MATLAB EXPO 2018

Introduction to Machine Learning and Deep Learning

Valerie Leung

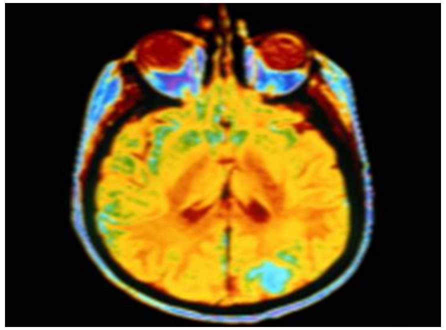
Machine learning in action

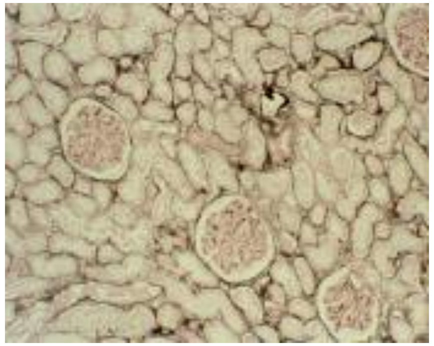
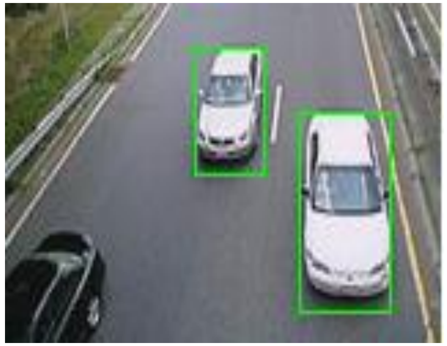
<u>CamVid Dataset</u>

1. Segmentation and Recognition Using Structure from Motion Point Clouds, ECCV 2008
2. Semantic Object Classes in Video: A High-Definition Ground Truth Database, Pattern Recognition Letters

Machine learning is everywhere

- Image recognition
- Speech recognition
- Stock prediction
- Medical diagnosis
- Predictive maintenance
- Language translation
- and more…



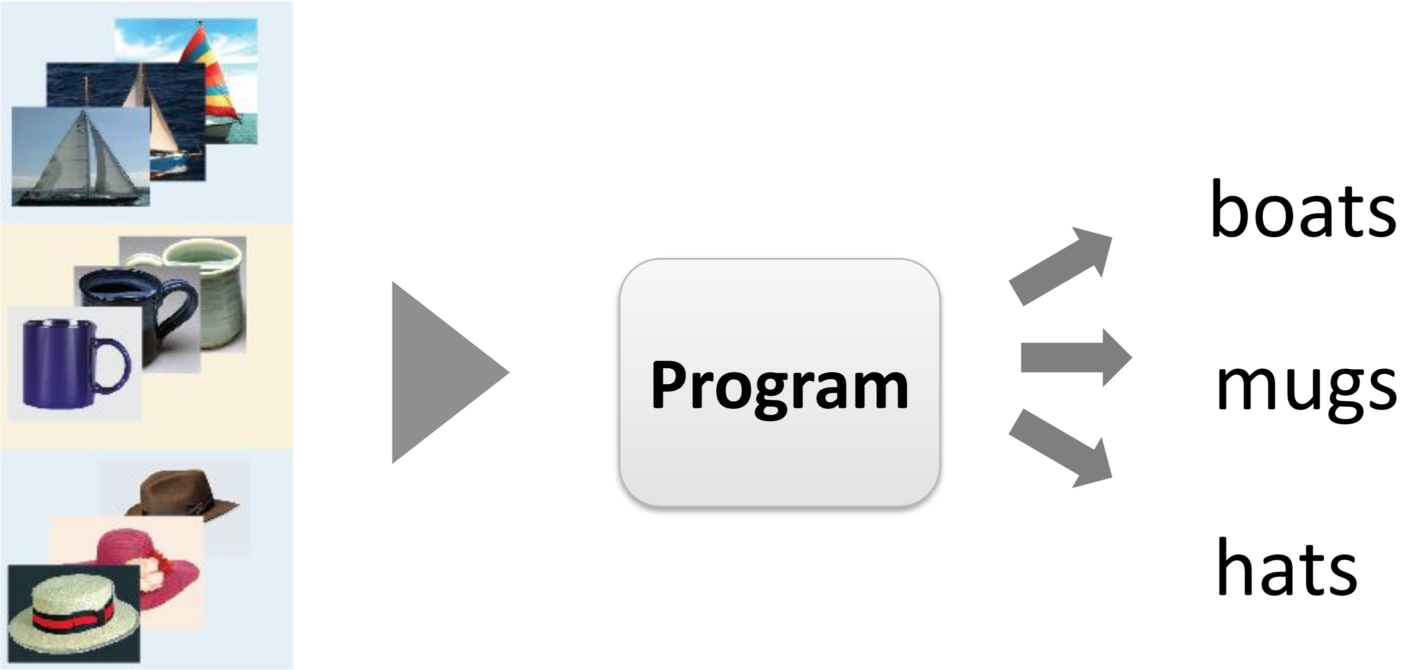
Agenda

- Machine Learning
  - What it is
  - Example : object classification
- Deep Learning
  - What it is and why it is popular
  - Object classification revisited
- Tackling time series with deep learning

What is machine learning?

Machine learning uses **data** and produces a **program** to perform a **task**

Task: Image Category Recognition



What is machine learning?

Machine learning uses **data** and produces a **program** to perform a **task**

Task: Image Category Recognition

Explicit Approach

Hand Written Program → boats / mugs / hats

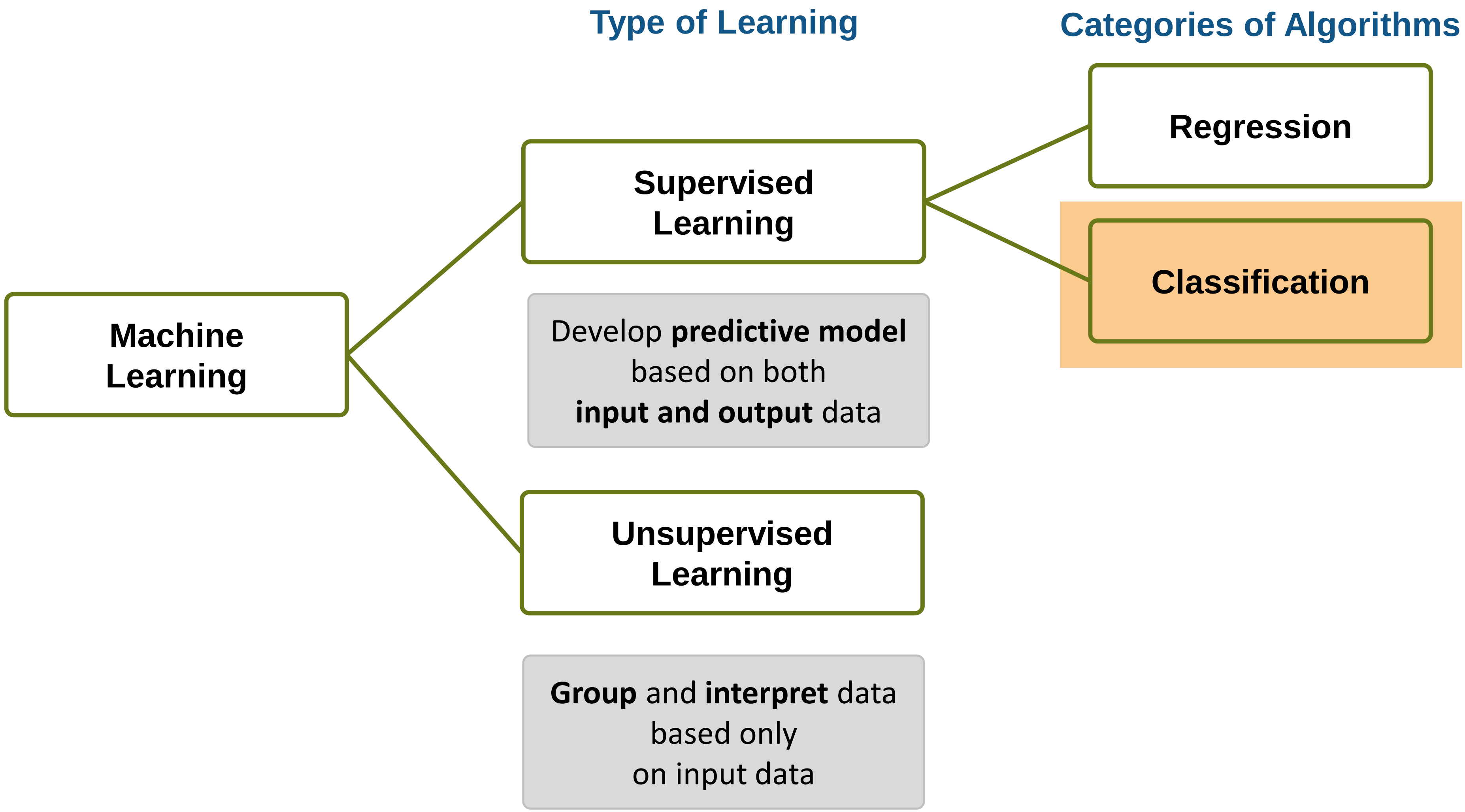
If **brightness** > 0.5
then 'hat'
If **edge_density** < 4 and **major_axis** > 5
then "boat"
…

