Deep Learning for Images and Video

Bo Luan
Application Engineer
boluan@mathworks.com

Sharon Kim
Application Engineer
sharonk@mathworks.com

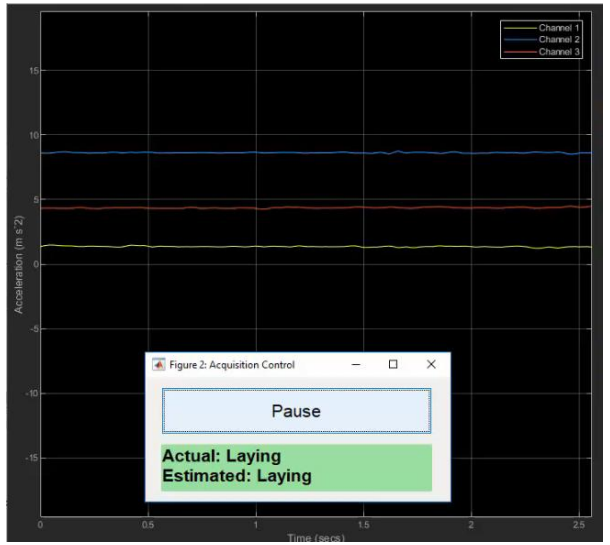
Images

Object Localization and Detection

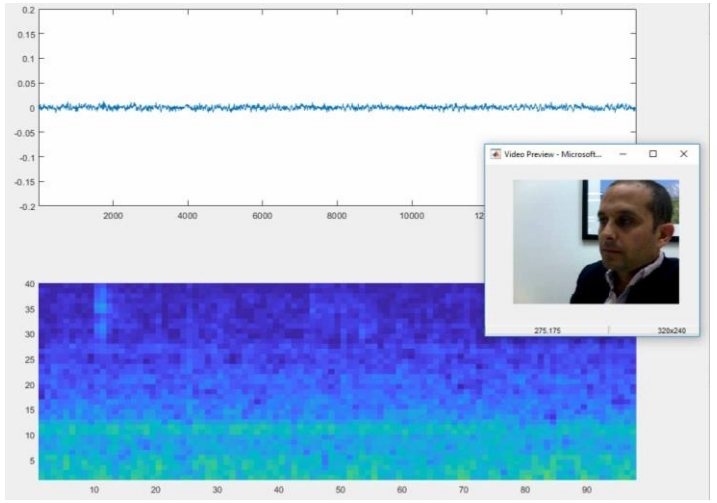
Semantic Segmentation

Signals

Signal Classification

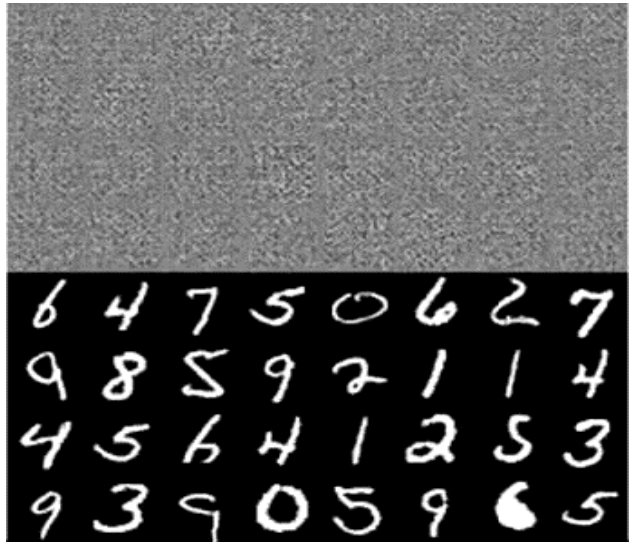


Speech Command Recognition

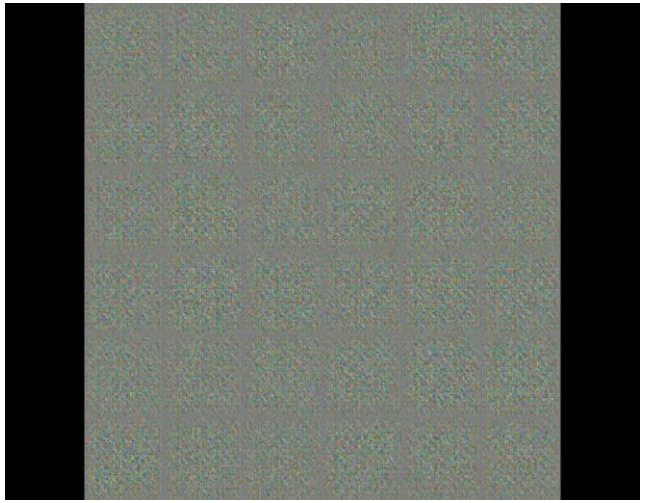


Data Synthesis

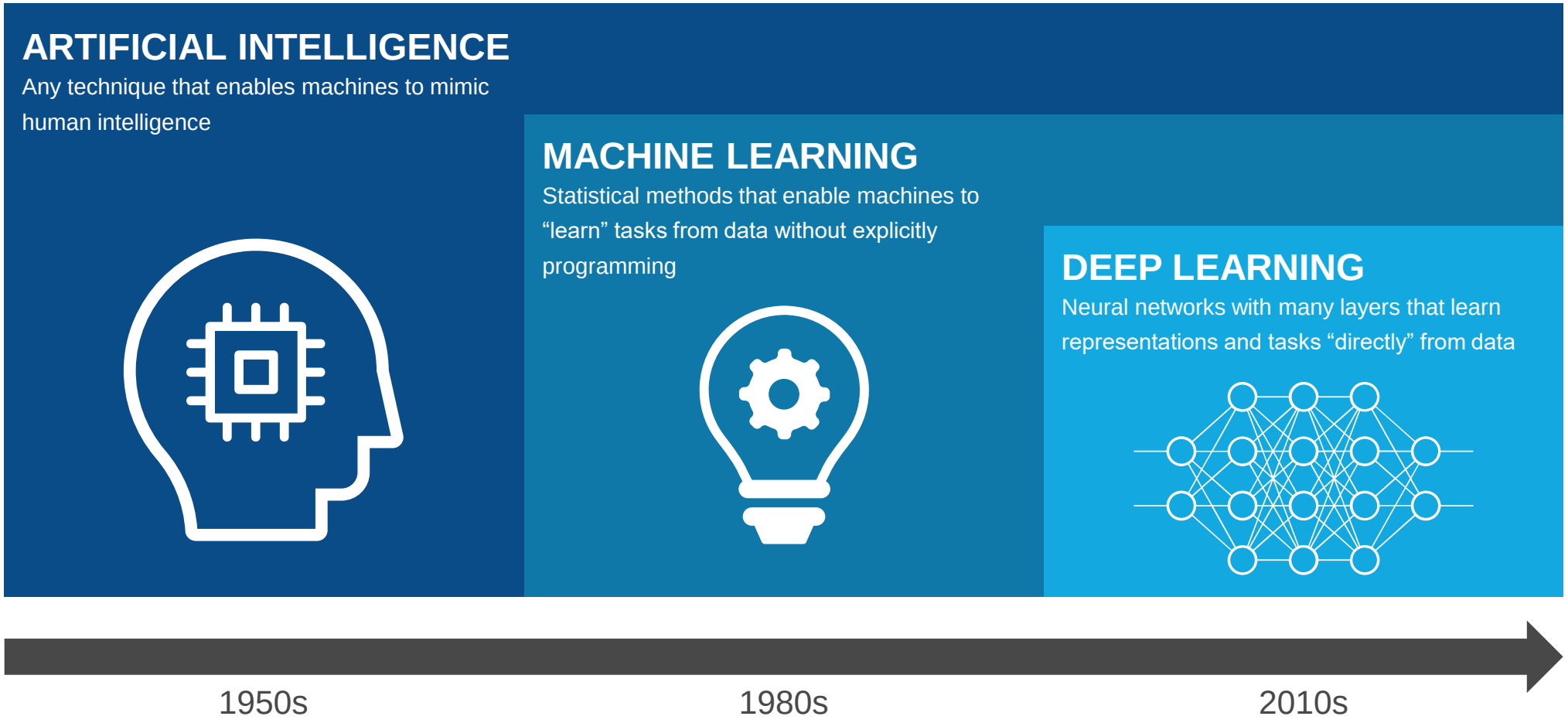
Generative Adversarial Networks (GAN)



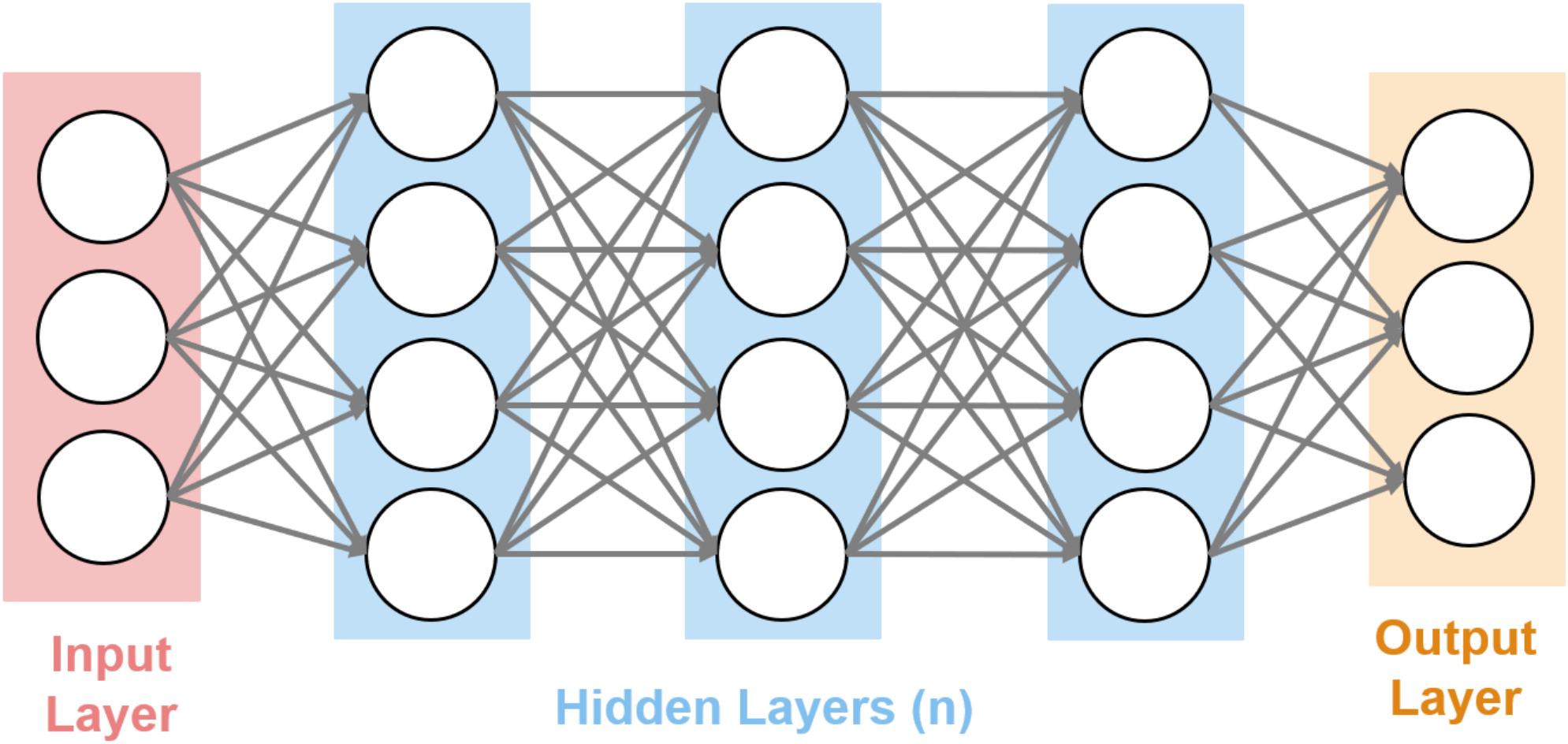
Deep Convolutional GAN



Deep learning is a key technology driving the AI megatrend

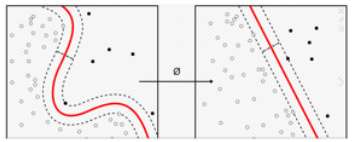


Deep Learning is a subset of machine learning that uses neural networks to extract features from data



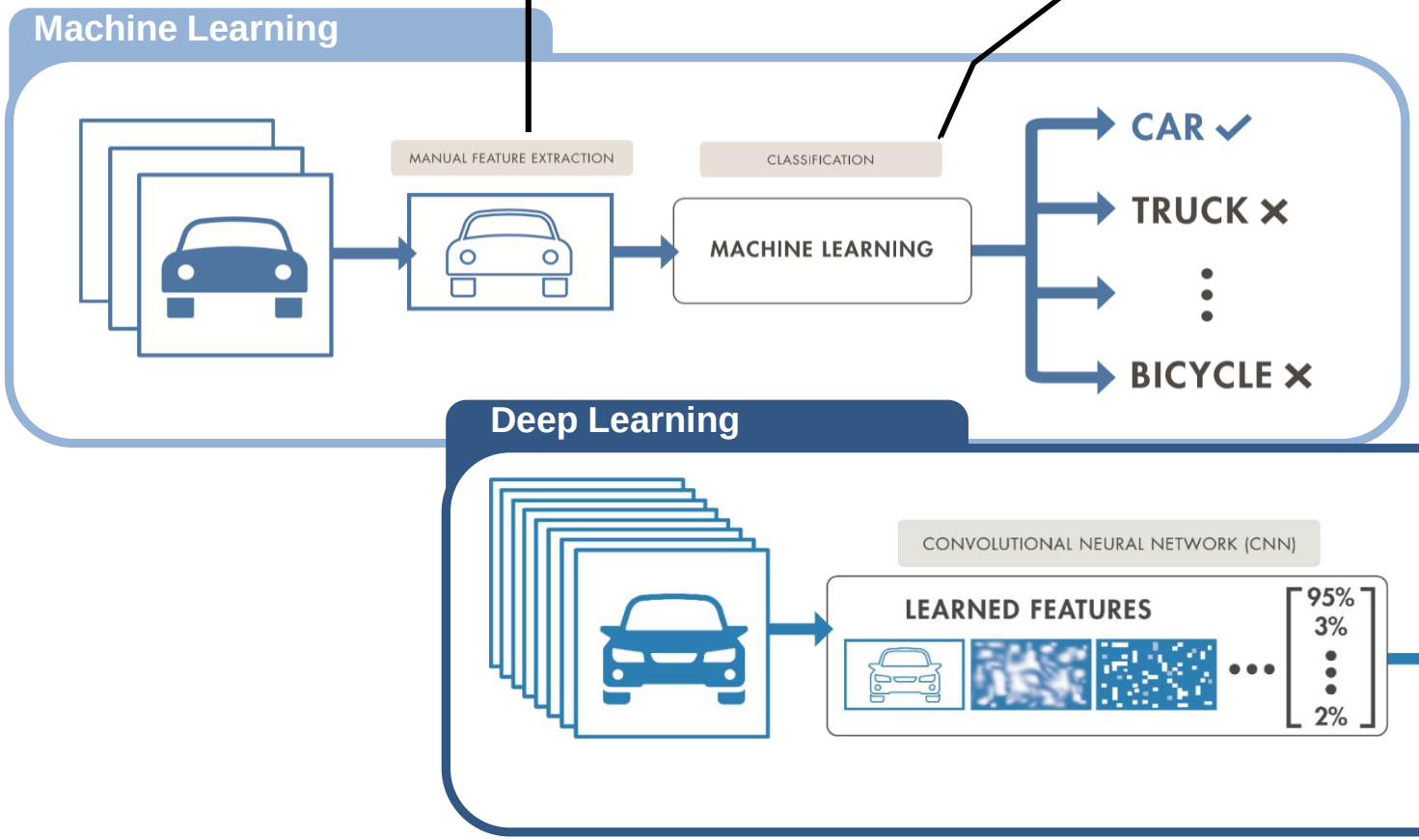
Machine Learning vs Deep Learning

Deep learning performs end-to-end learning by learning features, representations and tasks directly from **images, time-series, and text data**



We specify the nature of the features we want to extract…

…and the type of model we want to build.



