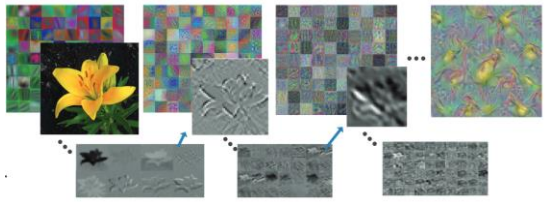
Deep Learning

A Practical Approach in MATLAB



David Willingham
david.willingham@mathworks.com

Senior Data Analytics Engineer

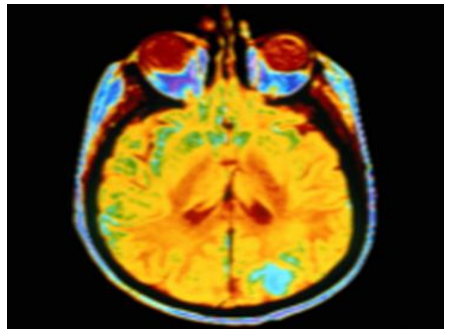
WITH SOFTWARE

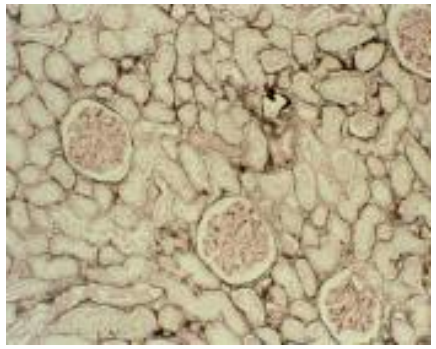
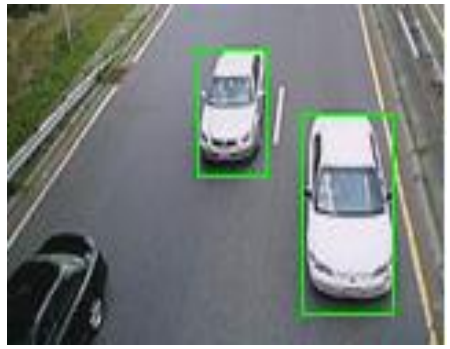
(and smart people)

ANYTHING IS POSSIBLE

Artificial Intelligence is Everywhere

- Image Recognition
- Speech Recognition
- Stock Prediction
- Medical Diagnosis
- Data Analytics
- Robotics
- and more…




50 km/h - sudden brake

What's in the trunk (aka boot) of an autonomous vehicle?

- SELF-DRIVING CARS USE CRAZY AMOUNTS OF POWER, AND IT'S BECOMING A PROBLEM – WIRED 6/2/2018

Deep Learning - end to end product development

Access and Explore Data

Files

Databases

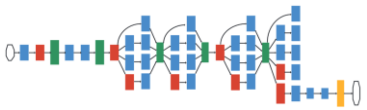
Sensors

Label and Preprocess Data

Clean Data

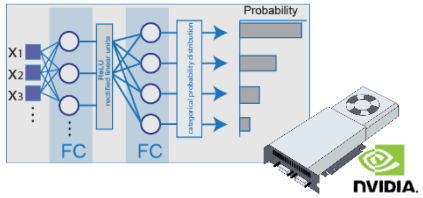
Feature Engineering

Select Model



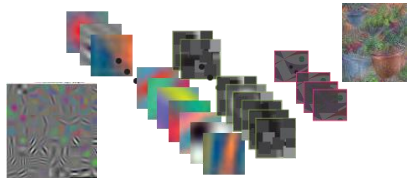
Develop Predictive Models

Train



Tune

Visualise

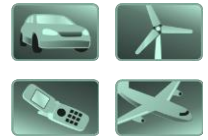


Integrate Models with Systems

Desktop Apps

Web Based Systems

On Device



Data Access

MATLAB Analytics work with business and engineering data

Business and Transactional Data

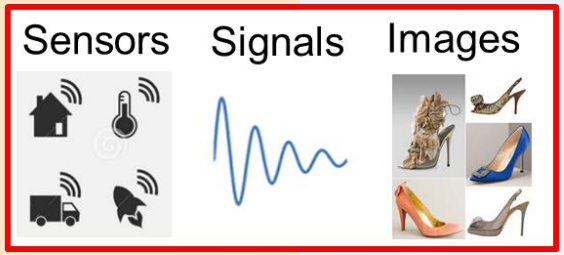
Repositories
- Databases (SQL)
- NoSQL
- Hadoop