Machine Learning Approach

Computer Vision → Machine Learning → boats / mugs / hats

Model = ⟨ Machine Learning Algorithm ⟩ (data, label)

Machine Learning: problem specific overview



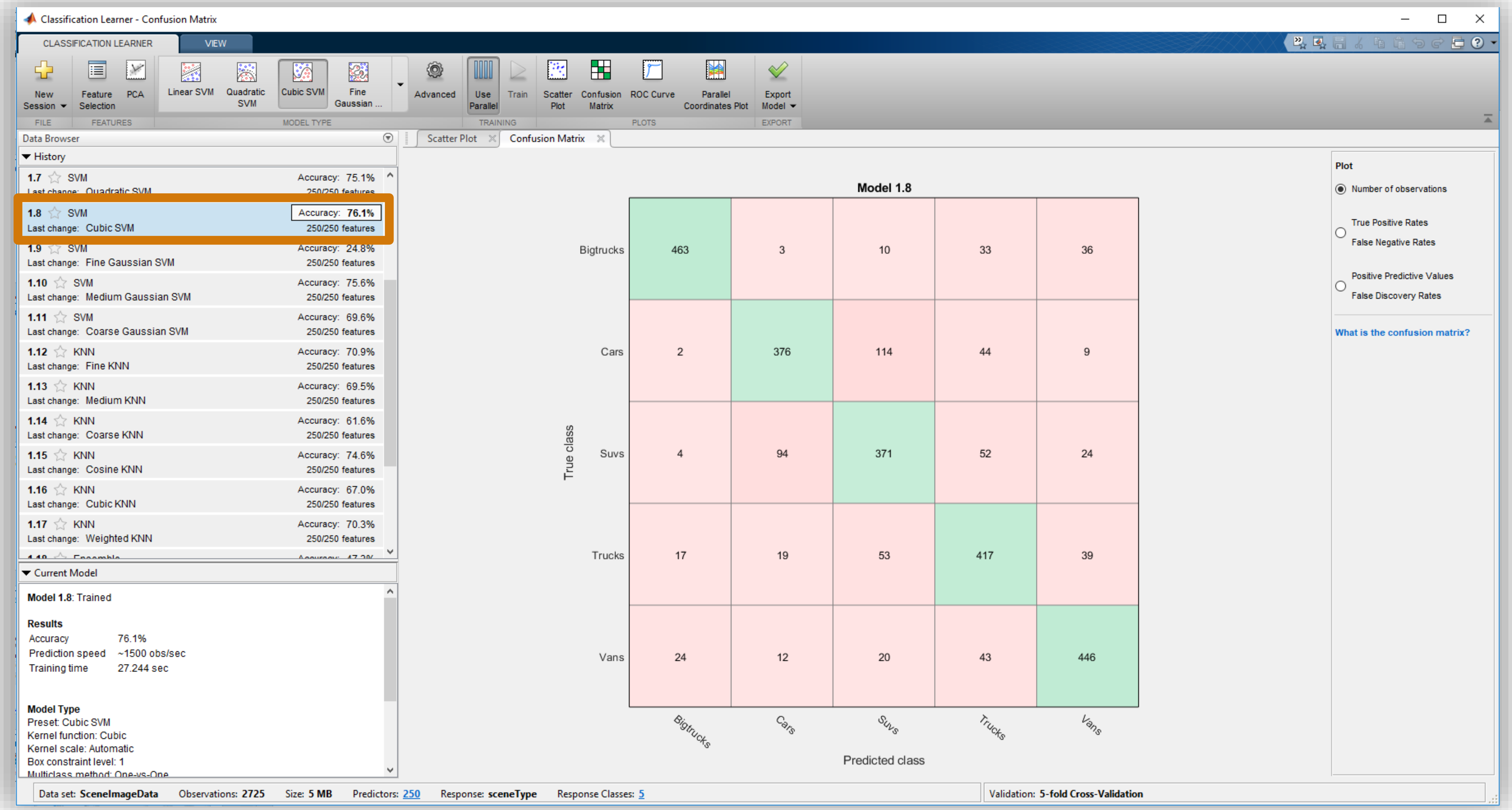
Example: Object Classification with Machine Learning