Deep Learning Workflow

Data Preparation

- Data cleansing and preparation
- Human insight
- Simulation-generated data

AI Modeling

- Model design and tuning
- Hardware accelerated training
- Interoperability

Simulation & Test

- Integration with complex systems
- System simulation
- System verification and validation

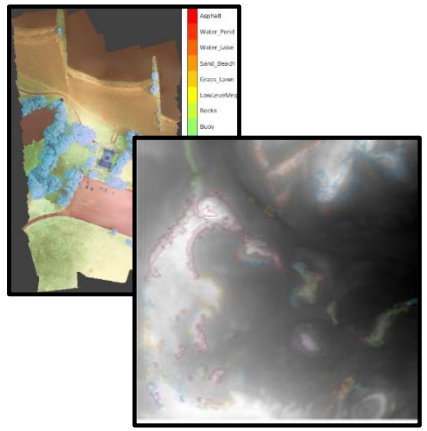
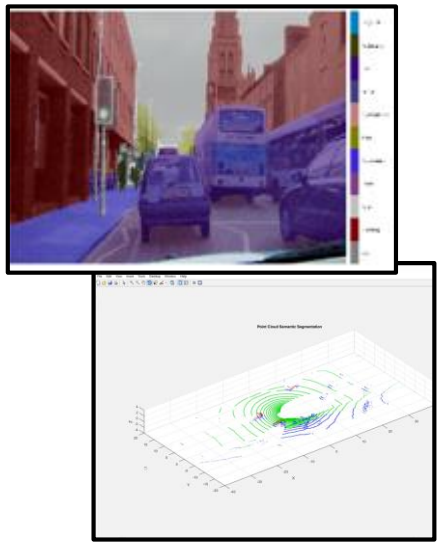
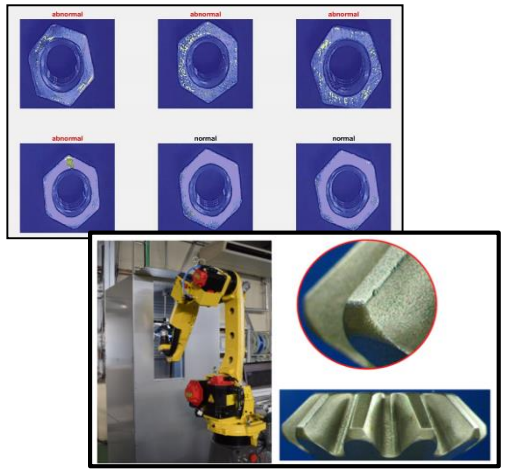
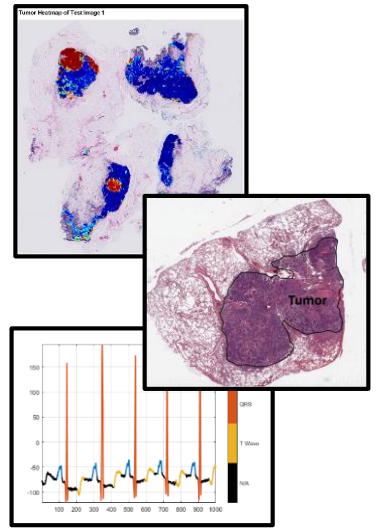
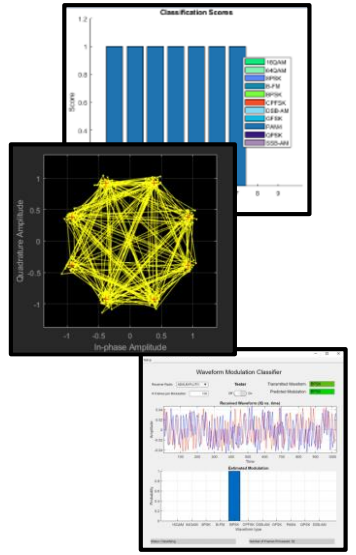
Deployment

- Embedded devices
- Enterprise systems
- Edge, cloud, desktop

Iteration and Refinement

Deep learning applications can be found across many industries

Industries

Applications	Aerospace & Defense	Automotive	Industrial Automation	Medical Devices	Communications
	Airborne Image Analysis 	**Autonomous Driving** 	**Defect Detection** 	**Medical Image and Signal Segmentation** 	**Modulation Classification** 

Airbus Uses Artificial Intelligence and Deep Learning for Automatic Defect Detection

Nicolas Castet, Airbus

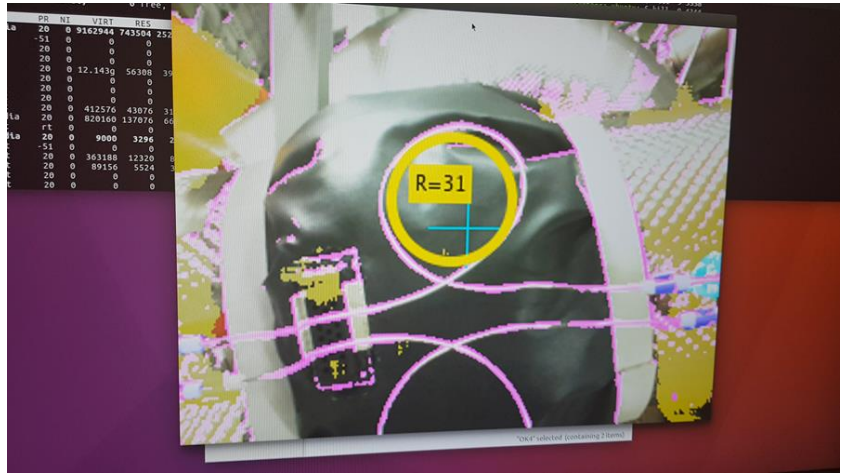
How do you build a robust end-to-end AI model to automatically detect defects in pipes in an aircraft? That was the big challenge for Airbus, which used MATLAB® to quickly prototype and develop deep learning models to meet their needs.

Working with the MathWorks Consulting Services team, Airbus adopted MATLAB to address the three main steps in the process. The first step was to have an integrated tool to build and train deep learning models from scratch for approaches such as semantic segmentation, as well as an easy and interactive environment for labeling videos. The positions of the ventilation holes and the wires on the pipe, found by the deep learning model in MATLAB, were used to measure distances and angles required by industry standard. Next, they needed to be able to display the analysis of the defects in real time. The final step was to translate the MATLAB code to CUDA code automatically, without requiring any coding skills, to deploy it directly on the embedded system.

Advantages of using MATLAB:

- Use an integrated tool to design, train, and deploy deep learning models
- Perform interactive prototyping and testing in a very short amount of time
- Directly translate MATLAB code to CUDA code

> “Having the ability to **test**, **modify**, **train**, and **test again** the code in a **short timeframe** was key to success.”



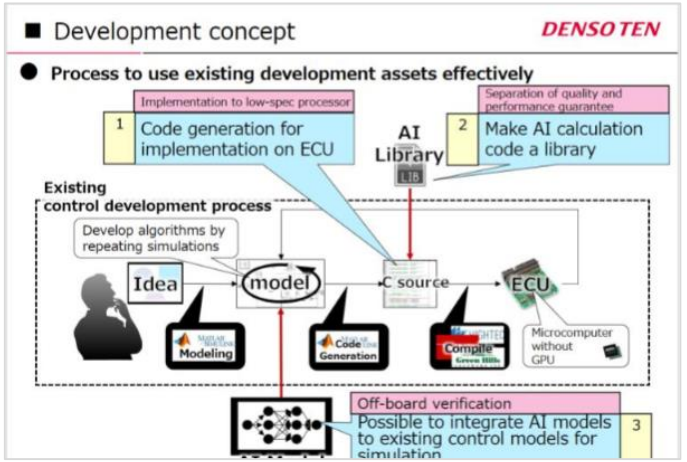
» **Learn about deep learning modeling**

Presented at MATLAB EXPO 2019 France

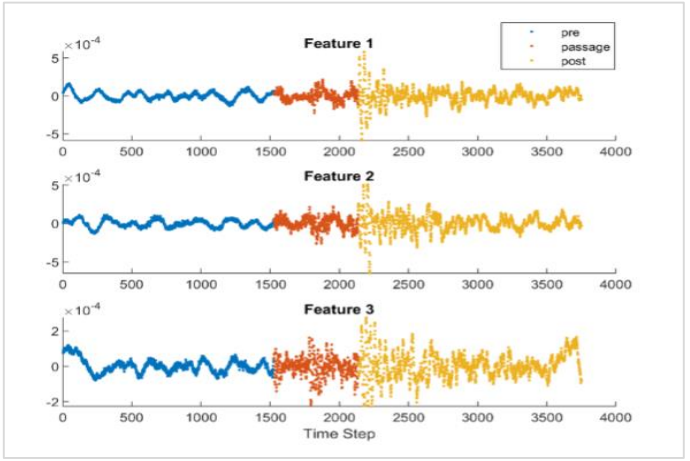
Watch video (18:25)

MATLAB Deep Learning used in Industry

Automatic Defect Detection
Airbus



ECU Vehicle Control
Denso



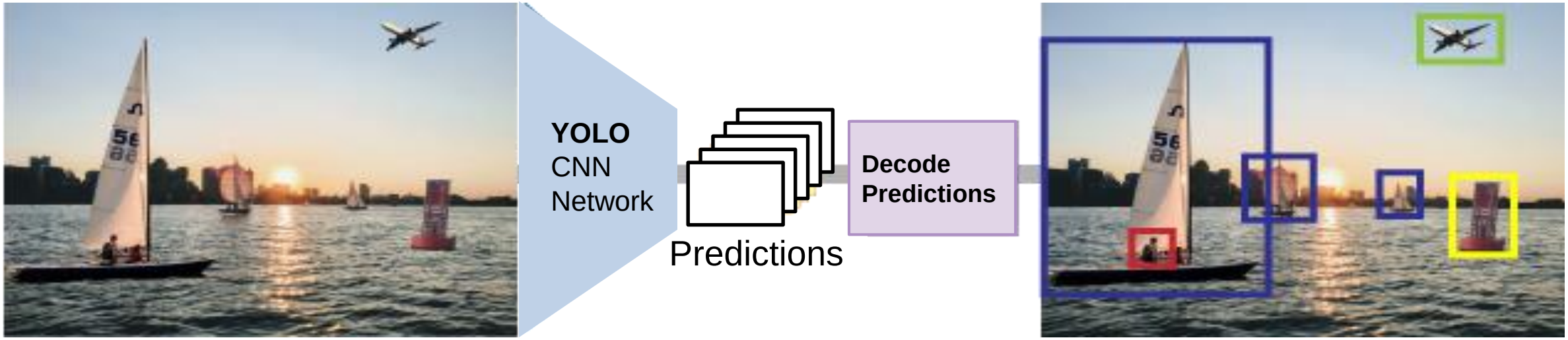
Seismic Event Detection
Shell

9

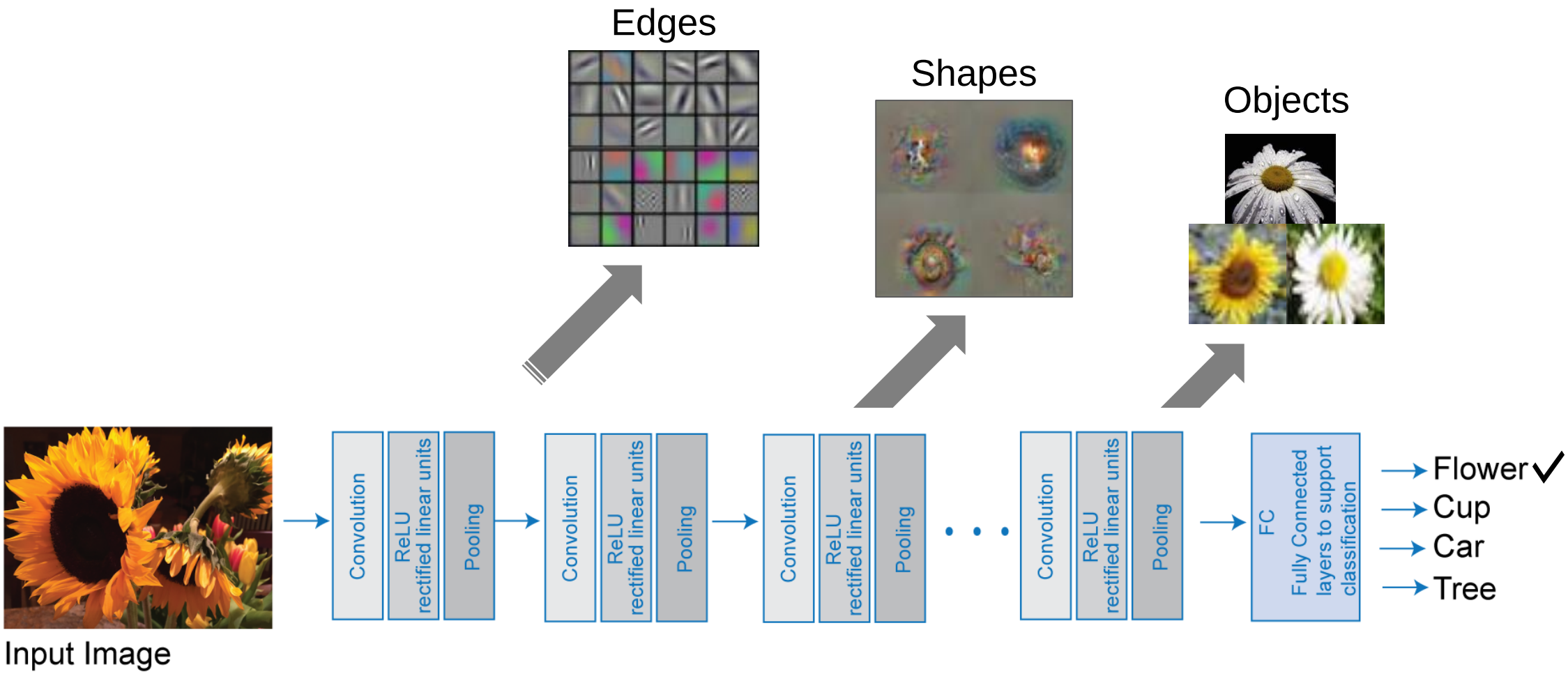
Featured Example: Detecting Objects with YOLO v2

Build, test, and deploy a deep learning solution that can detect objects in images and video.

