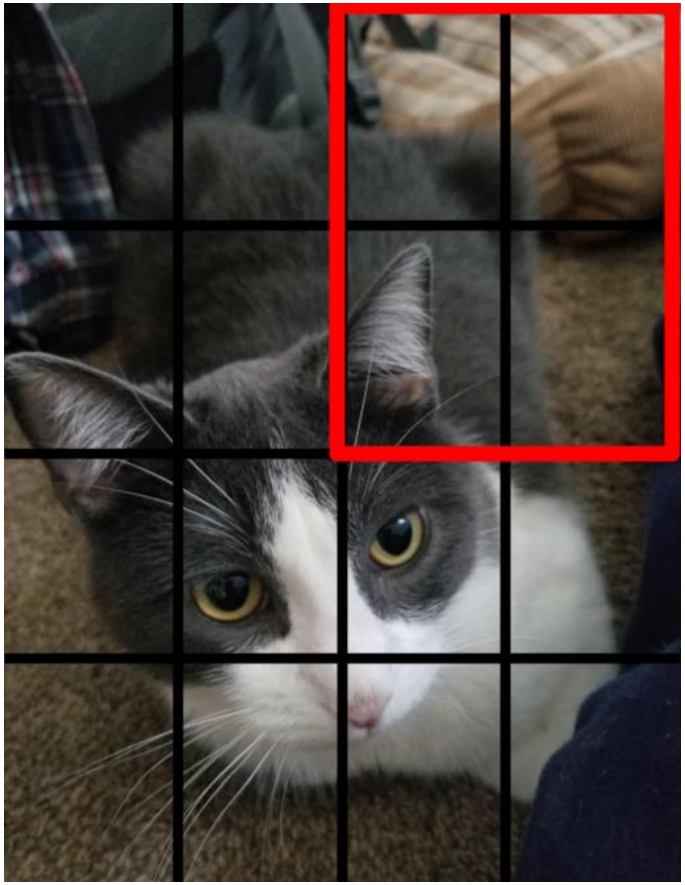
Object Classification Problem



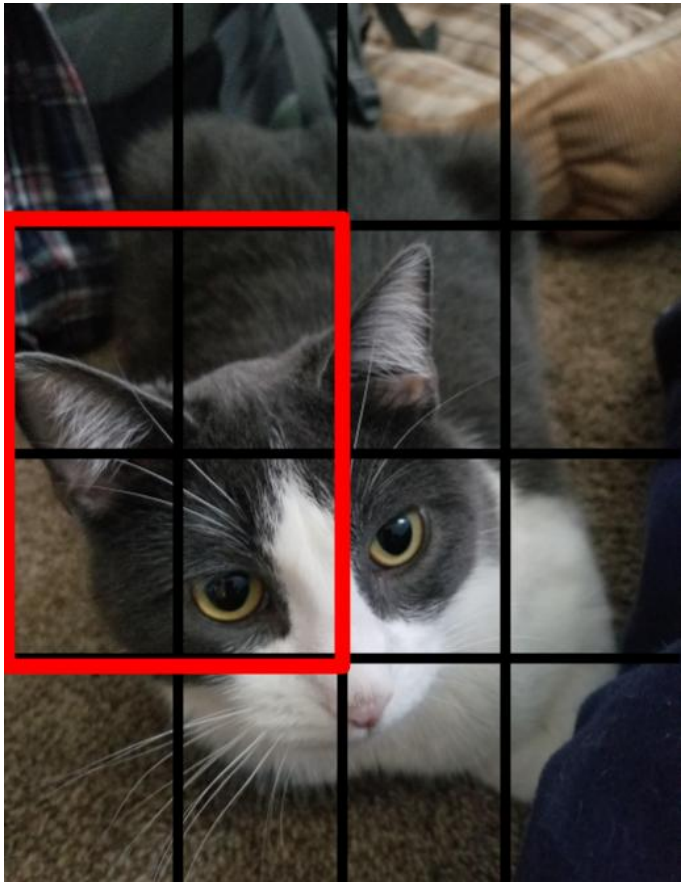
Object Classification Problem



Object Classification Problem



Object Classification Problem