Task: Distinguish between 5 categories of vehicles

Machine Learning Model

- car
- suv
- pickup
- van
- truck

Using the Classification Learner App to determine the best model

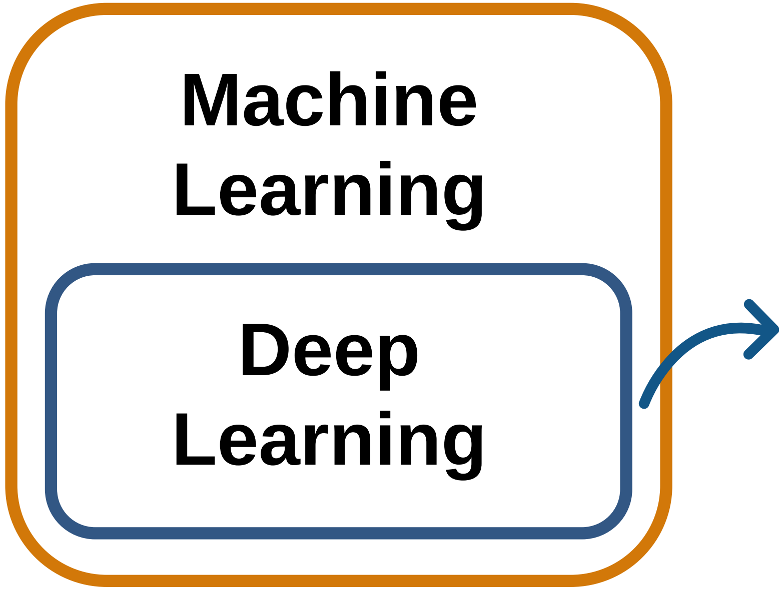


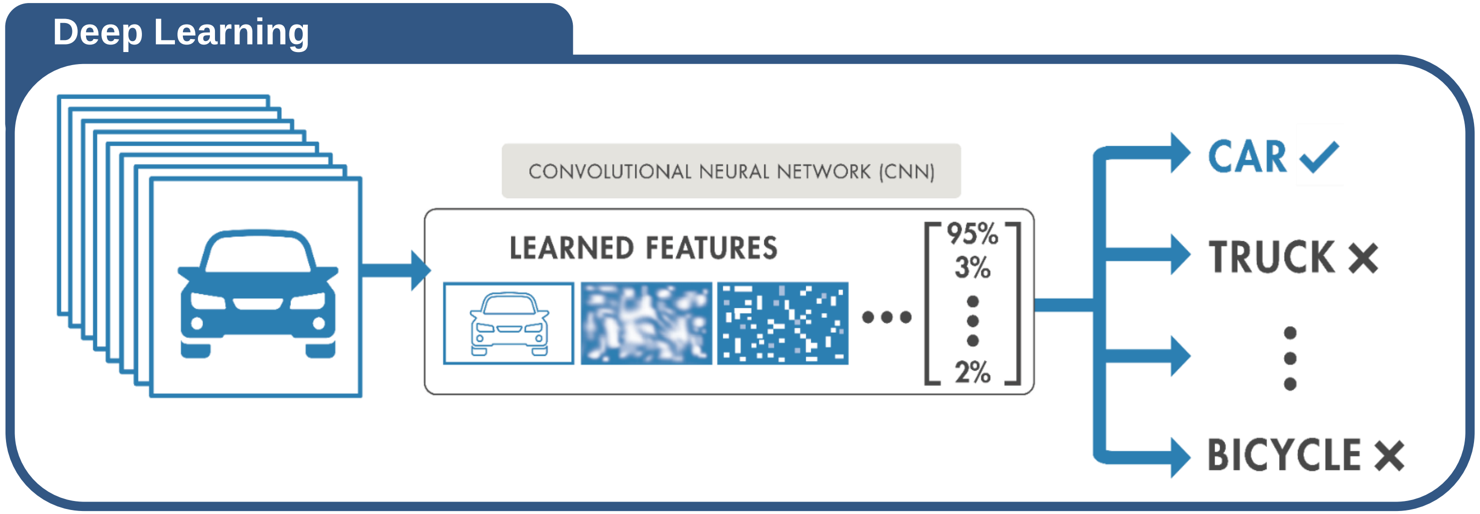
Agenda

- Machine Learning
  - What it is
  - Example : object classification
- Deep Learning
  - What it is and why it is popular
  - Object classification revisited
- Tackling time series with deep learning

Deep Learning for Classification
What is Deep Learning ?

- Subset of machine learning with automatic feature extraction
  - Learns features and tasks directly from data
  - More data = better model





Deep Learning for Classification

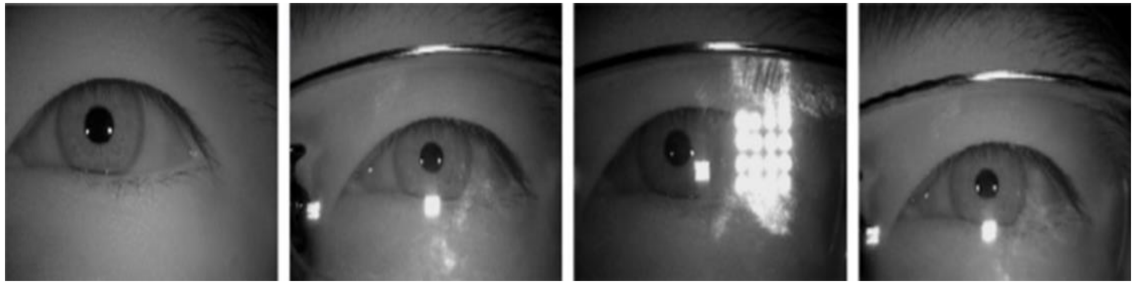
Applications: autonomous driving, image classification, etc.

Detection of cars and road in autonomous driving systems

Single Image Age Estimation[3,4]

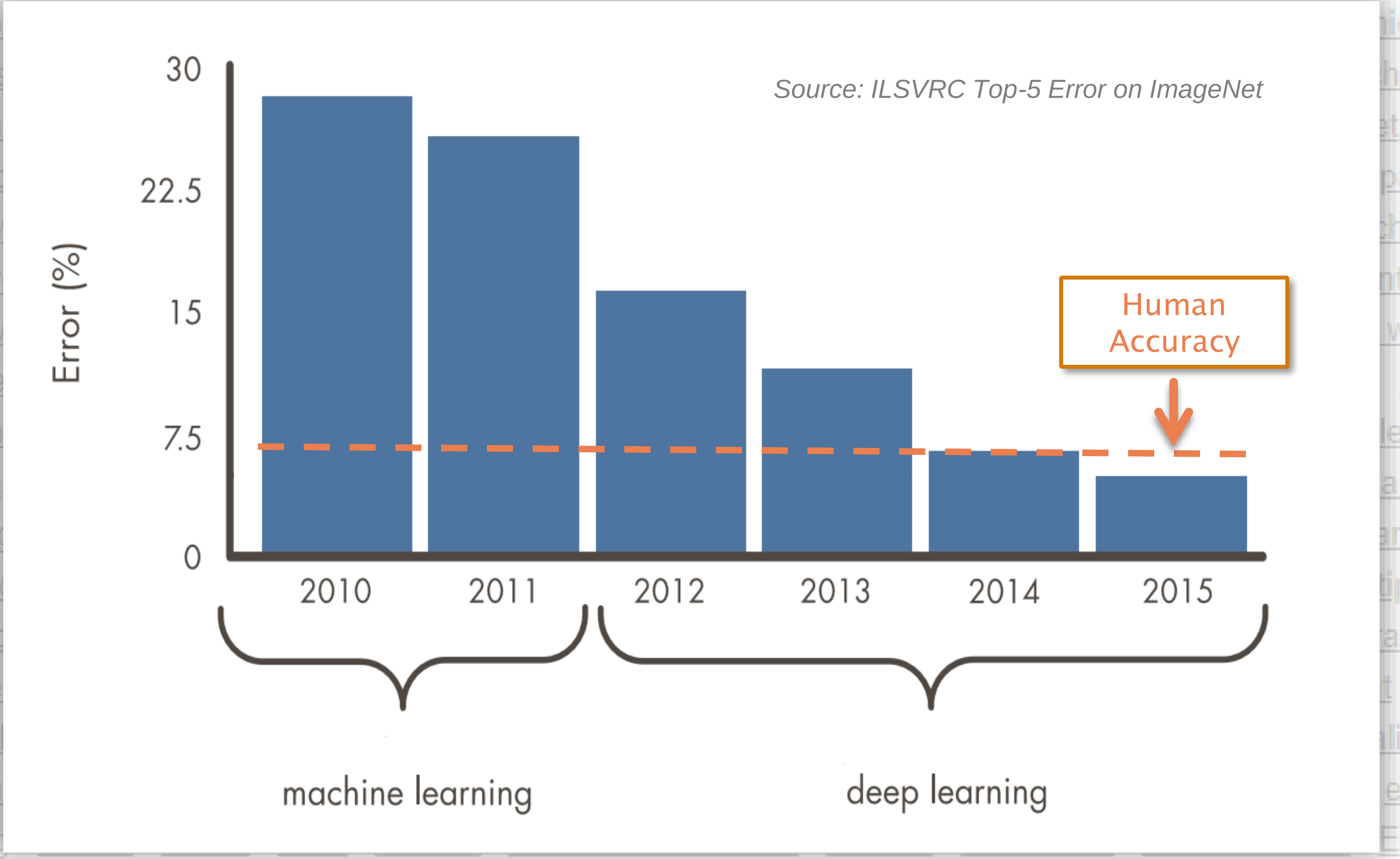
Rain Detection and Removal[1]

Iris Recognition – 99.4% accuracy[2]



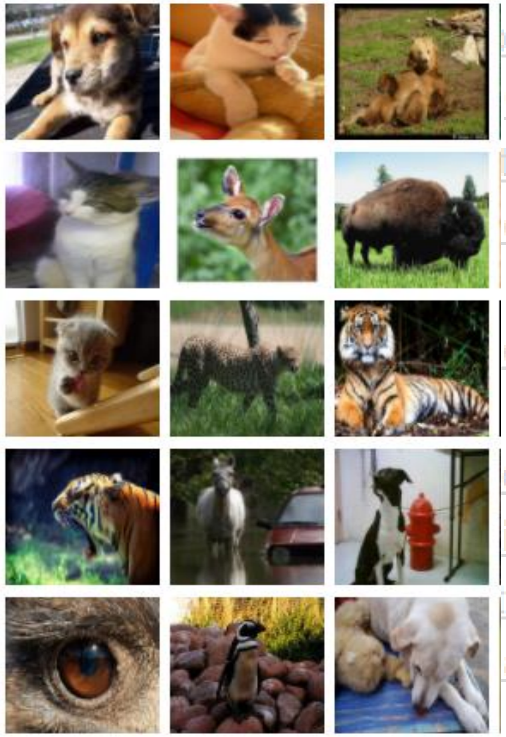
1. *"Deep Joint Rain Detection and Removal from a Single Image" Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan*
2. *Source: An experimental study of deep convolutional features for iris recognition Signal Processing in Medicine and Biology Symposium (SPMB), 2016 IEEE Shervin Minaee ; Amirali Abdolrashidiy ; Yao Wang; An experimental study of deep convolutional features for iris recognition*
3. *"DEX: Deep Expectation of apparent age from a single image", Rasmus Rothe, Radu Timofte and Luc Van Gool, Looking at People Workshop, ICCV 2015*
4. *Image source : https://en.wikipedia.org/wiki/Emmanuel_Macron*