- You Only Look Once
- Real-time object detector
- Autonomous driving, traffic monitoring
- 1000x faster than R-CNN



Convolutional Neural Networks (CNN)

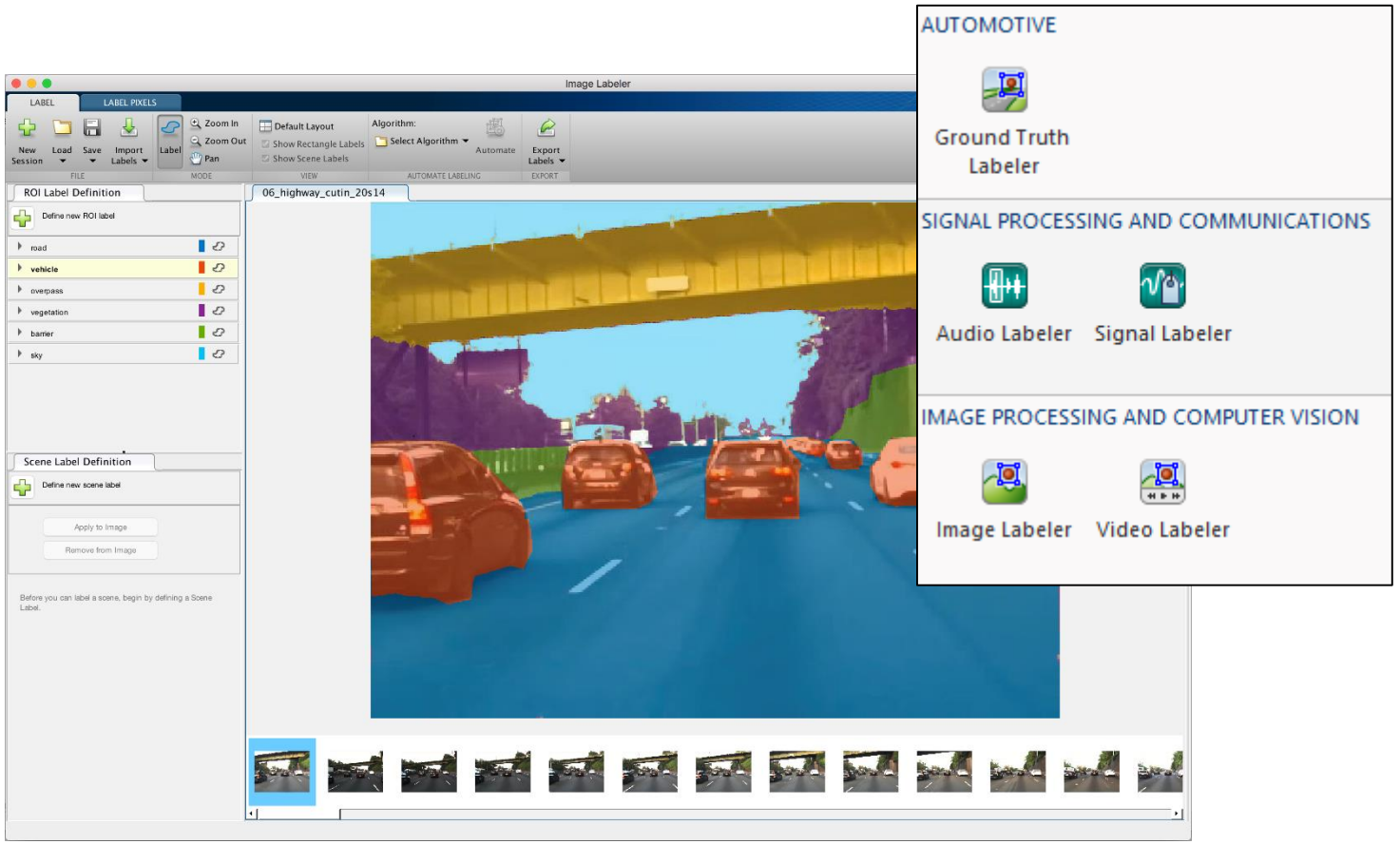


Spend less time preprocessing and labeling data

Synchronize disparate time series, filter noisy signals, automate labeling of video, and more.

Data Preparation

- Data cleansing and preparation
- Human insight
- Simulation-generated data

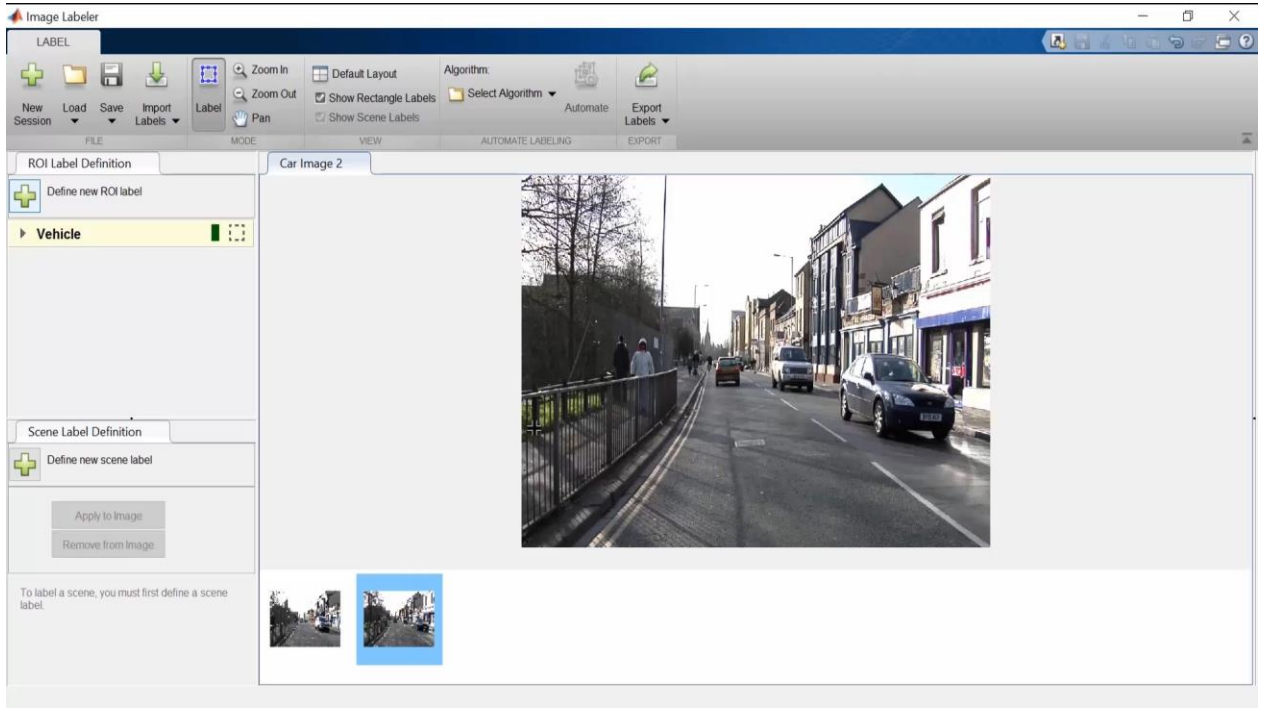


Use labeling apps for deep learning workflows like semantic segmentation

13

Using Apps for Ground Truth Labeling

Image and Video Data – Object Detection



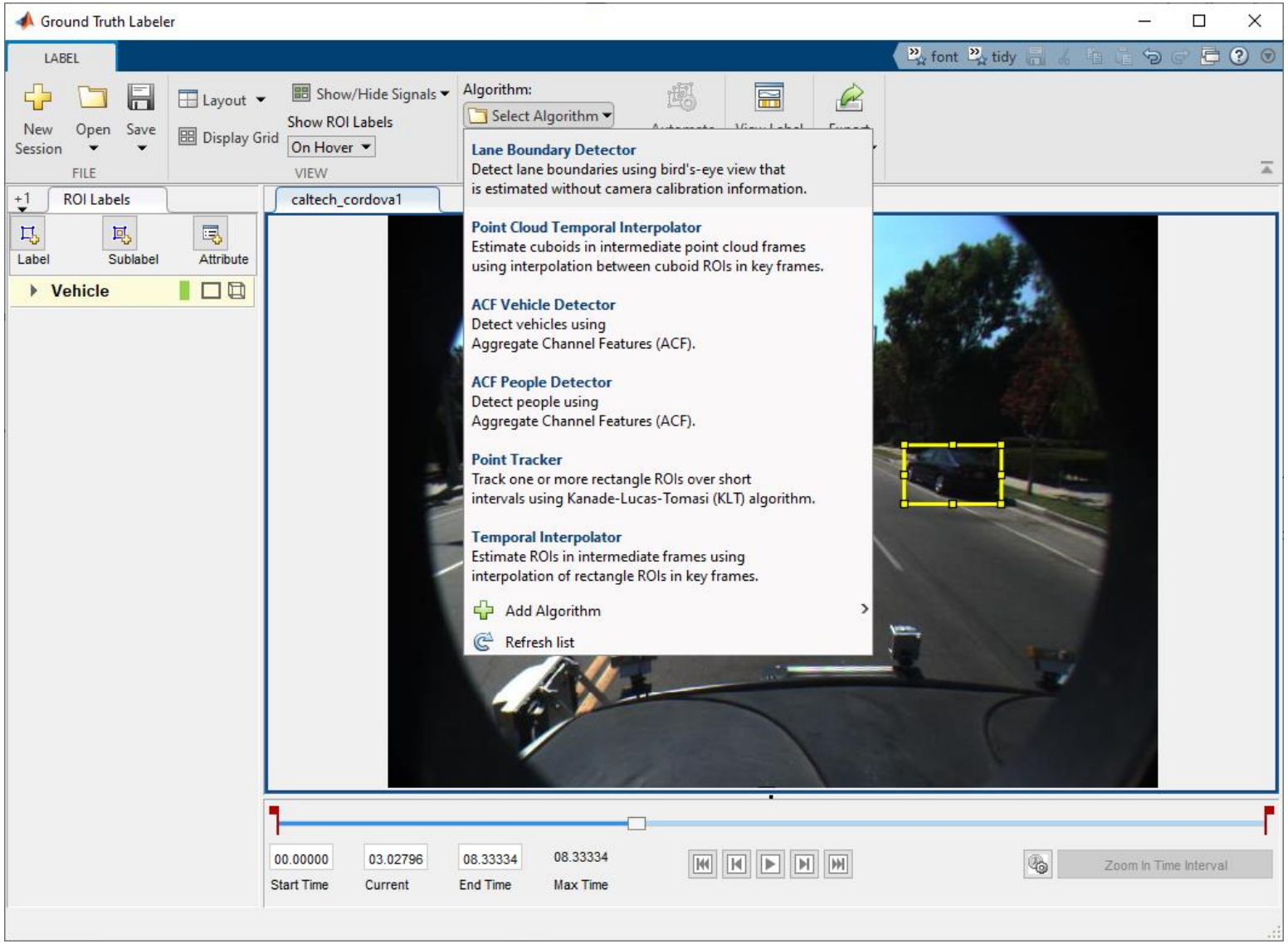
Built-in Vehicle Detector

Built-In Algorithms for Label Automation

- ACF vehicle detector
- ACF people detector
- Viola-Jones Algorithm for human face or torso detection
- HOG features and SVM for upright human detection