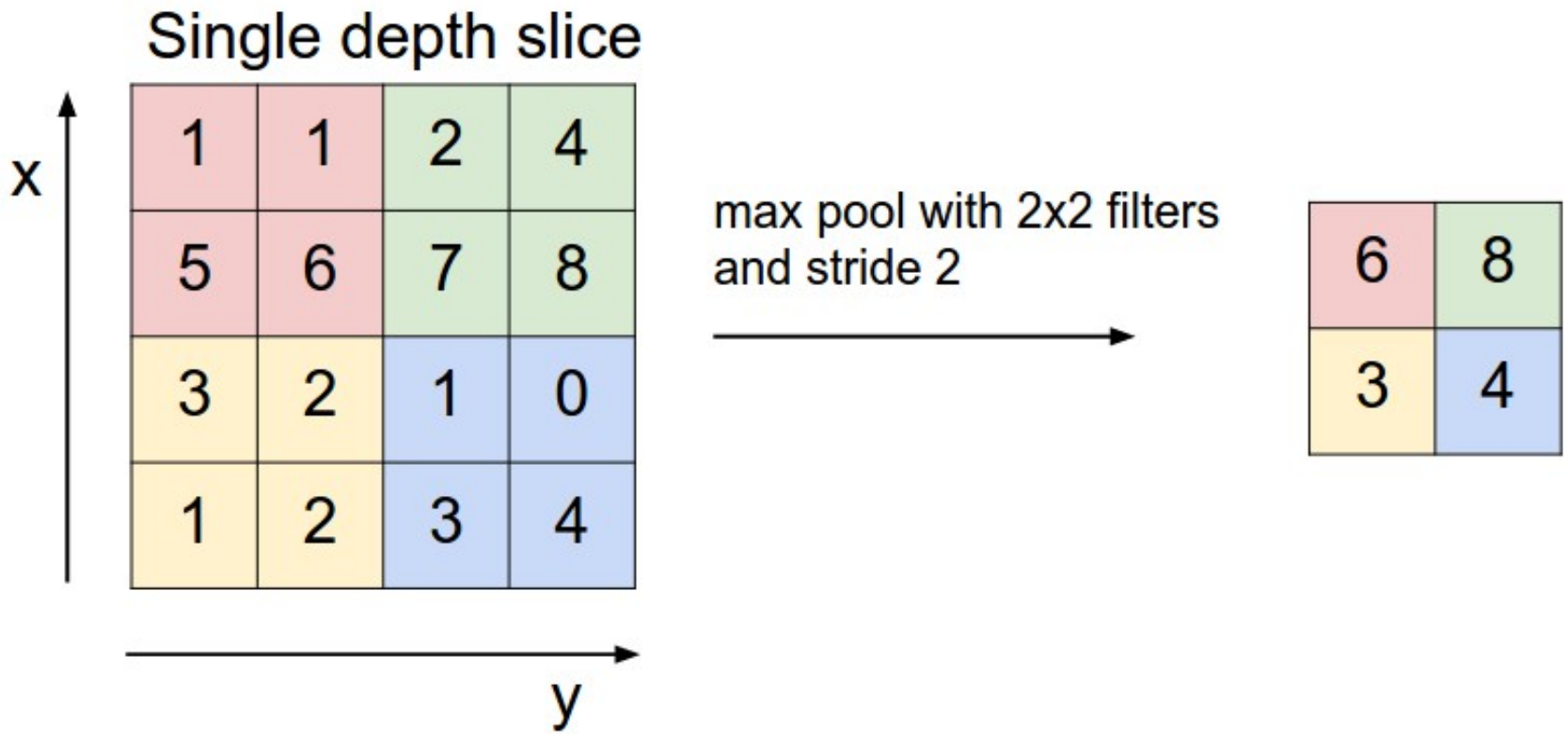


Activation Layer

- Introducing non-linearity into the network
- Most common activation function used is called Rectified Linear Unit (ReLU)