File I/O
- Text
- Spreadsheet
- XML

Web Sources
- RESTful
- JSON
- HTML
- Mapping
- Financial datafeeds

Cloud
- Amazon S3
- Azure Blob

Streaming
- Amazon Kinesis
- Azure Event Hub
- Kafka
- MQTT



Engineering, Scientific, and Field Data

File I/O
- Text
- Spreadsheet
- XML
- CDF/HDF
- Image
- Audio
- Video
- Geospatial

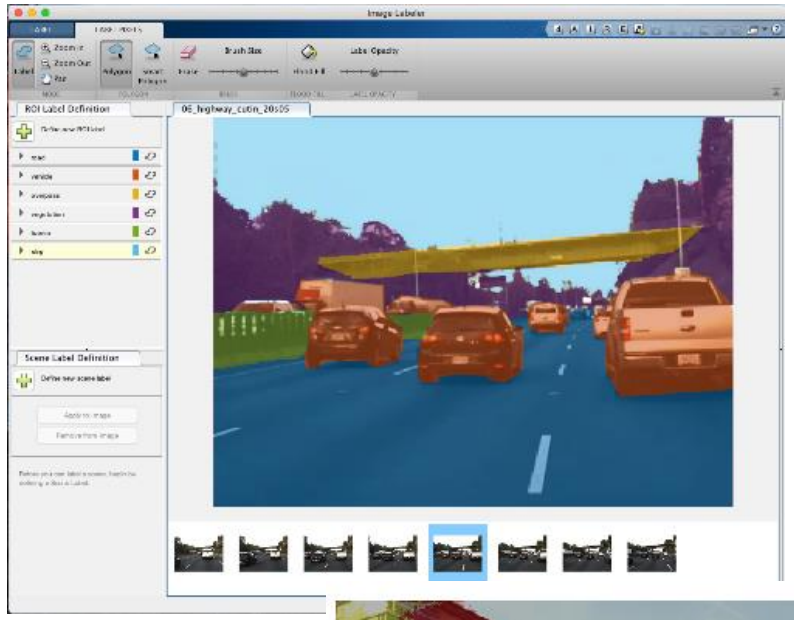
Communication Protocols
- CAN (Controller Area Network)
- DDS (Data Distribution Service)
- OPC (OLE for Process Control)
- XCP (eXplicit Control Protocol)

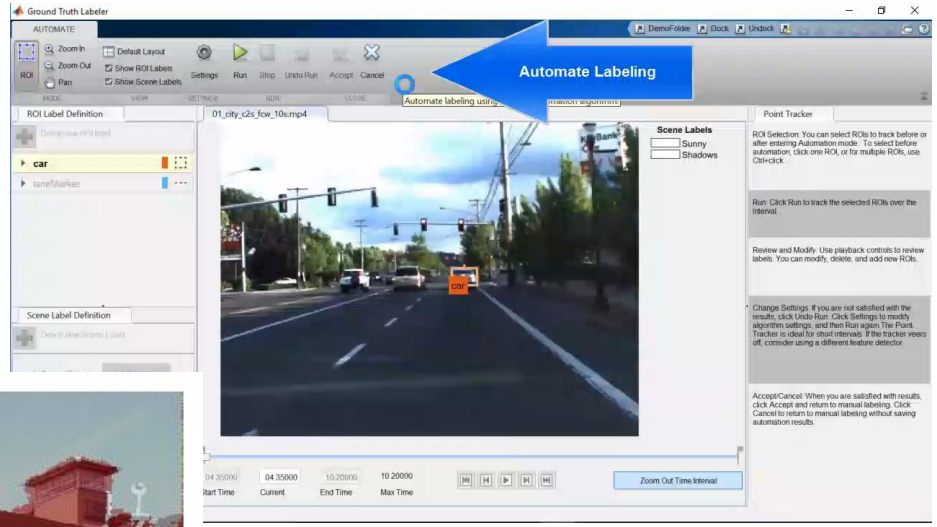
Real-Time Sources
- Sensors
- GPS
- Instrumentation
- Cameras
- Communication systems
- Machines (embedded systems)