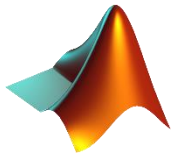
Using Apps for Ground Truth Labeling

Image and Video Data – Object Detection / Tracking



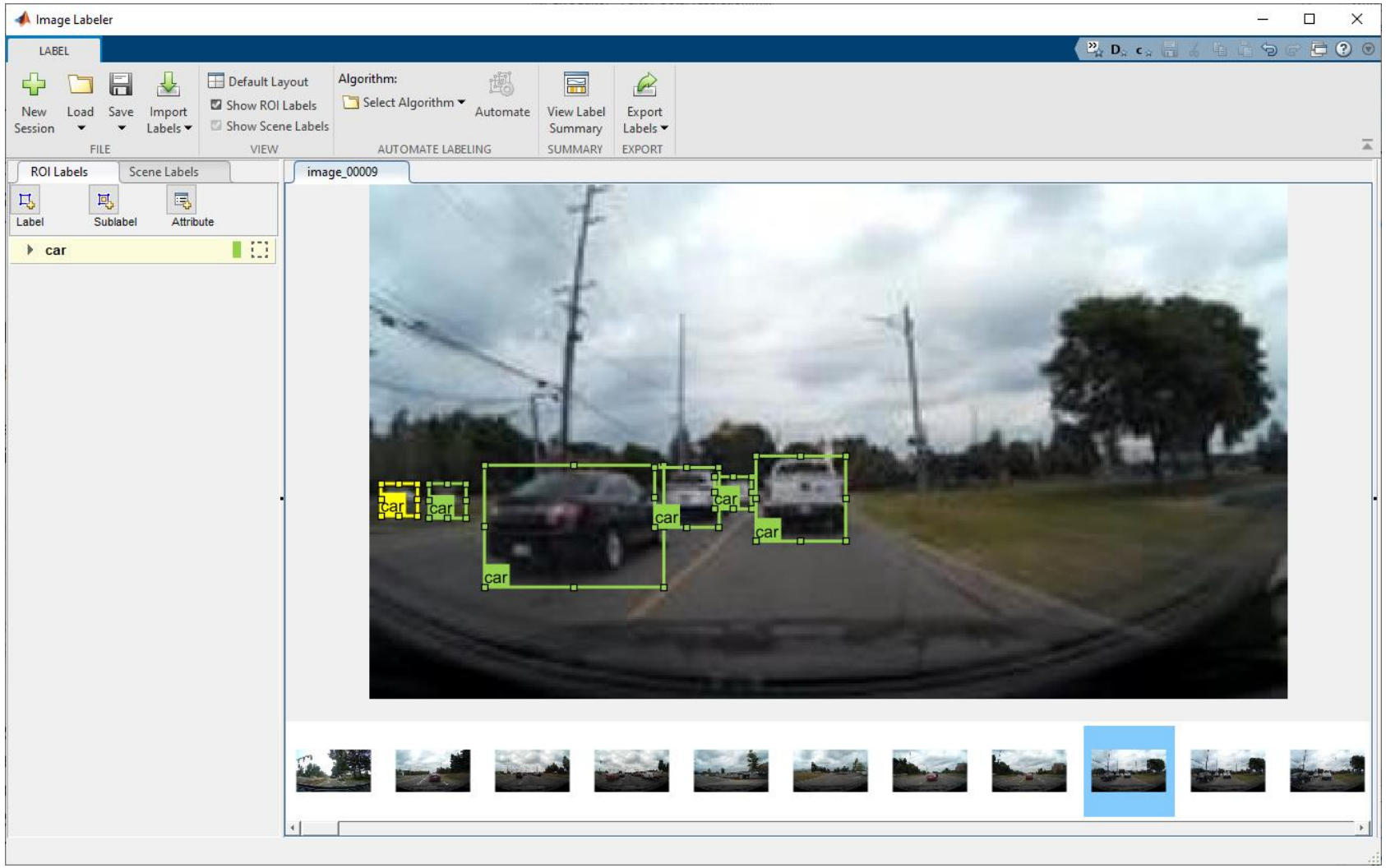
Data Preparation Demo

Poll



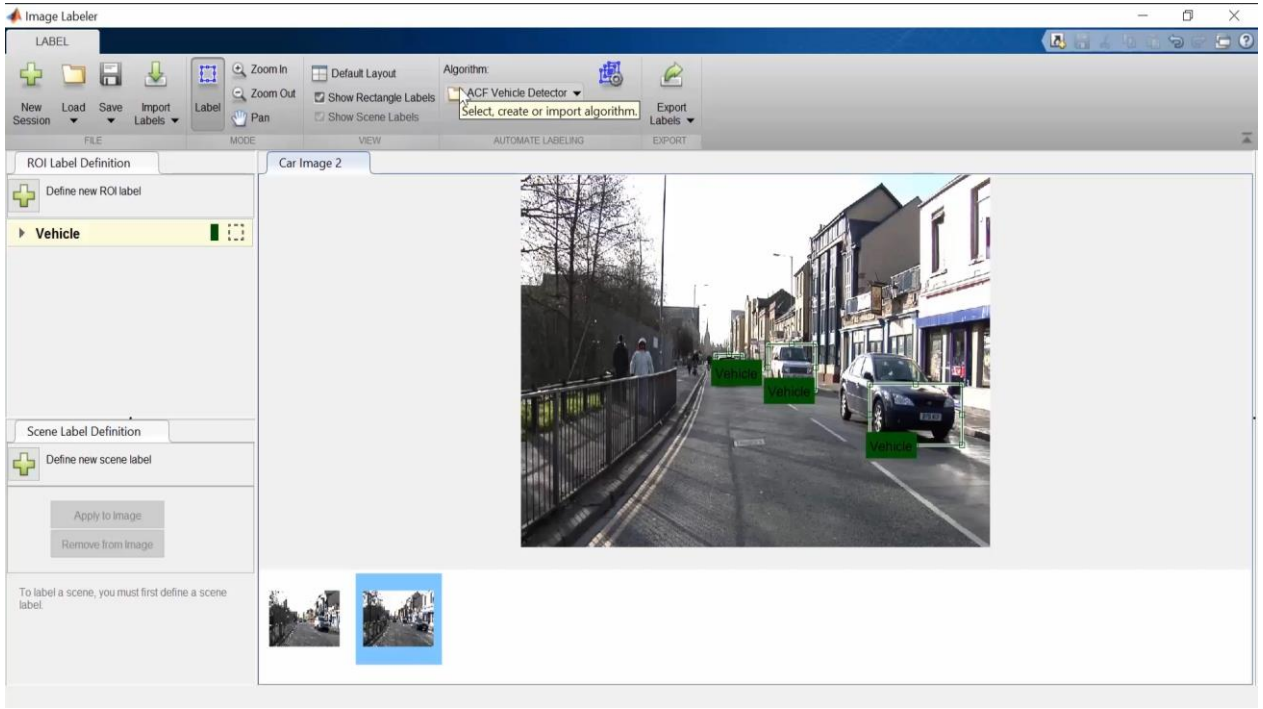
Data Preparation

- Data cleansing and preparation
- Human insight
- Simulation-generated data



Using Apps for Ground Truth Labeling

Image and Video Data – Semantic Segmentation



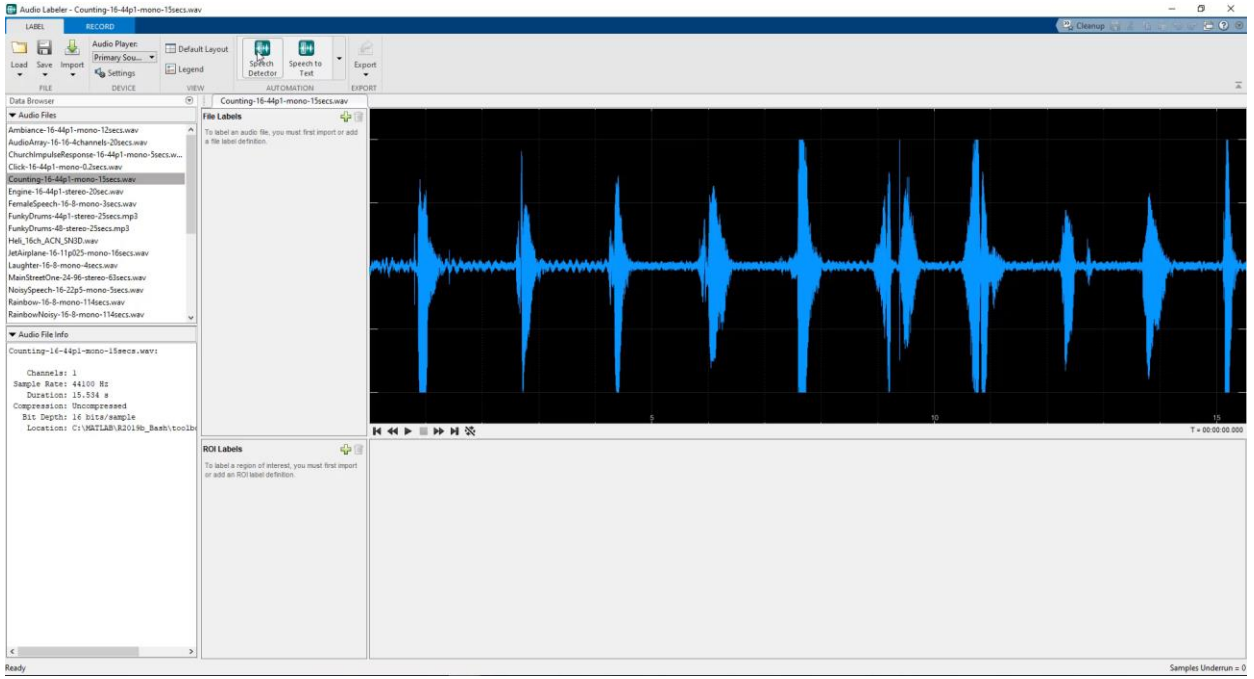
Pixel / ROI Labeling

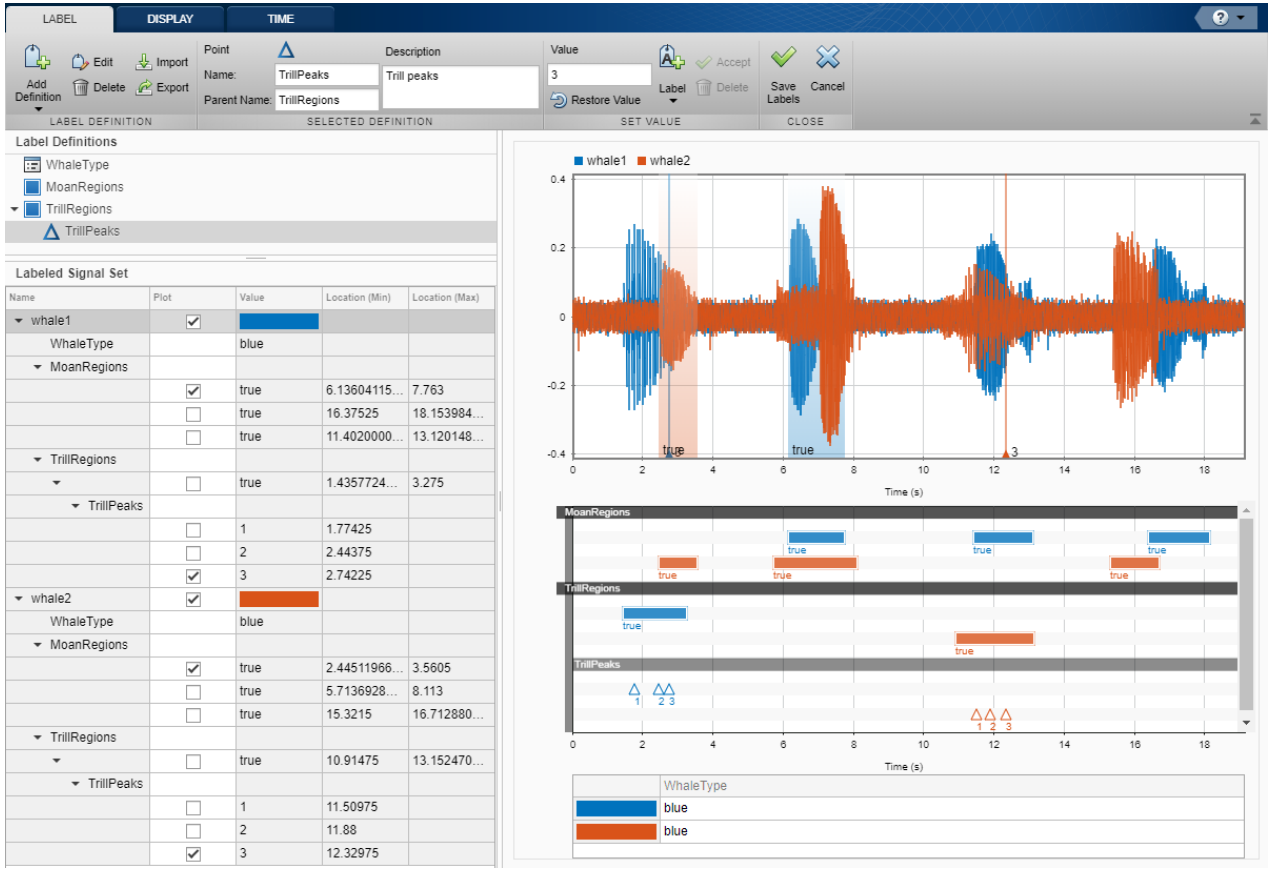
Custom automation algorithms for labeling with **`vision.labeler.AutomationAlgorithm`** class

- Create Automation Algorithm for Labeling
- Automate Ground Truth Labeling for Semantic Segmentation

Using Apps for Ground Truth Labeling

Audio and Signal Data





Handle big data with Datastores and Tall Arrays

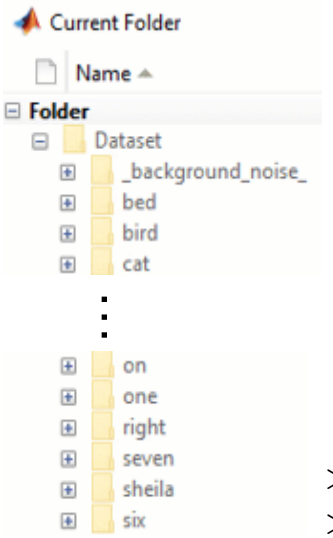
Handle out-of-memory data and large data collections easily without reworking your code

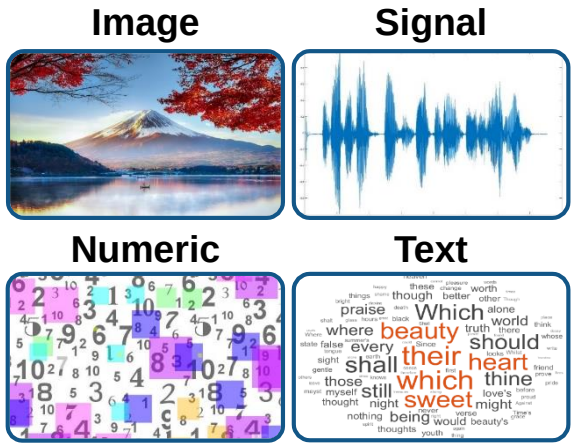
Data Preparation

- Data cleansing and preparation
- Human insight
- Simulation-generated data

Datastores

- Handle out-of-memory data
- Flexible across many data formats



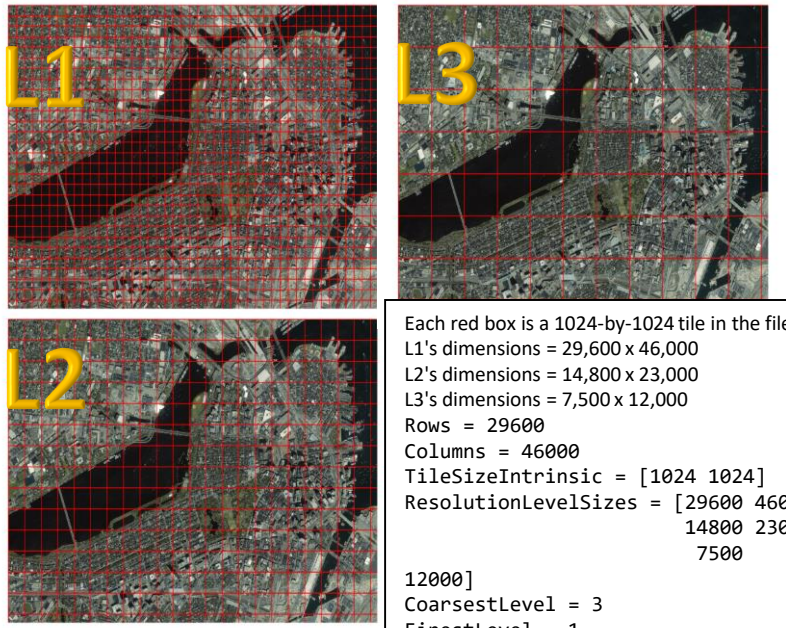



```
>> imageDatastore
>> audioDatastore
>> fileDatastore
   ...
Custom Datastores also available
```

Documentation: Datastores for Various File Formats

bigimage

- Work with very large, tiled and multi-resolution images




```
Each red box is a 1024-by-1024 tile in the file.
L1's dimensions = 29,600 x 46,000
L2's dimensions = 14,800 x 23,000
L3's dimensions = 7,500 x 12,000
Rows = 29600
Columns = 46000
TileSizeIntrinsic = [1024 1024]
ResolutionLevelSizes = [29600 46000
                        14800 23000
                         7500
12000]
CoarsestLevel = 3
FinestLevel = 1
PixelSpacings = [1 1; 2 2; 3.947
3.833]
```

Documentation: bigimageDatastore

20

Augment existing data or synthesize additional data

Augment existing data or synthesize more to address data imbalance or lack of data.

Data Preparation

- Data cleansing and preparation
- Human insight
- Simulation-generated data

Data Augmentation

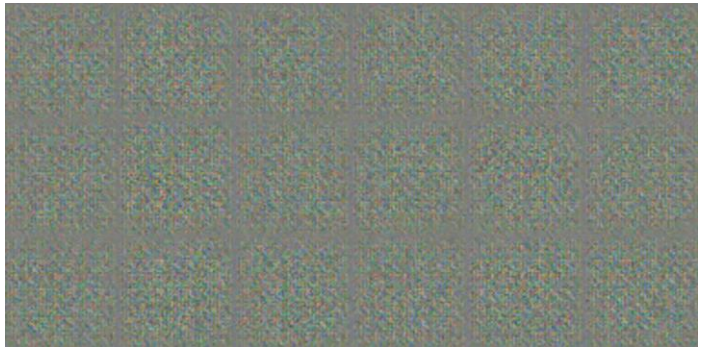
Built in functionalities:

- Transform (e.g. rotate, reflect, warp, jitter)
- Replicate
- Class weighting

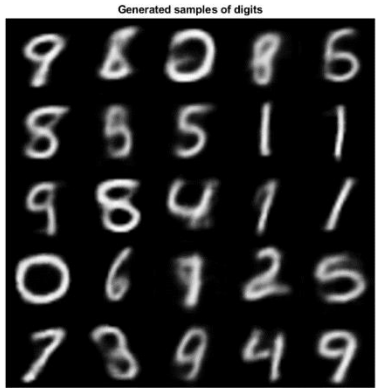
Article: Data Augmentation for Image Classification

Data Synthesis

Generative Adversarial Networks



Variational Auto Encoders



21

Start with a complete set of algorithms and pre-built models

AI Modeling


Model design and tuning


Hardware accelerated training


Interoperability

Algorithms

Machine learning
Trees, Naïve Bayes, SVM…

Deep learning
CNNs, GANs, LSTM, MIMO…

Reinforcement learning
DQN, A2C, DDPG…

Regression
Linear, nonlinear, trees…

Unsupervised learning
K-means, PCA, GMM…

Predictive maintenance
RUL models, condition indicators…

Bayesian optimization

Pre-built models

Image classification models
AlexNet, GoogLeNet, VGG, SqueezeNet, ShuffleNet, ResNet, DenseNet, Inception…

Reference examples

Object detection
Vehicles, pedestrians, faces…

Semantic segmentation
Roadway detection, land cover classification, tumor detection…

Signal and speech processing
Denoising, music genre recognition, keyword spotting, radar waveform classification…

…and more…

22

Transfer learning workflow

Load pretrained network

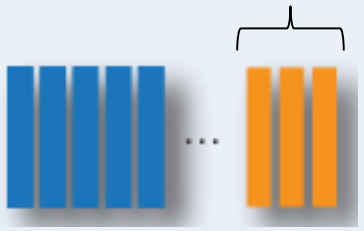
Early layers that learned low-level features (edges, blobs, colors)