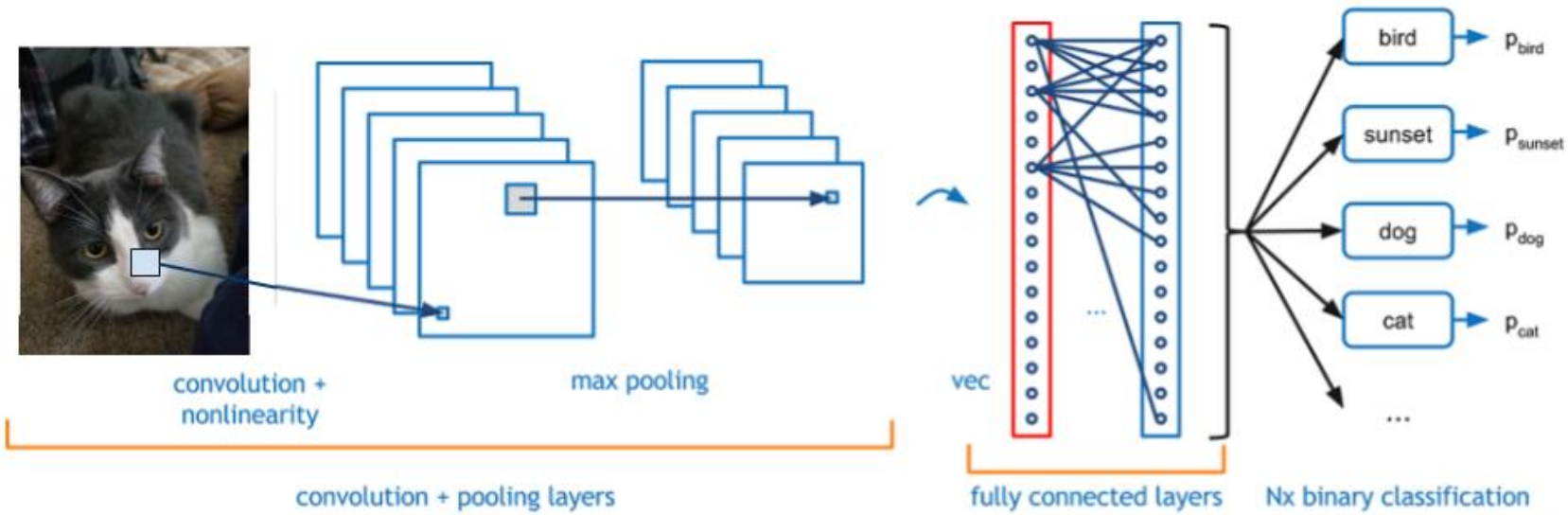
Pooling Layer

- Downsizing layer, reducing spatial size of the representation



Fully-Connected and Softmax Layer

- As the name implies, nodes in this layer have full connections to all nodes in the previous layer



Dropout

- Avoids overfitting
- Deactivate a certain percentage of nodes during the training phase
- Forces the network to learn more robust features

Fully-Convolutional Network (FCN)

- Modification of CNN architecture
- Replace all fully-connected layer(s) with backward convolution (up-sample or deconvolution) layer(s)
- FCN architecture is split into the following components:
  - Encoder
  - Decoder