Deep learning models can surpass human accuracy



IMAGENET[1]

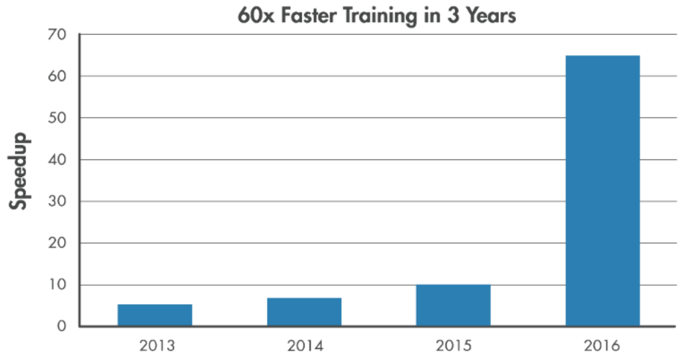
Database for visual object recognition research



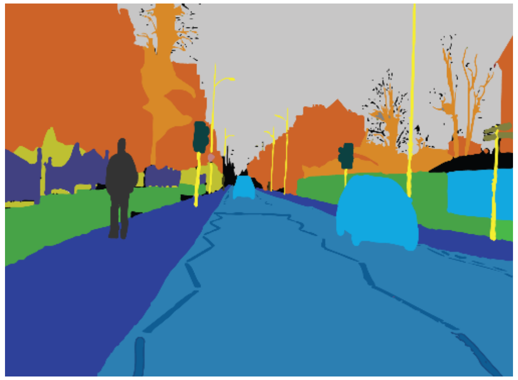
1. http://www.image-net.org/

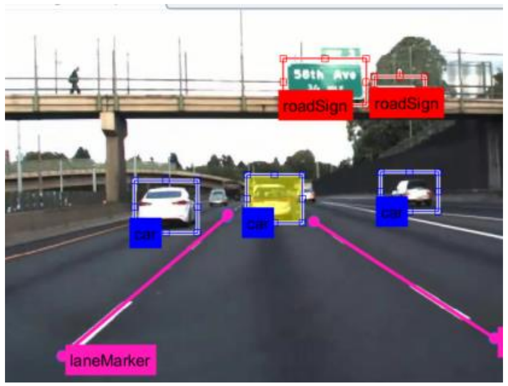
Deep Learning for Classification
Deep Learning Enablers