Last layers that learned task specific features

1 million images
1000s classes

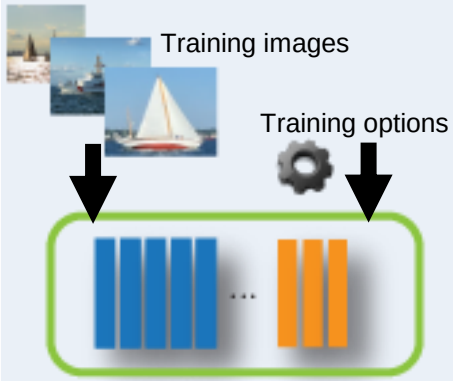
Replace final layers

New layers to learn features specific to your data



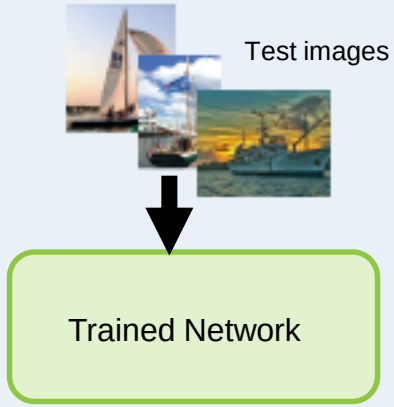
Fewer classes
Learn faster

Train network



100s images
10s classes

Predict and assess network accuracy

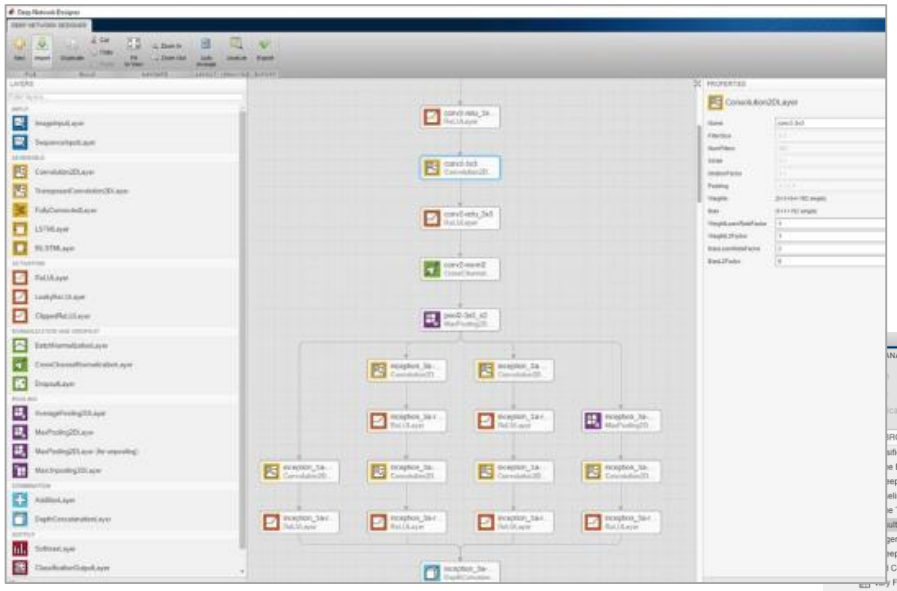


Increase productivity using Apps for design and analysis

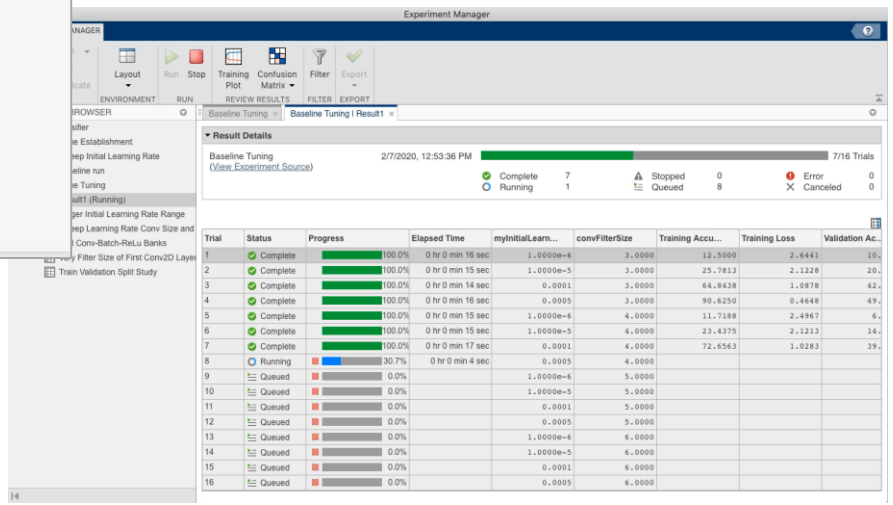
Use MATLAB Apps to design deep learning networks, explore a wide range of classifiers, train regression models, train an optical character recognition model, and more.

AI Modeling

- Model design and tuning
- Hardware accelerated training
- Interoperability



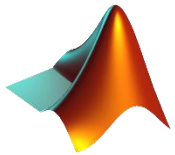
Deep Network Designer app to build, visualize, and edit deep learning networks



Experiment Manager app to manage multiple deep learning experiments, analyze and compare results and code

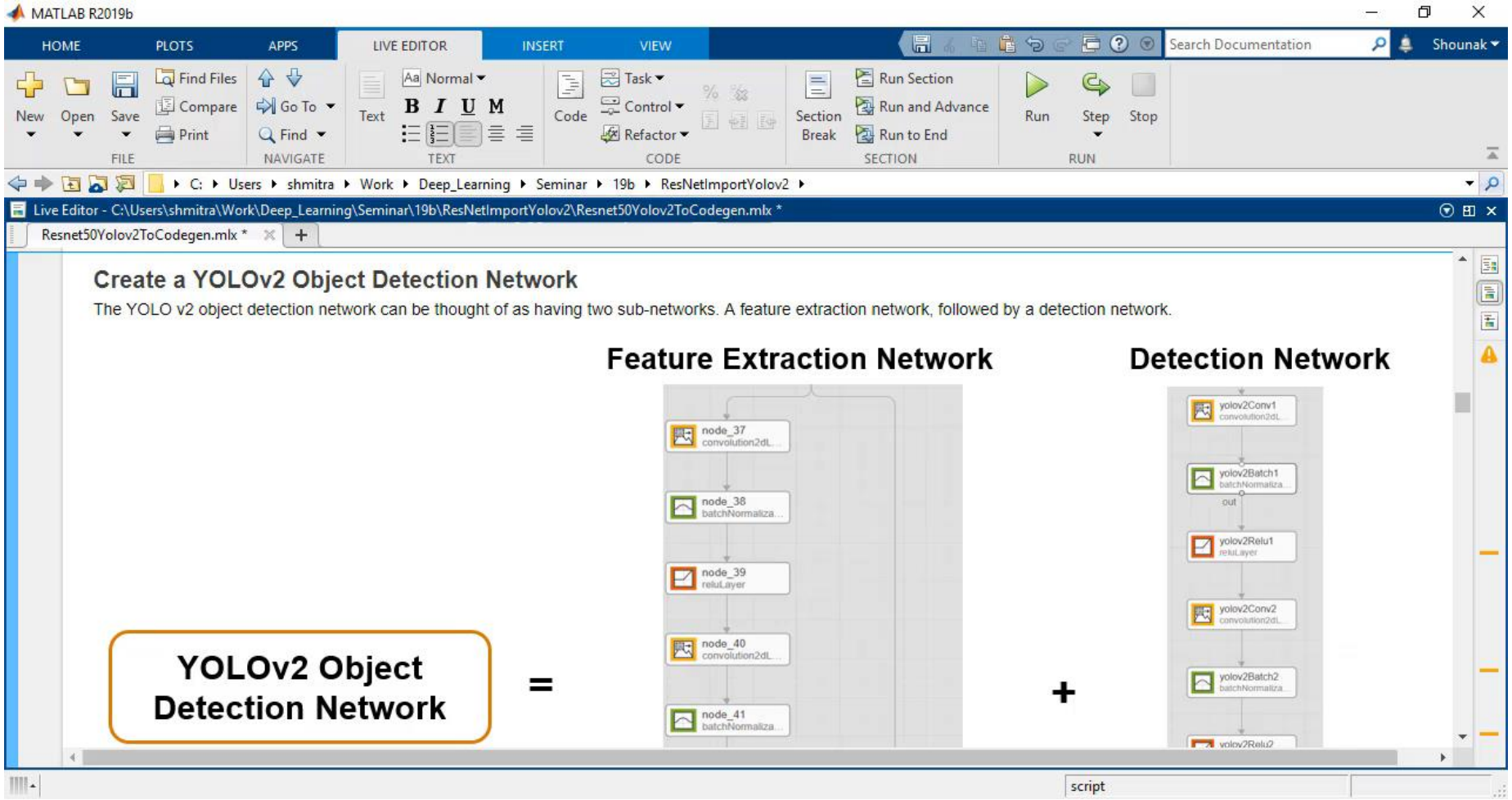
25

Modeling Demo

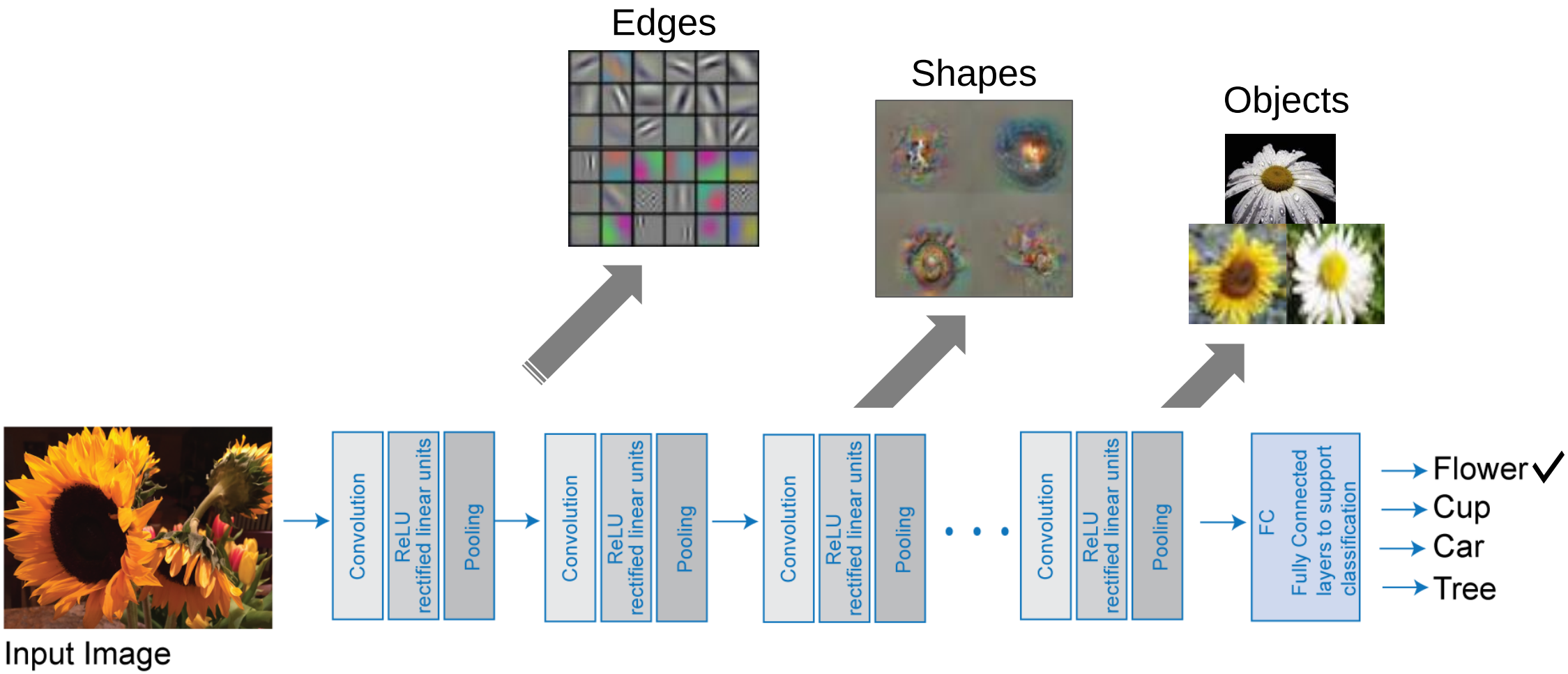


AI Modeling

- Model design and tuning
- Hardware accelerated training
- Interoperability

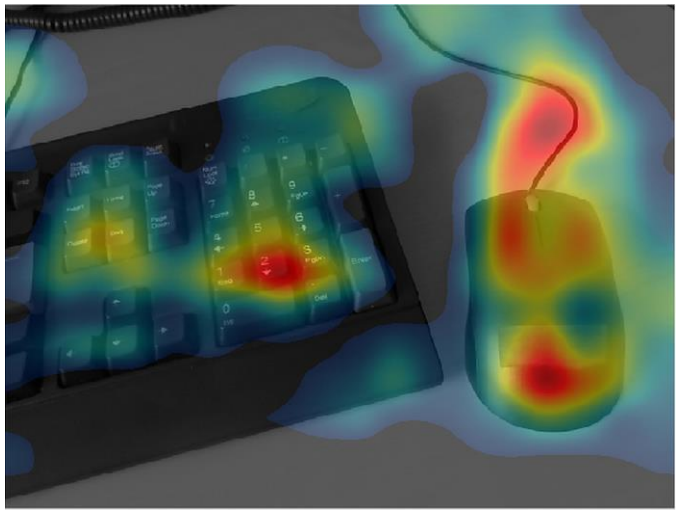


Convolutional Neural Networks (CNN)

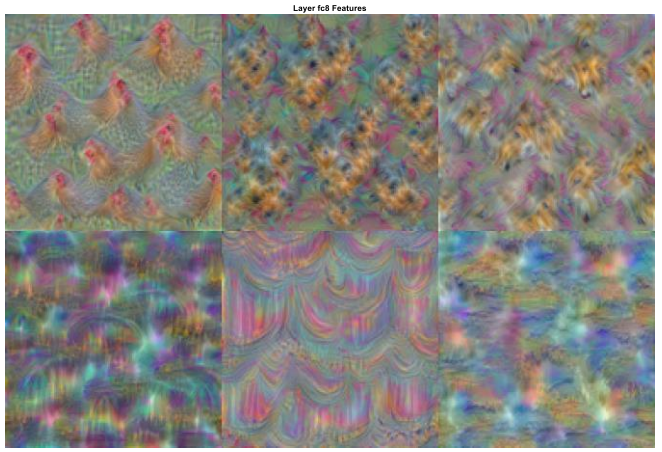


CNN Visualization Methods

Layer Activations

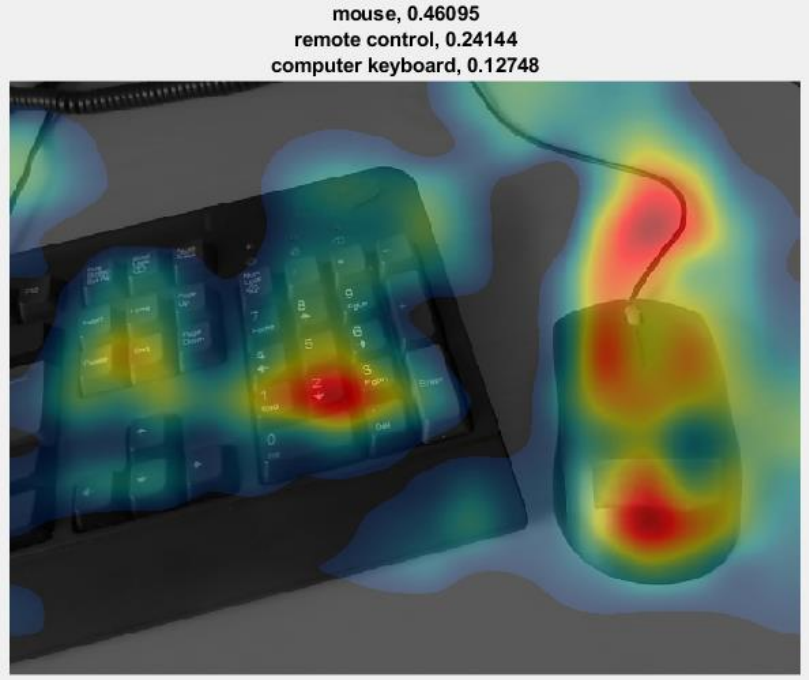


Class Activations

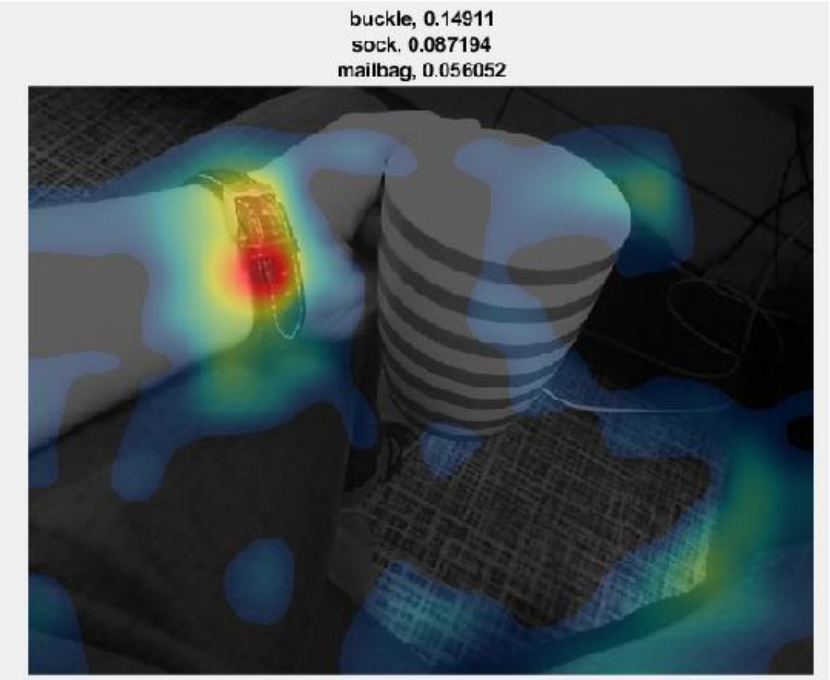


DeepDream Images

Class Activation Mapping to Investigate Network Predictions



Classified as “mouse”- incorporated presence of keyboard



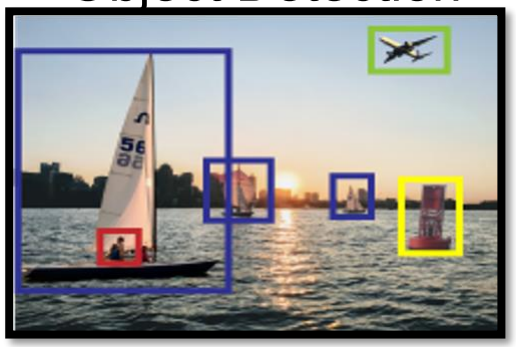
Incorrectly classified “coffee mug” as “buckle” due to the watch