Up-Sample Layer

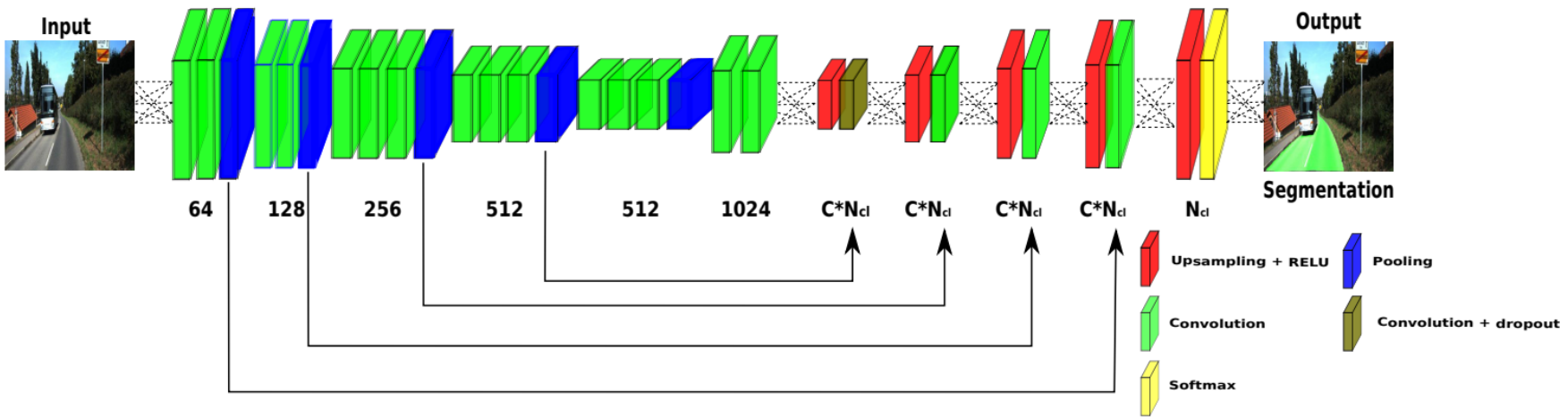
- Pooling layer aims to reduce the spatial dimension of input
- Up-sampling allows the output to match the spatial dimension of the input

Overall Goal

- FCNs, have seen a rise in popularity in road segmentation
- We are exploring two approaches to road segmentation
  - Up-Convolution Network
  - LIDAR only Deep Neural Network
- Comparing time and accuracy

Outline

Up-Convolutional Network

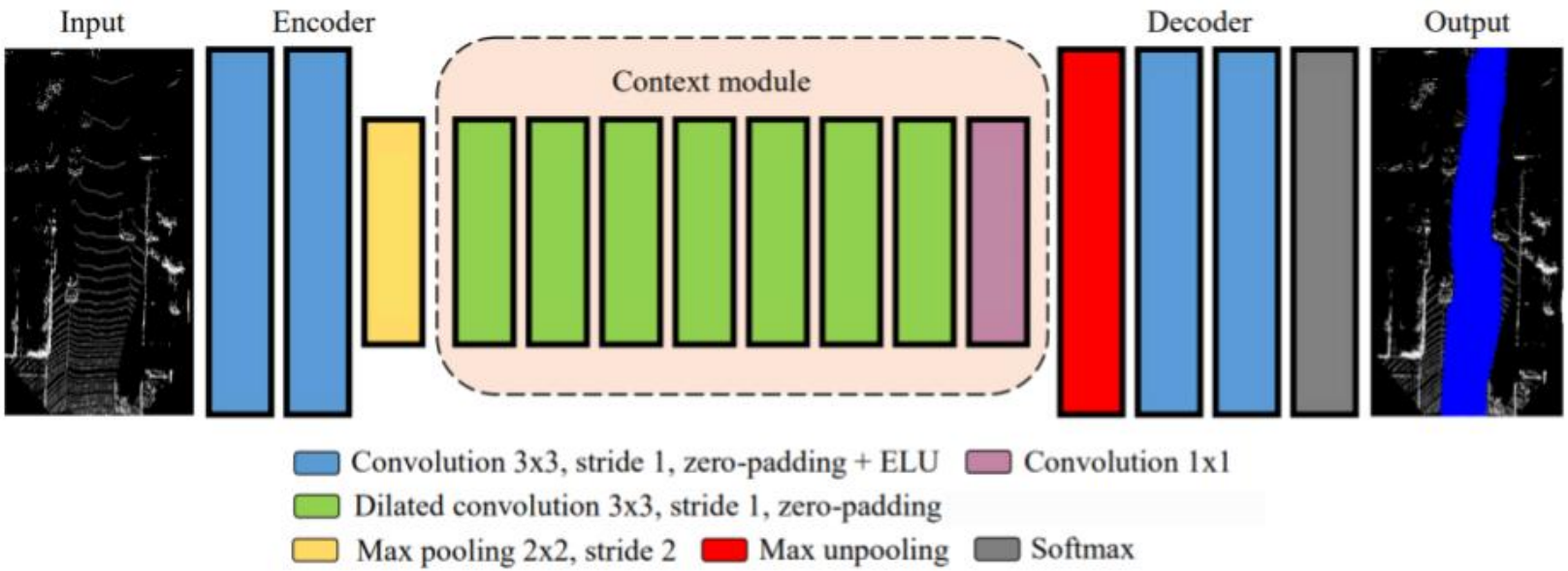


Outline

LIDAR only Deep Neural Network (LoDNN)