Increased GPU acceleration



Labeled public datasets





World-class models to be leveraged

AlexNet
PRETRAINED MODEL

VGG-16
PRETRAINED MODEL

ResNet
PRETRAINED MODEL

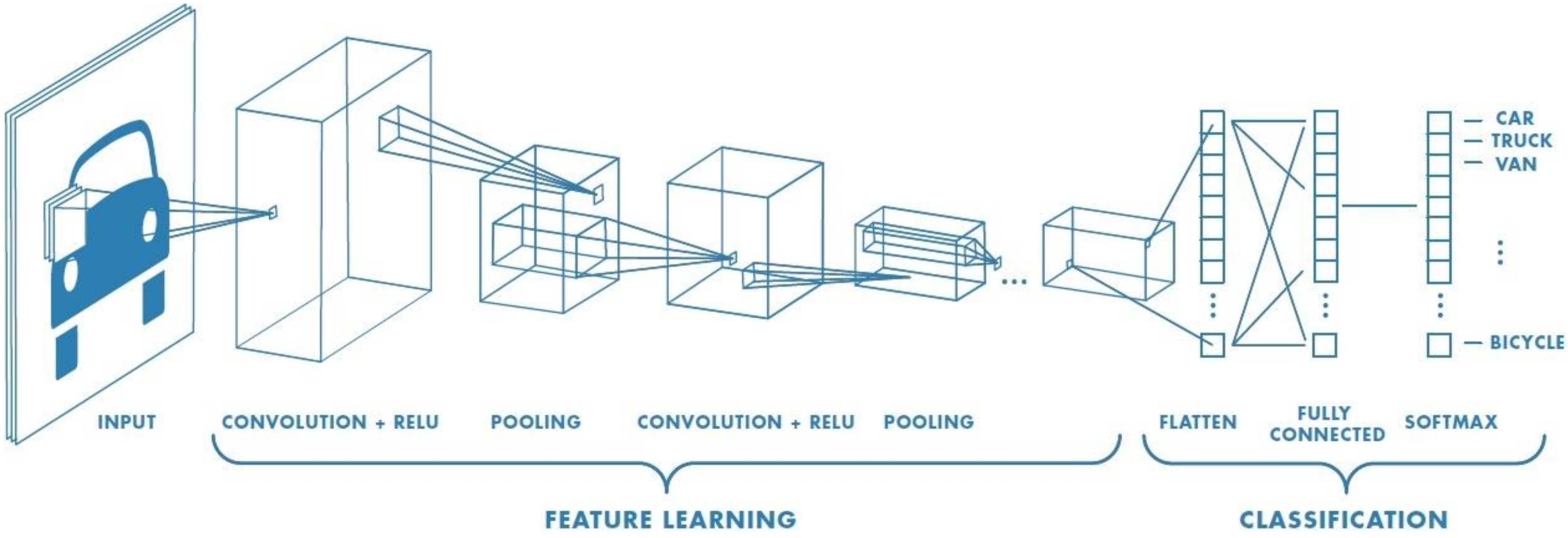
Caffe
MODELS

GoogLeNet
PRETRAINED MODEL

TensorFlow/Keras
MODELS

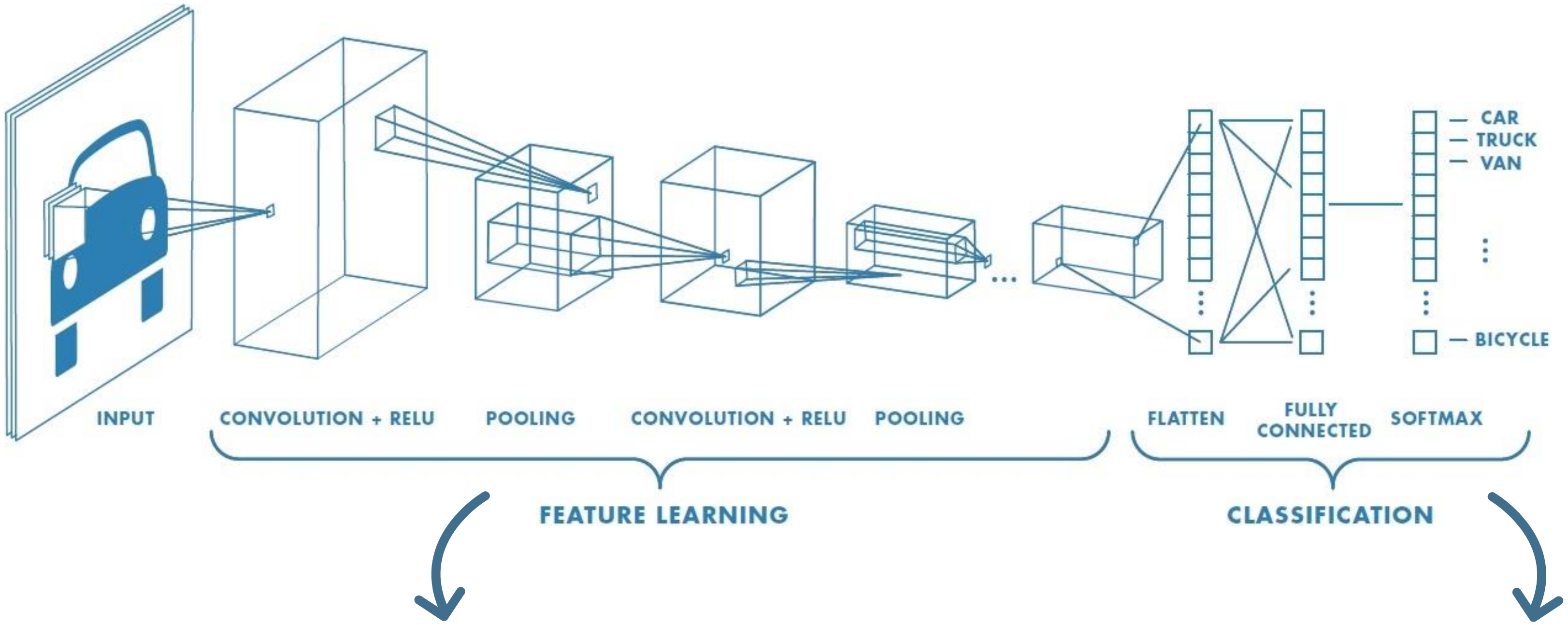
Deep Learning for Classification

What does a Convolutional Neural Network (CNN) architecture look like?



Deep Learning for Classification

What does a Convolutional Neural Network (CNN) architecture look like?



Convolution puts the input images through a set of convolutional filters, each of which activates certain features from the images.

Rectified linear unit (ReLU) allows for faster and more effective training by mapping negative values to zero and maintaining positive values.

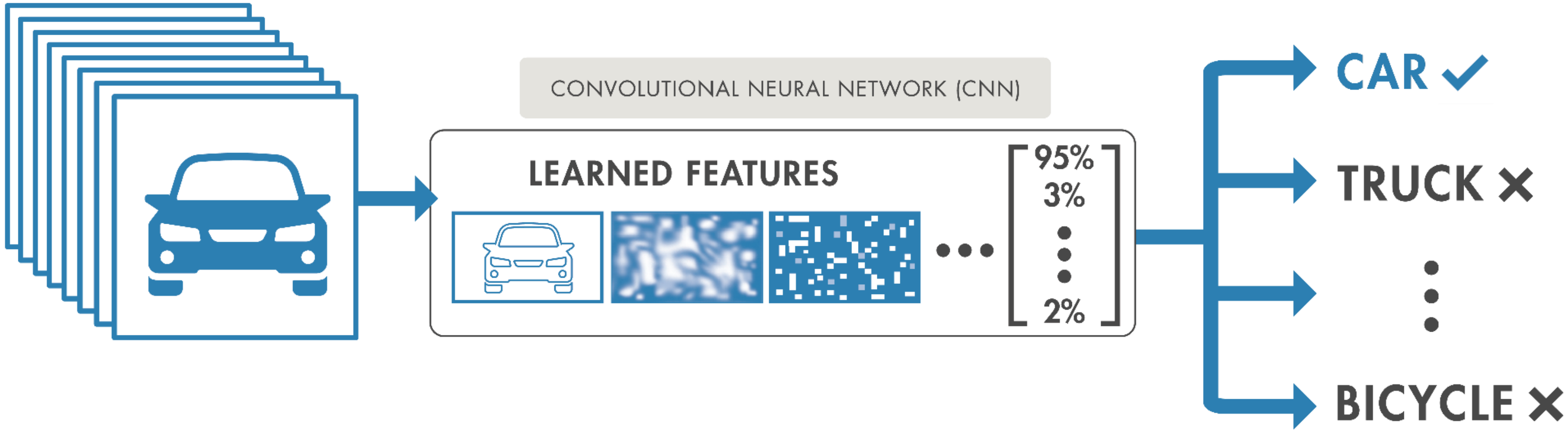
Pooling simplifies the output by performing nonlinear downsampling, reducing the number of parameters that the network needs to learn about.

The next-to-last layer is a **fully connected layer** (FC) that outputs a vector of K dimensions where K is the number of classes that the network will be able to predict. This vector contains the probabilities for each class of any image being classified.

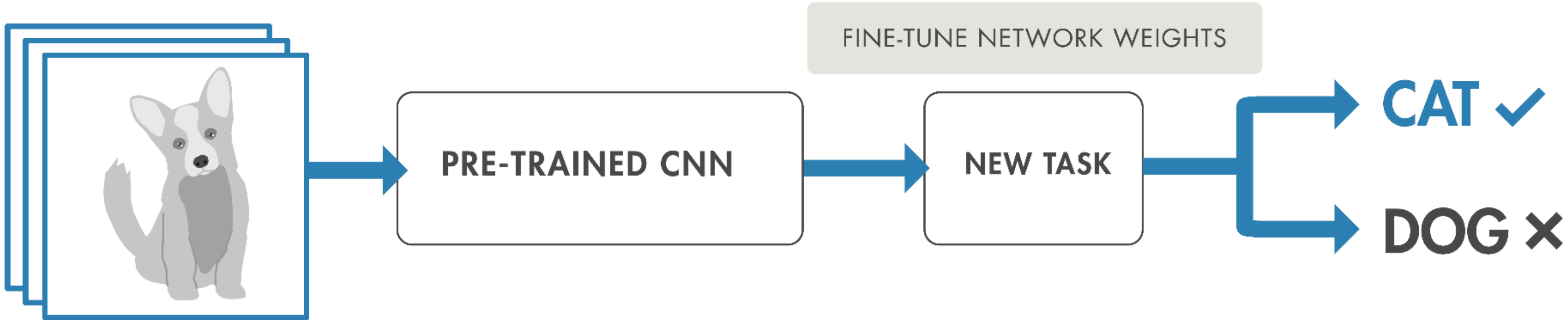
The final layer of the CNN architecture uses a **softmax** function to provide the classification output.

Deep Learning for Classification
Two approaches

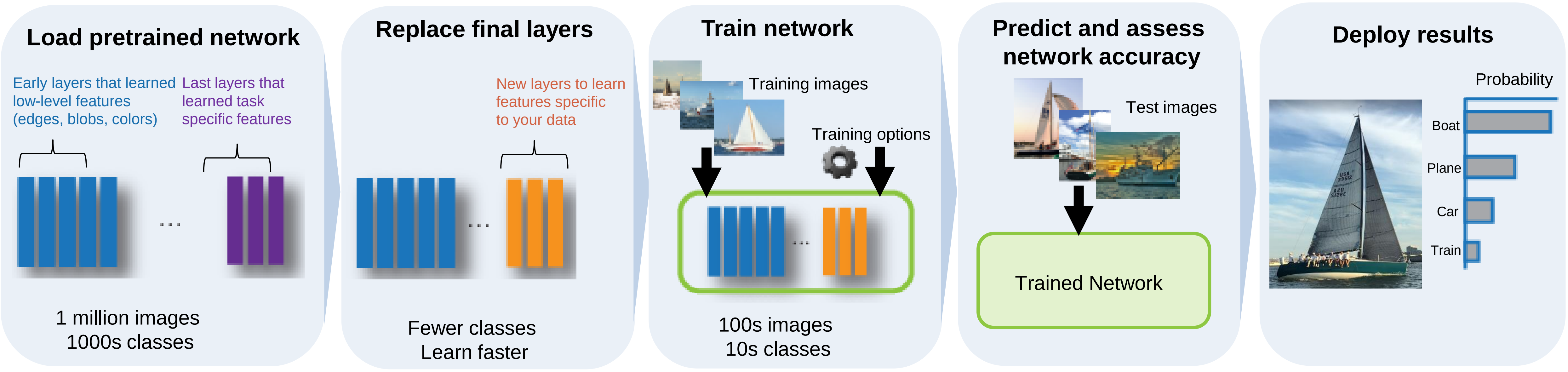
1. Train a Deep Neural Network from Scratch