Deep Learning can be applied to object recognition, detection, image regression, and segmentation problems

Object recognition

Image Classification

Object Detection



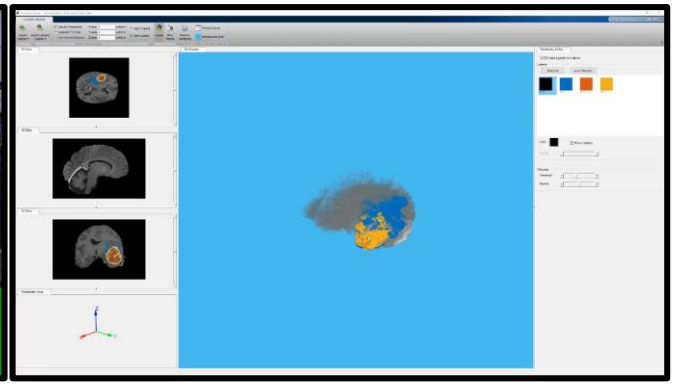
YOLO v2 Object Detector

Image to Image Regression

Image Super-Resolution using VDSR Network

Semantic Segmentation

Free Road Detection



Brain Tumor Segmentation

Object detection systems are ideal for applications like autonomous driving and traffic monitoring

YOLO v2 (You Only Look Once)

Yolo-v2

Yolo-v3

Single Shot Detection

RCNN

Faster-RCNN

Fast-RCNN

Prominent Algorithms for Object detection workflows

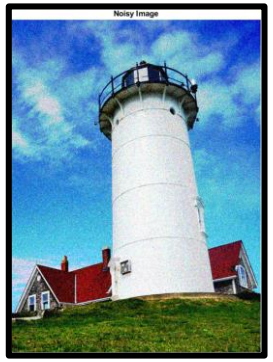
How it works

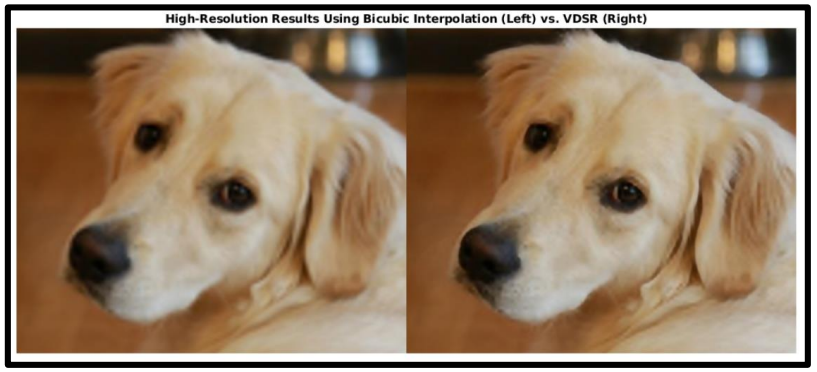
YOLOv2 Object Detection Network = **Feature Extraction Network** + **Object Detection Network**

Typically, ResNets, VGGs, or Inception models

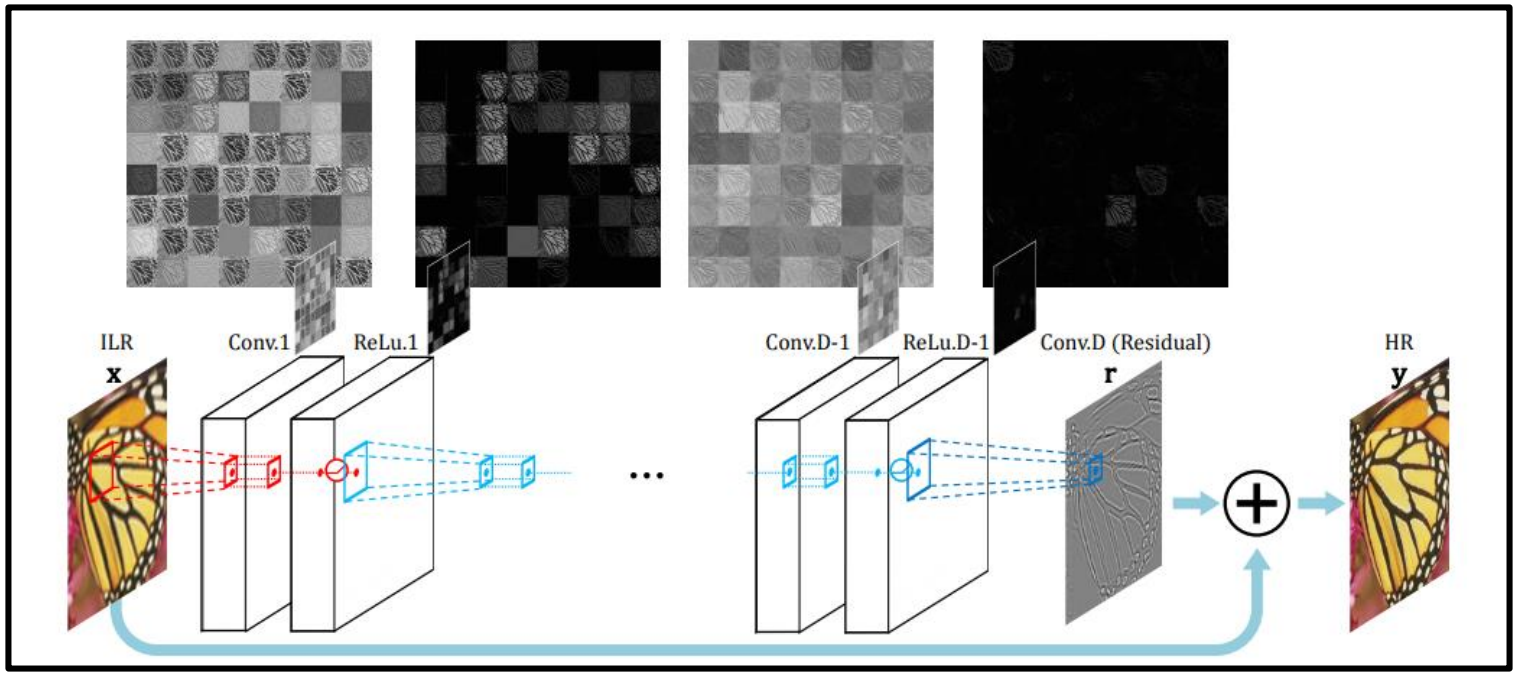
Contains YOLOv2 specific layer with anchor boxes

Image-Image Regression is effective in applications like estimating a high-resolution image from a single low-resolution image





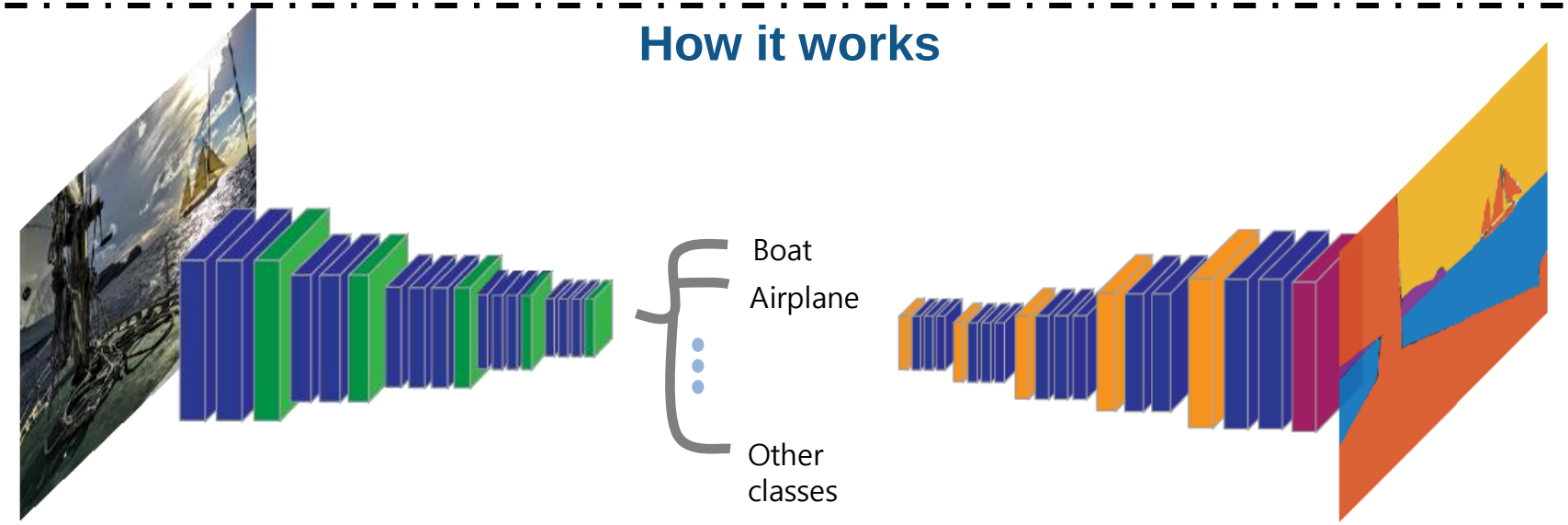
How it works



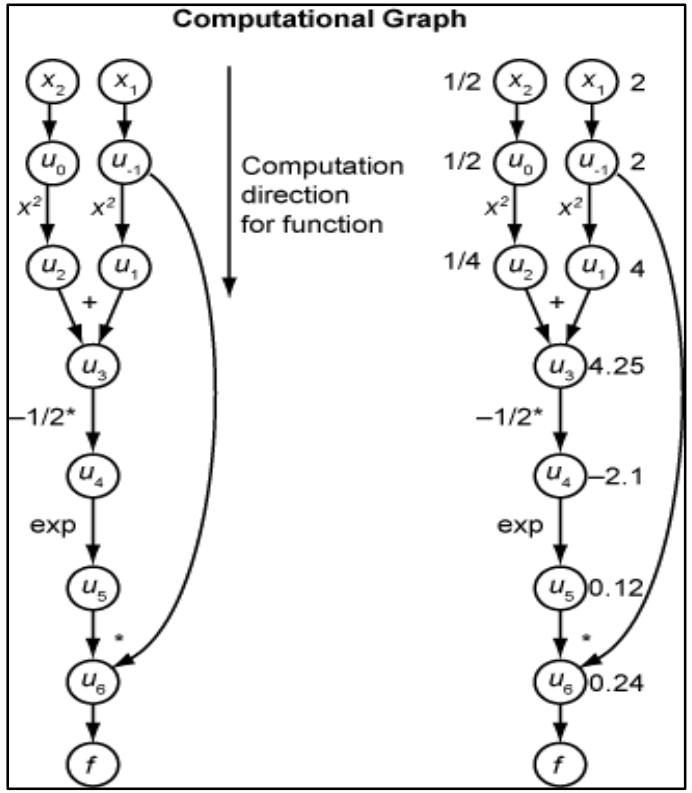
Semantic Segmentation plays a central role in applications such as Automated Driving

Semantic Segmentation using SegNet

How it works



MATLAB provides flexibility to design and train advanced networks



Automatic Differentiation

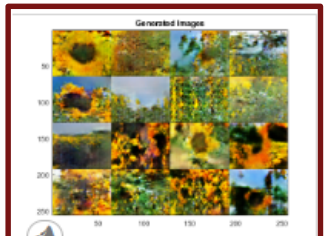
[Documentation]

```
% Loop over epochs.
for epoch = 1:numEpochs
    % Shuffle data.
    idx = randperm(numel(YTrain));
    XTrain = XTrain(:,:,:,idx);
    YTrain = YTrain(idx);

    % Loop over mini-batches.
    for i = 1:numIterationsPerEpoch
        iteration = iteration + 1;

        % Read mini-batch of data and convert the labels to dummy
        % variables.
        idx = (i-1)*miniBatchSize+1:i*miniBatchSize;
        X = XTrain(:,:,:,idx);
```

Custom Training Loops

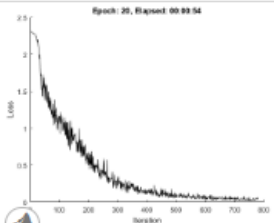
[Documentation Example]

Design and build advanced network architectures like GANs, Siamese Networks, Variational Autoencoders



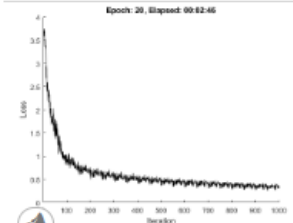
Train Generative Adversarial Network (GAN)

Train a generative adversarial network (GAN) to generate images.



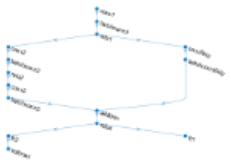
Train Network Using Custom Training Loop

Train a network that classifies handwritten digits with a custom learning rate schedule.



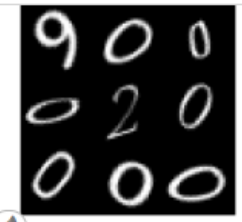
Train Network Using Model Function

Create and train a deep learning network by using functions rather than a layer graph or a dlnetwork. The advantage of using functions is



Train Network with Multiple Outputs

Train a deep learning network with multiple outputs that predict both labels and angles of rotations of handwritten digits.

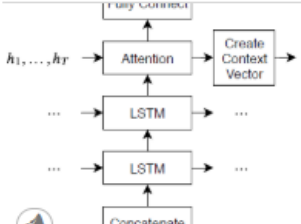


Train a Siamese Network for Dimensionality Reduction

Train a Siamese network to compare handwritten digits using dimensionality reduction.

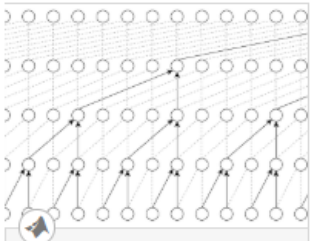
Train a Siamese Network to Compare Images

Train a Siamese network to identify similar images of handwritten characters.



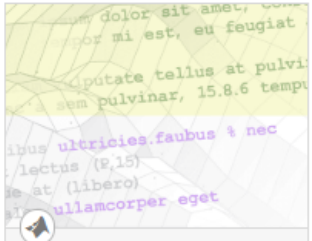
Sequence-to-Sequence Translation Using Attention

Convert decimal strings to Roman numerals using a recurrent sequence-to-sequence encoder-decoder model with attention.



Sequence-to-Sequence Classification Using 1-D Convolutions

Classify each time step of sequence data using a generic temporal convolutional network (TCN).

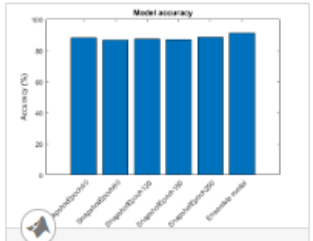


Train Network in Parallel with Custom Training Loop

Set up a custom training loop to train a network in parallel. In this example, parallel workers train on portions of the overall mini-batch. If

Train Variational Autoencoder (VAE) to Generate Images

Create a variational autoencoder (VAE) in MATLAB to generate digit images. The VAE generates hand-drawn digits in the style of the



Train Network Using Cyclical Learn Rate for Snapshot Ensembling

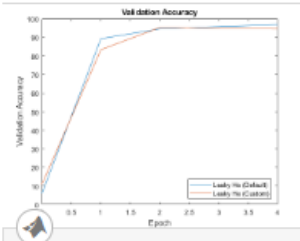
Train a network to classify images of objects using a cyclical learn rate schedule and snapshot ensembling for better test accuracy. In the

Make Predictions Using dlnetwork Object

Make predictions using a dlnetwork object by splitting data into mini-batches.