- Compress LIDAR point cloud into a grid in the x-y plane; some basic statistics are then computed for each grid cell
- Context module using dilated convolutions

LIDAR only Deep Neural Network

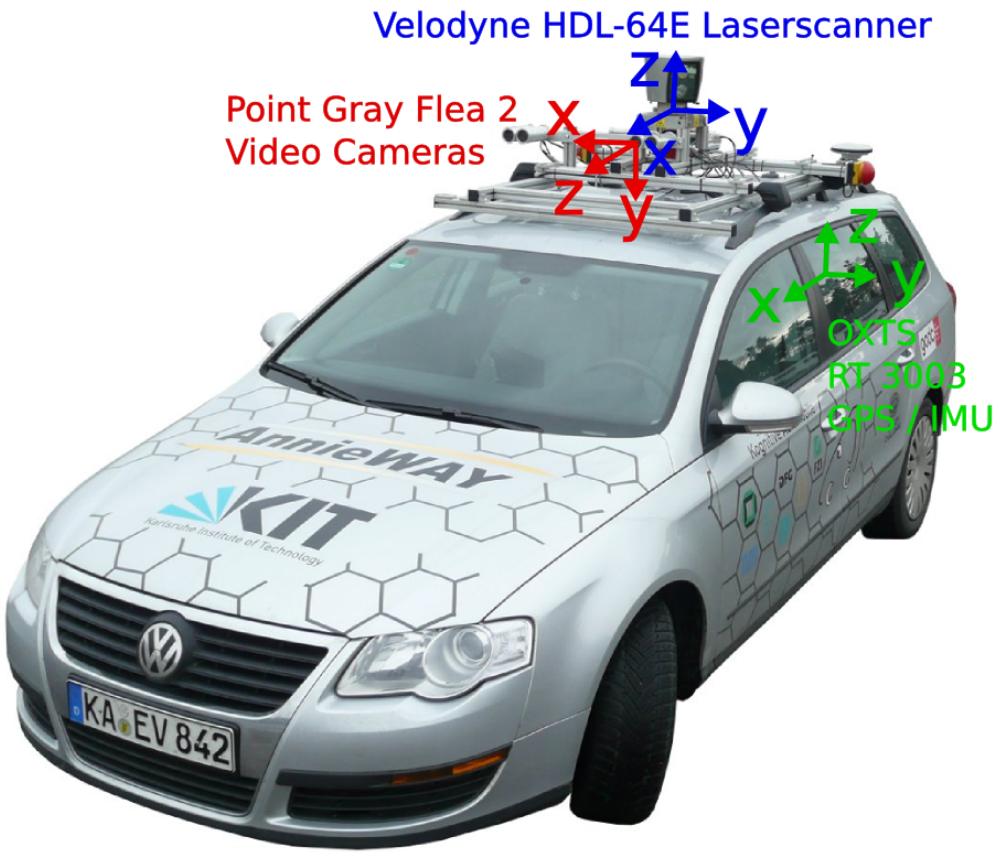


Outline

KITTI Road Benchmark

- Designed to benchmark road detection
- Manually segmented images are available for training
- Available in camera image and LIDAR point cloud format

KITTI Road Benchmark



Results and Conclusion

- Precision (PRE), recall (REC), and Maximum F1 Measure were used as metrics for evaluation

Method	Maximum F1 Measure	PRE	REC	Time (ms)
LoDNN	94.07	92.81	95.37	18
Up-Conv-Poly	93.83	94.00	93.67	80

Table: KITTI Road Benchmark Results (In %) On Urban Road Category

Acknowledgement

Thank you to my advisor Peter Dolan and Elena Machkasova for your guidance and feedback

Discussion

Questions?