2. Fine-tune a pre-trained model (transfer learning)



Deep Learning for Classification
Transfer Learning workflow



Example: Object Classification revisited with Transfer Learning

Convolution → Activation → Pooling → Convolution → Activation → Pooling → Convolution → Activation → Pooling … Convolution → Activation → Pooling → Fully Connected Layers → 1000 Category Classifier

AlexNet

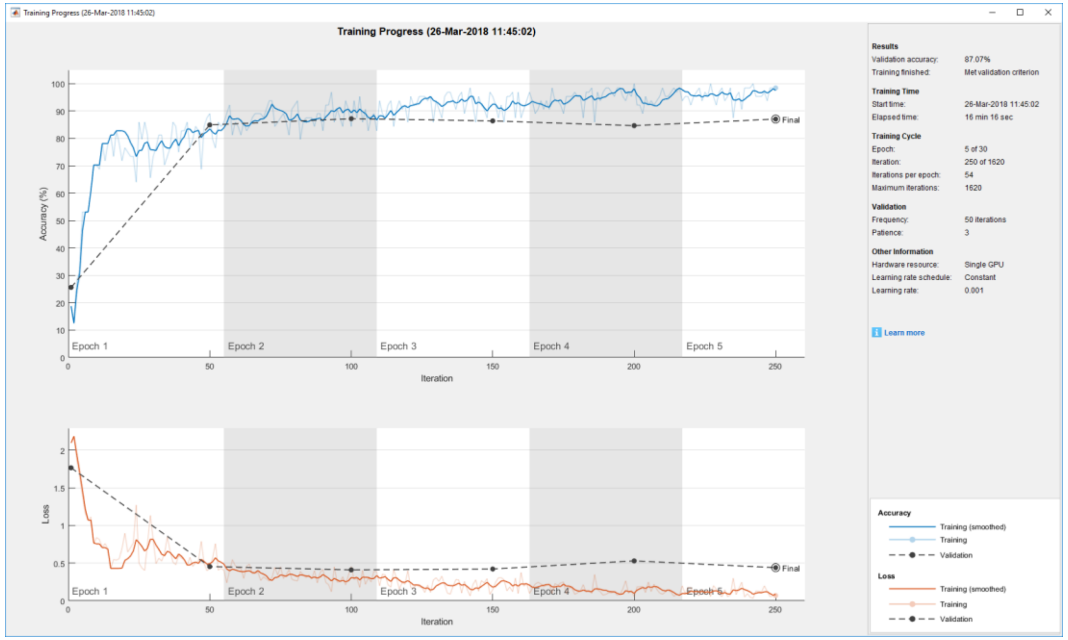
AlexNet

car
suv
pickup
van
truck

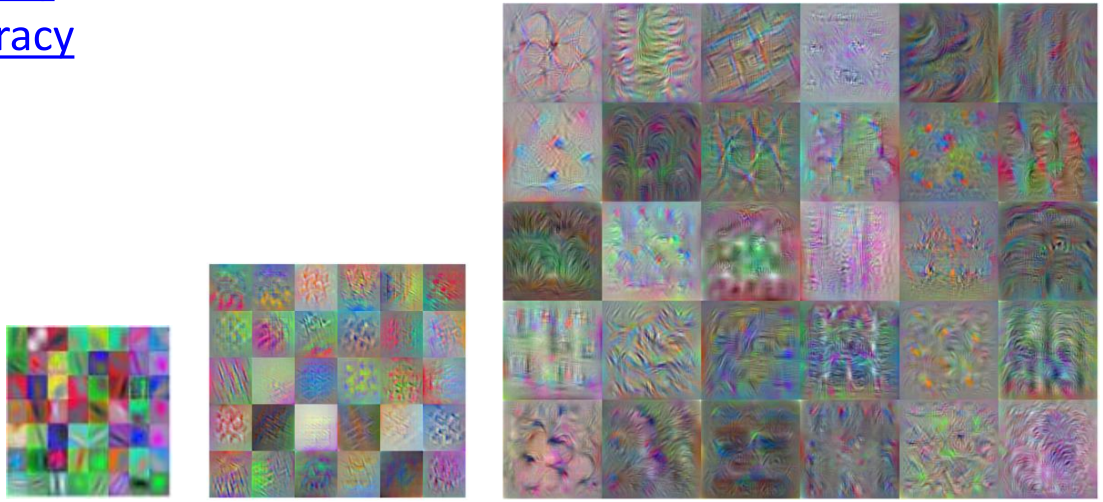
5 Category Classifier

New Data

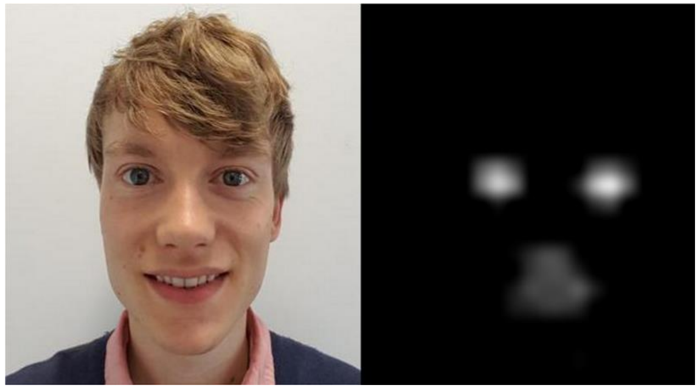
How do I visualize and debug deep neural networks?



Training Accuracy Plot



Feature Visualization



Network Activations

Deep Dream

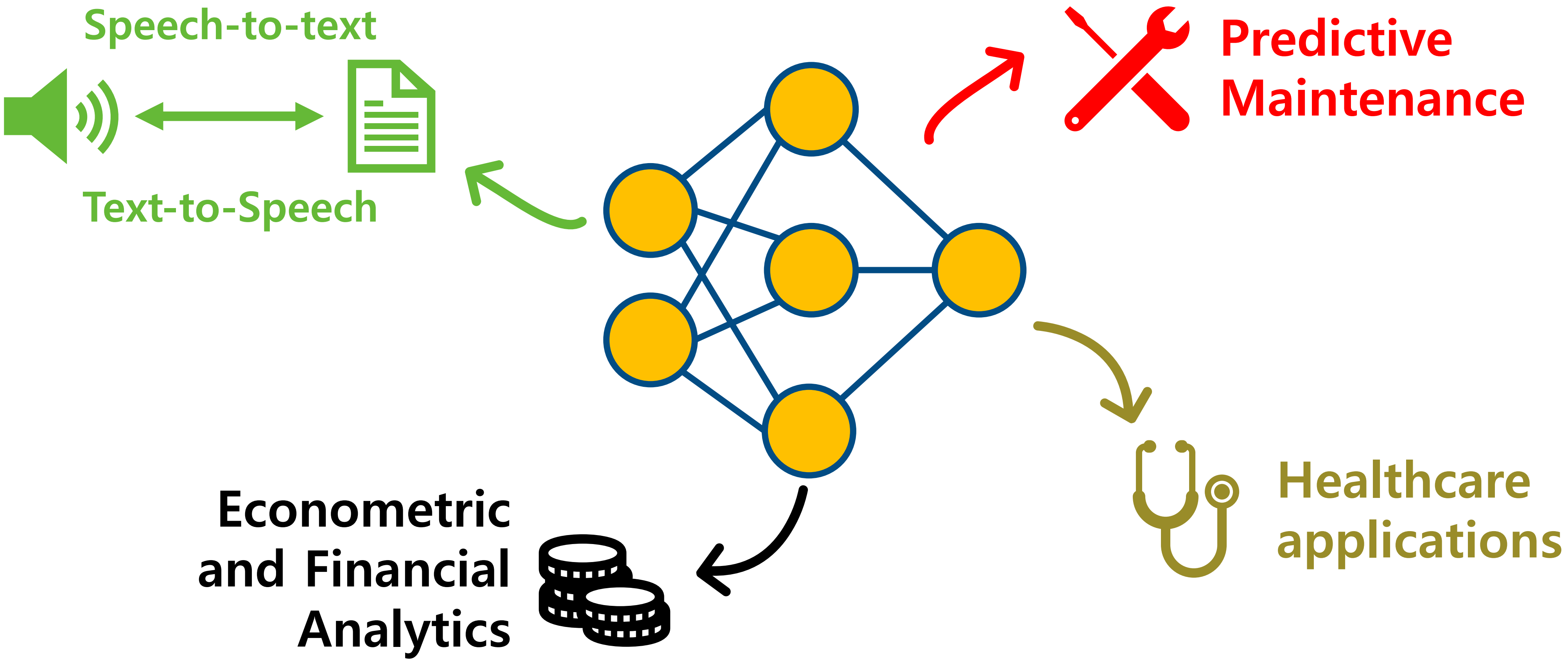
Recap: Deep Learning for Images

- Deep learning: end-to-end training including automatic feature extraction
- Two approaches:
  - Train from scratch - requires lots of labeled training data
  - Transfer learning – leverage existing models and adapt to new task
- Visualize and debug networks with built-in functions