Make Predictions Using Model Function

Make predictions using a model function by splitting data into mini-batches.



Specify Custom Weight Initialization Function

Create a custom He weight initialization function for convolution layers followed by leaky ReLU layers.

Hardware acceleration and scaling are critical for training

MATLAB accelerates AI training on GPUs, cloud, and datacenter resources without specialized programming.

AI Modeling

- Model design and tuning
- Hardware accelerated training
- Interoperability

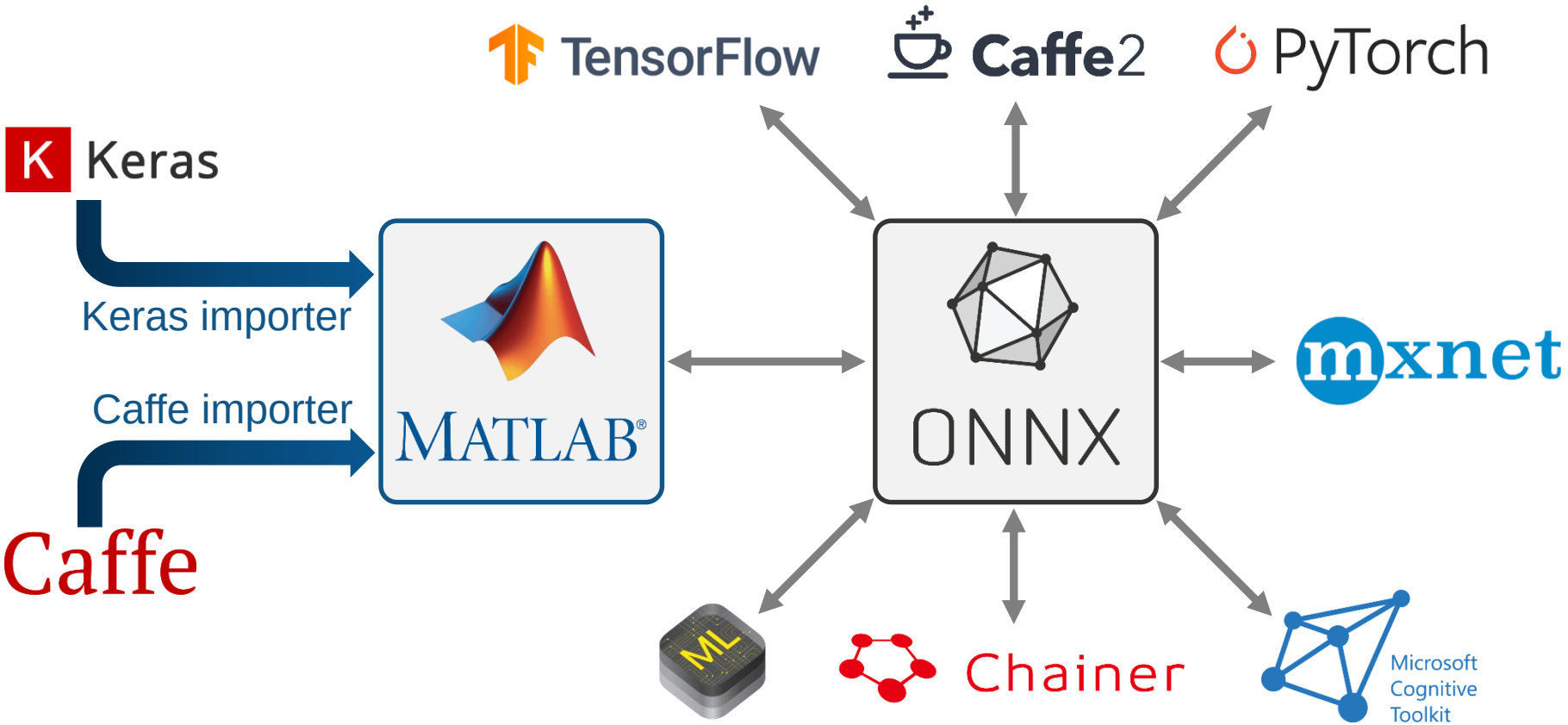
36

MATLAB interoperates with other frameworks

Supports ONNX and can exchange models with PyTorch, TensorFlow, and other frameworks.

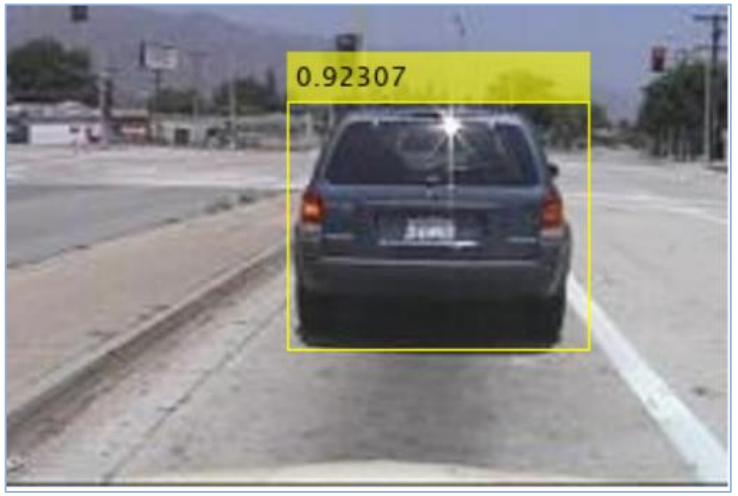
AI Modeling

- Model design and tuning
- Hardware accelerated training
- Interoperability



37

Modeling Demo



AI Modeling

- Model design and tuning
- Hardware accelerated training
- Interoperability

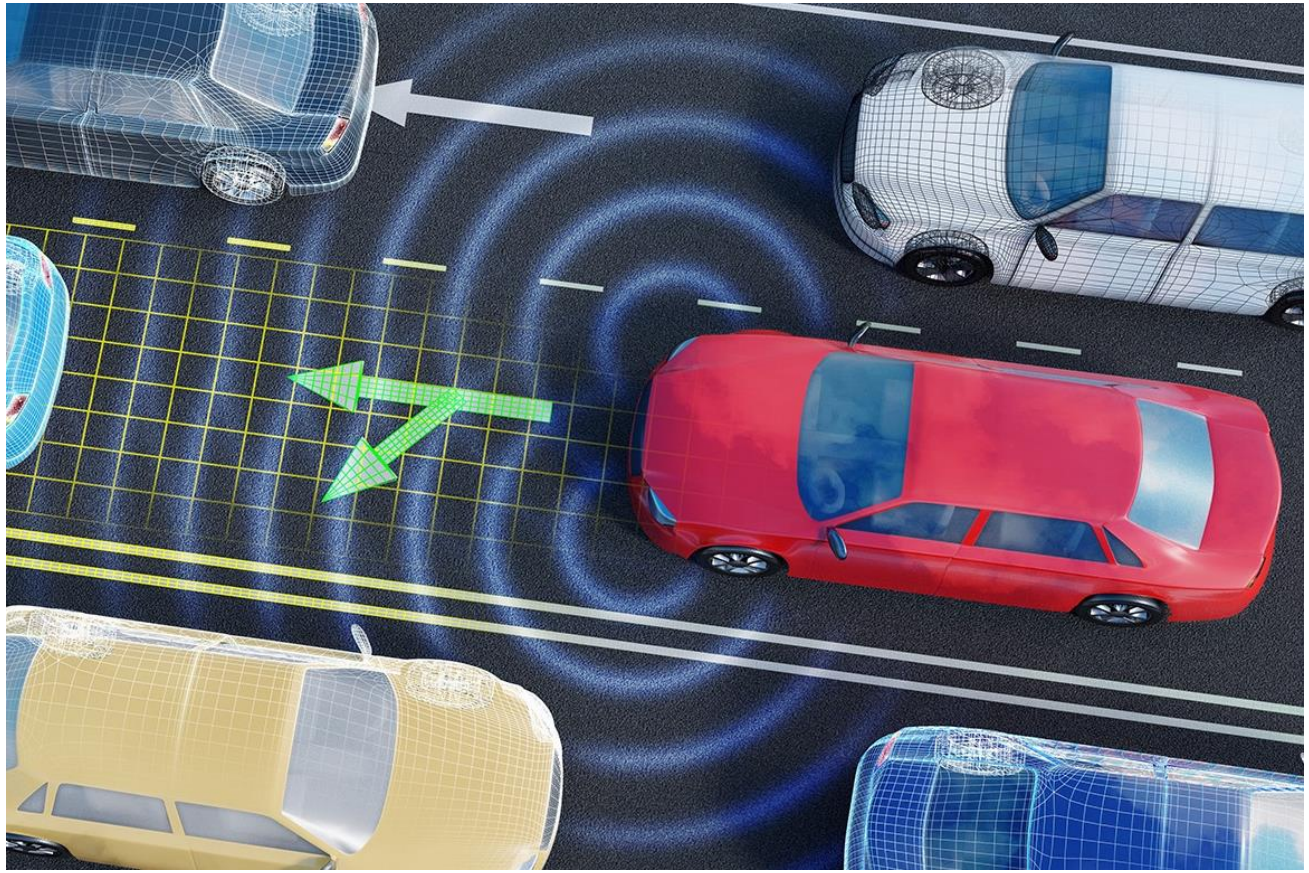
Models need to exist within a complete system

In automated driving systems, AI for perception must integrate with algorithms for path planning, braking, acceleration, and other controls.

- ISO 26262 support in MATLAB & Simulink

Simulation & Test

- Integration with complex systems
- System simulation
- System verification and validation



39

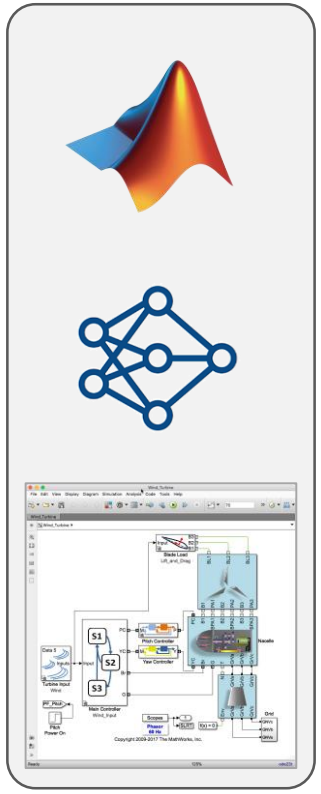
Deploy to any processor with best-in-class performance

AI models in MATLAB and Simulink can be deployed on embedded devices, edge devices, enterprise systems, the cloud, or the desktop.

Poll

Deployment

- Embedded devices
- Enterprise systems
- Edge, cloud, desktop



Code Generation

CPU

GPU

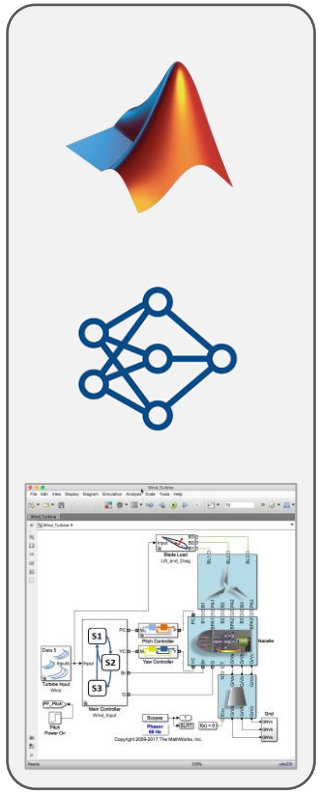
FPGA

40

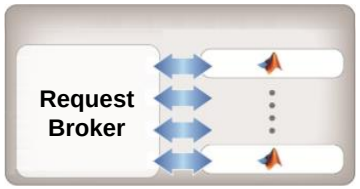
Deploy to enterprise IT infrastructure

Deployment

- Embedded devices
- Enterprise systems
- Edge, cloud, desktop



MATLAB Production Server

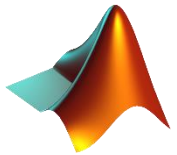


- Databases
- Streaming
- Cloud Storage
- Dashboards
- Containers
- Custom Tools

Cloud & Datacenter Infrastructure

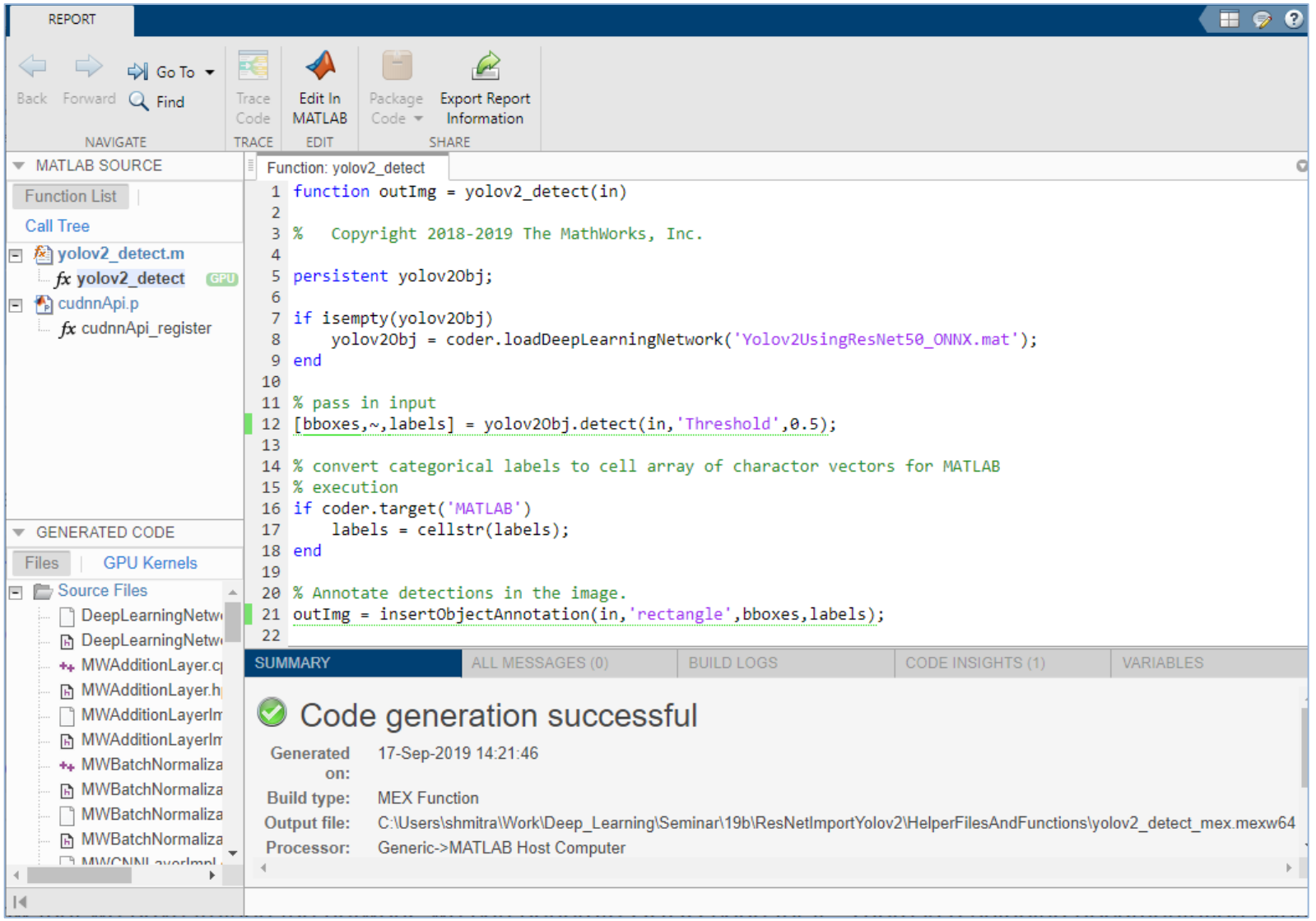
41

Deployment Demo



Deployment

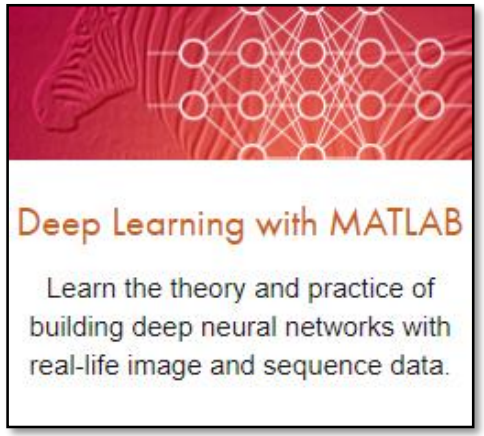
- Embedded devices
- Enterprise systems
- Edge, cloud, desktop

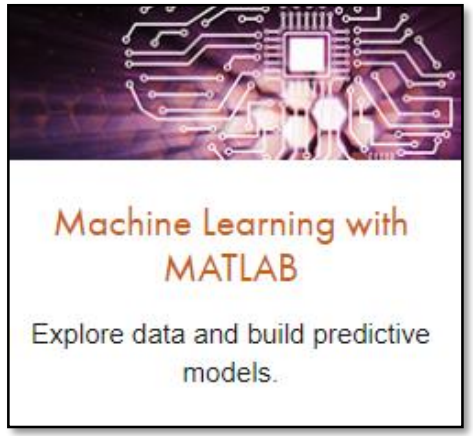


Further Learning

Poll

- Deep Learning Onramp
  - 2 hr online tutorial
- Hands-On Virtual Lab
  - 3 hr hands-on session
  - Contact us to schedule
- Deep Learning Training
  - 16 hr in-depth course
  - Online or Instructor-led





Onramps and Self-Paced Training

- Onramps

- Self-paced Training

- Instructor-led Training

https://matlabacademy.mathworks.com/

MATLAB Toolboxes for Deep Learning

- Deep Learning Toolbox
- Parallel Computing Toolbox

Image domain

- Image Processing Toolbox
- Computer Vision Toolbox
- Automated Driving Toolbox

Additional domains

- Reinforcement Learning Toolbox
- Text Analytics Toolbox
- Statistics and Machine Learning Toolbox

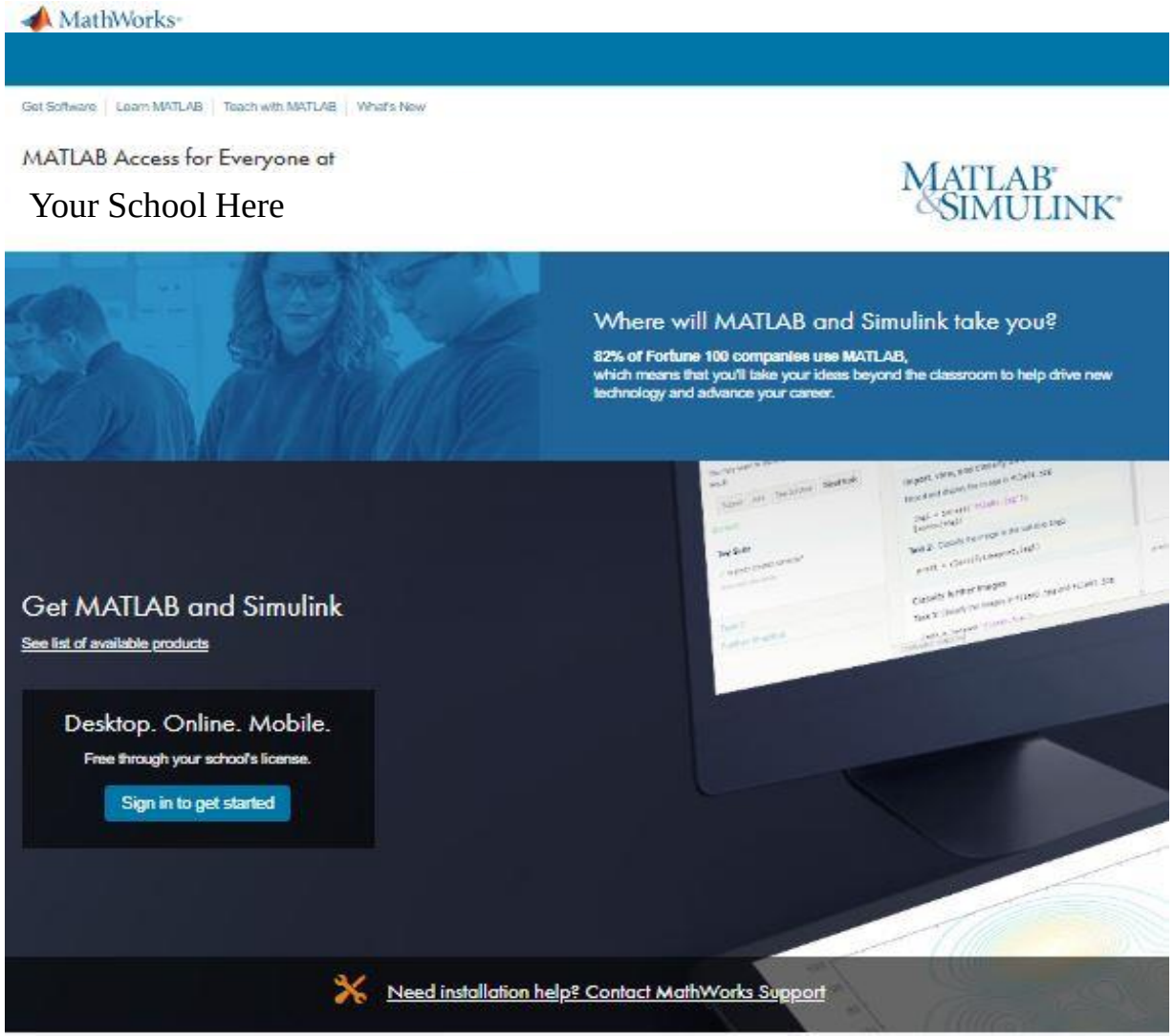
Signal domain

- Signal Processing Toolbox
- DSP System Toolbox
- Audio Toolbox
- Wavelet Toolbox
- Communications Toolbox
- Phased Array System Toolbox

Deployment

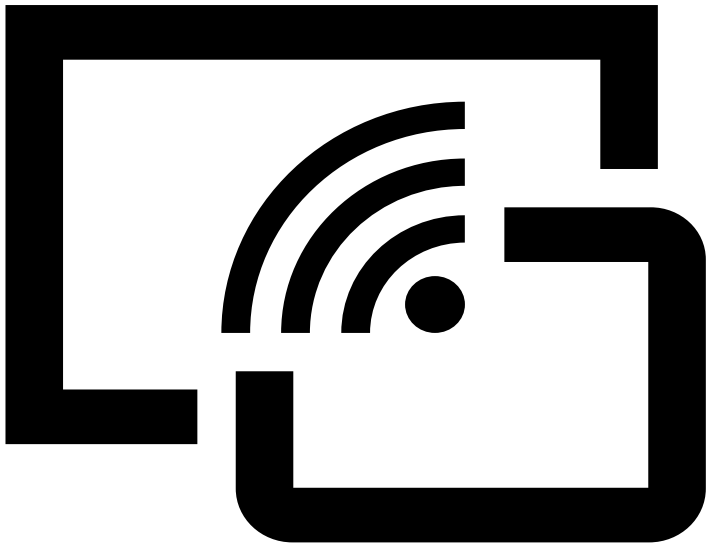
- MATLAB Coder
- GPU Coder
- MATLAB Parallel Server
- MATLAB Compiler

Campus-wide access

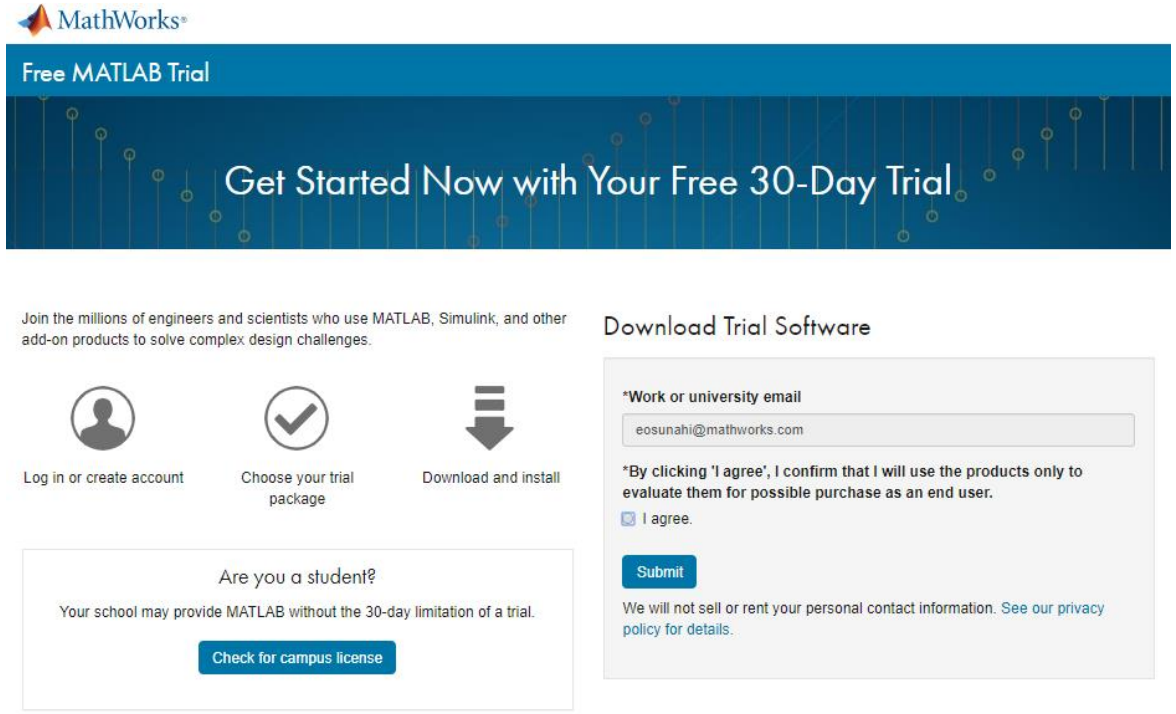


Access for universities without campus licenses

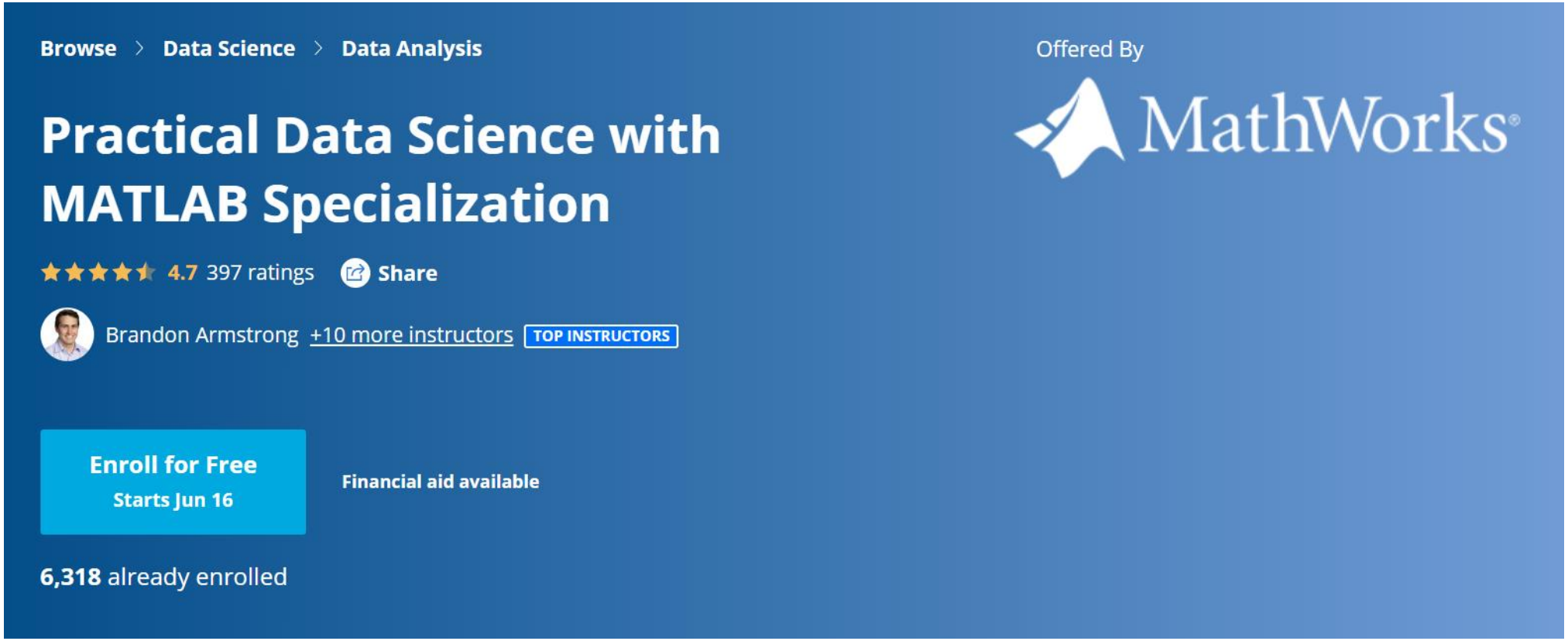
If available, access secure connection



For immediate needs, download 30 day trial



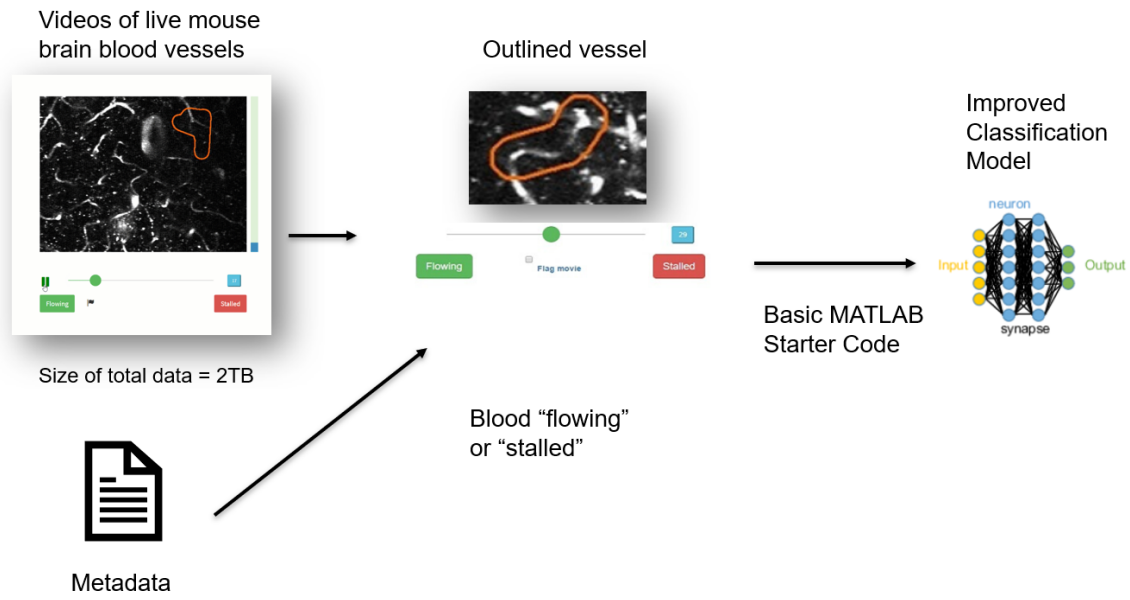
Data Science 4-Course Specialization



https://www.coursera.org/specializations/practical-data-science-matlab

Advance Alzheimer's Research Data Science Challenge

- Online data science competition hosted by DrivenData and supported by MathWorks
- No Eligibility Criteria
  - Open to students, professors, employers and at-home
  - Participate Individually or in a Team
  - Submit multiple entries
  - Online and around the globe
- Prizes worth $10K
- MathWorks support:
  - MATLAB benchmark code
  - Complimentary licenses
  - Technical assistance



Open to all until **August 3, 2020**

Register at : https://www.drivendata.org/competitions/65/clog-loss-alzheimers-research/

MathWorks Customer Success Engineers *education@mathworks.com*
consult with faculty and researchers to support them with their STEM initiatives,
including integrating computational or systems thinking into their curriculum.

MathWorks®

Thank you for joining us!

education@mathworks.com