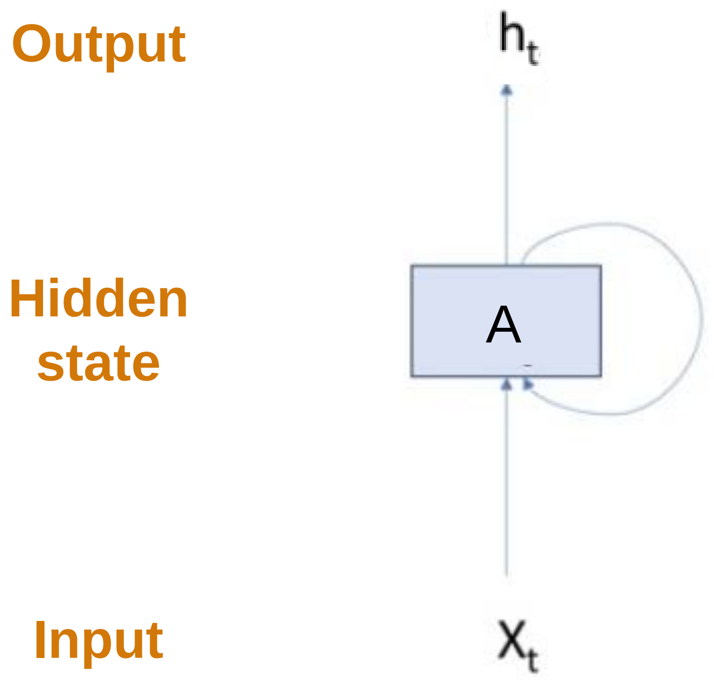
Agenda

- Machine Learning
  - What it is
  - Example : object classification
- Deep Learning
  - What it is and why it is popular
  - Object classification revisited
- Tackling time series with deep learning

Deep Learning for Time Series



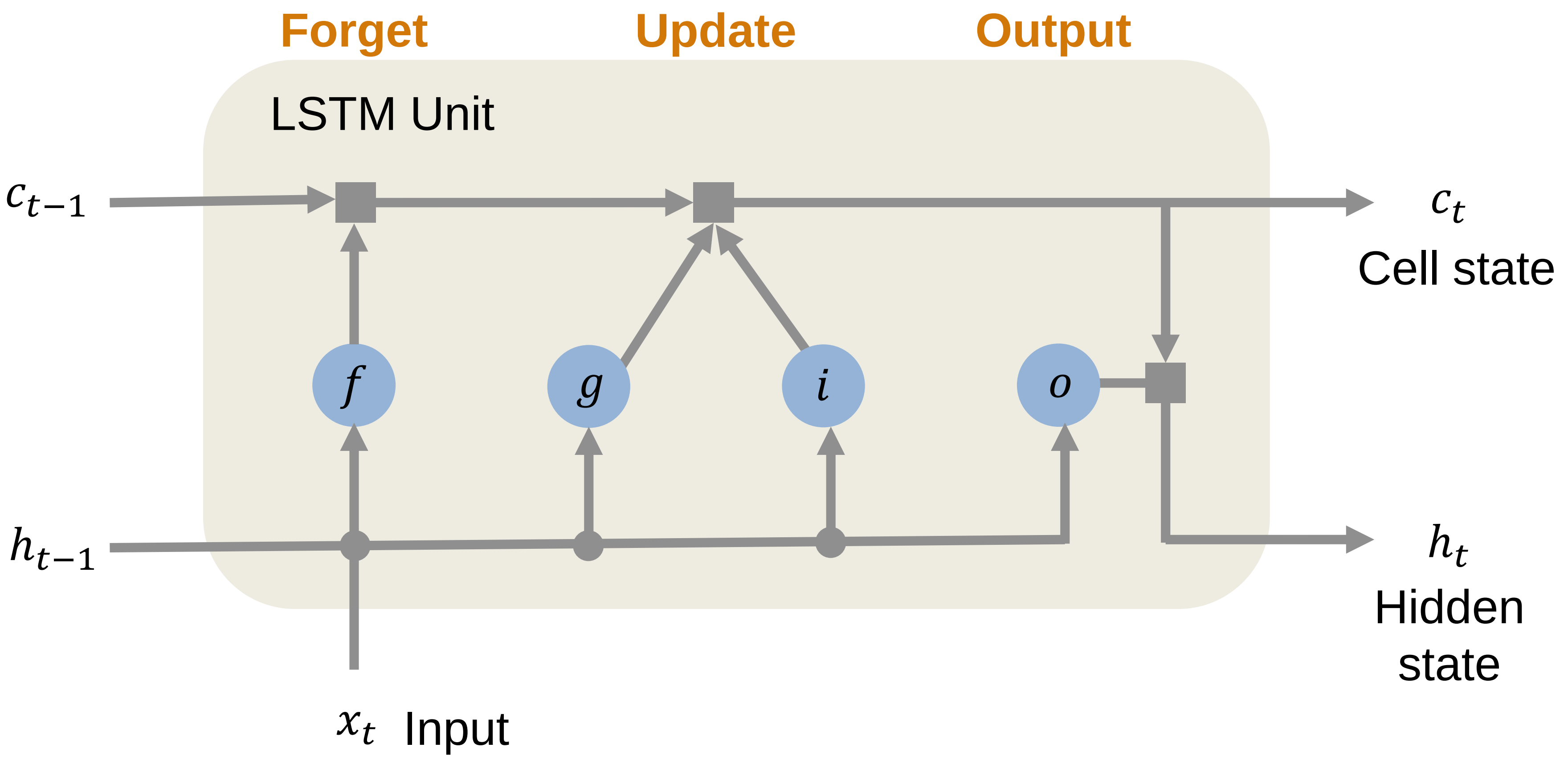
Basic Network: Recurrent Neural Network (RNN)



- Problems
  - Vanishing gradients: only short-term dependencies are captured – information from earlier time steps decays
  - Exploding gradients: error can grow drastically with each time step

Long Short-Term Memory (LSTM)

Selectively retains relevant information and forgets irrelevant information



f : Forget gate

g : Memory gate

i : Input gate

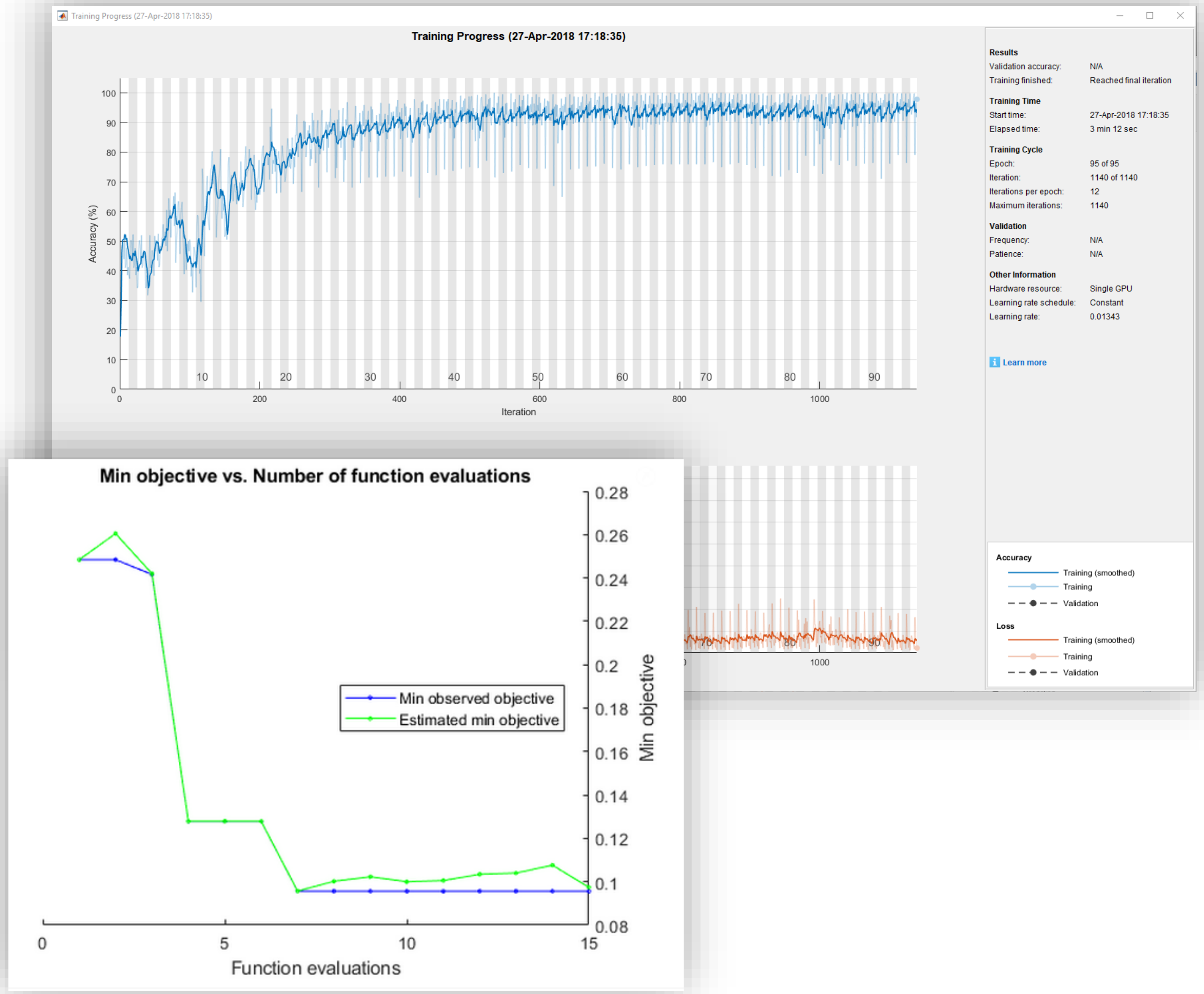
o : Output gate

Example: Human Activity Recognition from Mobile Phone Data

Objective: Train a classifier to classify human activity from sensor data

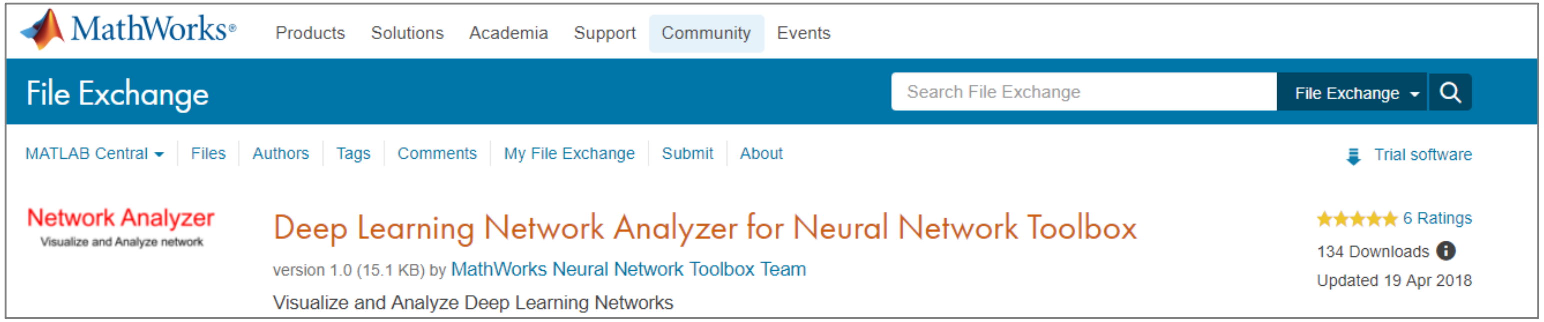
Data:

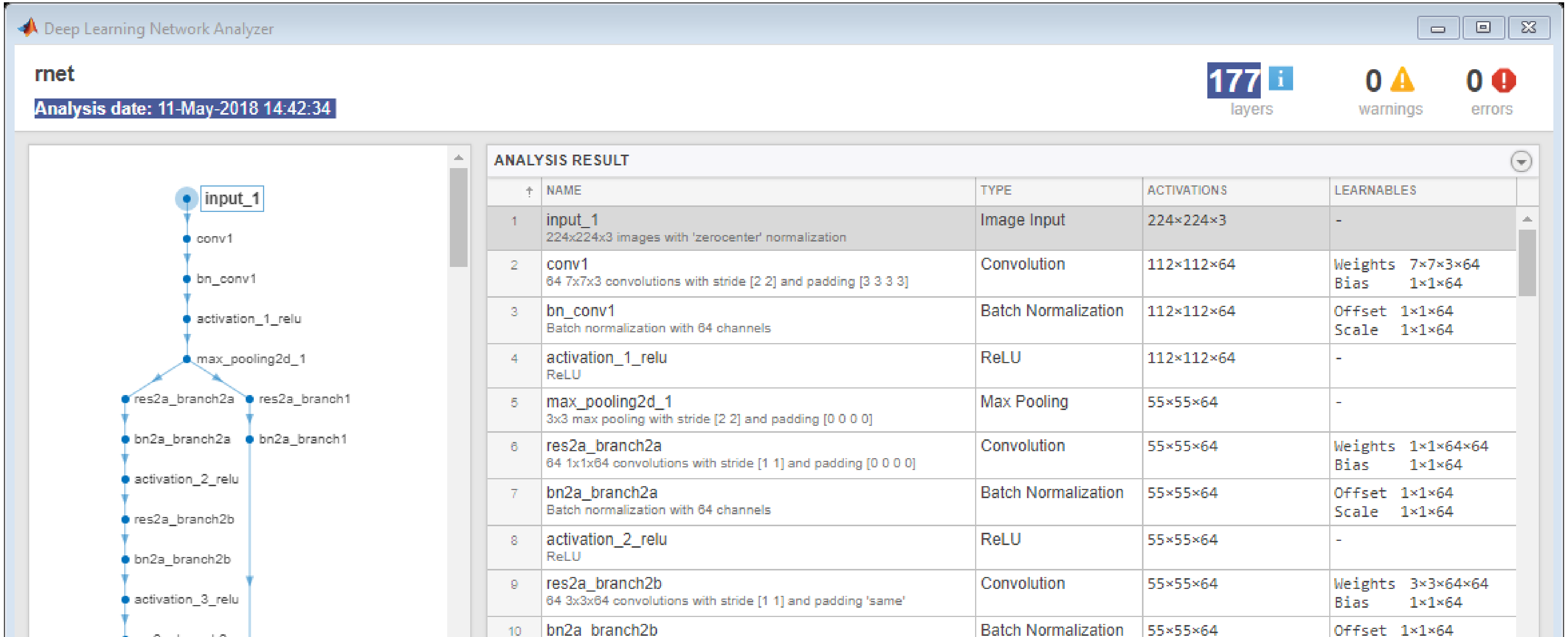
Predictors	3-axial Accelerometer and Gyroscope data 
Response	Activity: 



Data source: https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones

How do I know if my deep network is defined correctly?





Recap – Deep Learning for time series

- LSTMs good for handling temporal dependence

- Define with `lstmlayer` and `bilstmlayer`

Key Takeaways : Why use MATLAB for Deep Learning

- Handle and label large sets of data
  - Handle images with **`imageDatastore`**
- Accelerate deep learning with GPUs on desktops, clusters, and clouds
- Visualize and debug deep neural networks
- Access and use models from experts