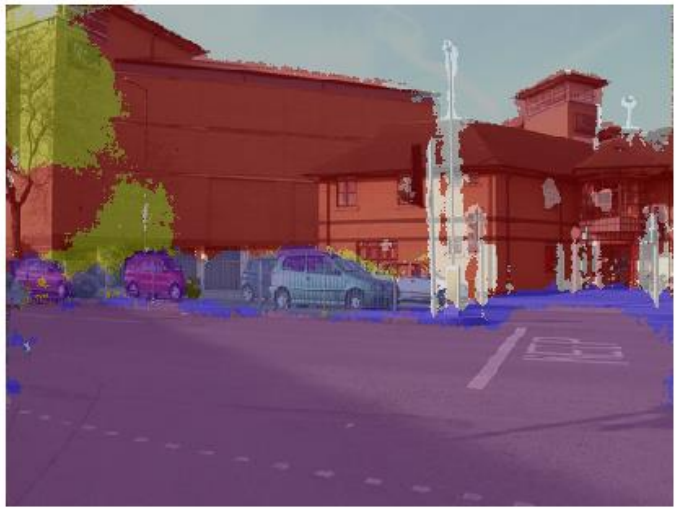
Pre Processing – MATLAB Automated Ground truth Labeling

New App for Ground Truth Labeling

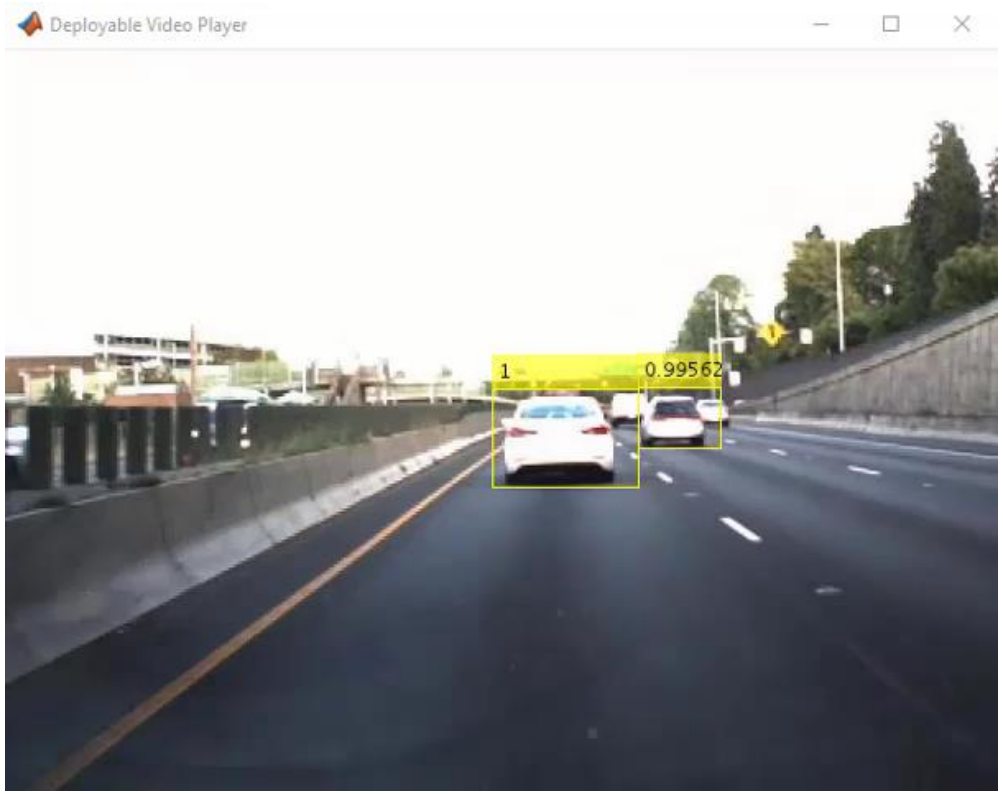
- **Automatically Create your own library of images from Video**
- Label pixels and regions for semantic segmentation







Detection and localization using deep learning

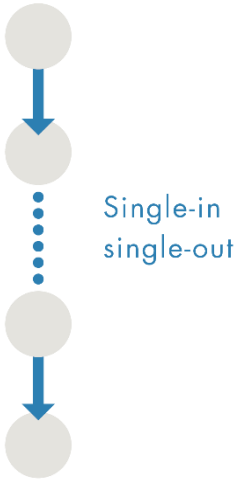


Regions with Convolutional Neural Network Features (R-CNN)

Semantic Segmentation using SegNet

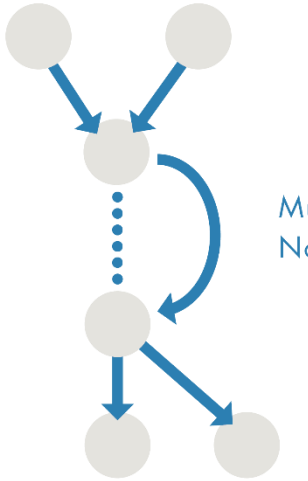
Deep Learning Model Support

SeriesNetwork



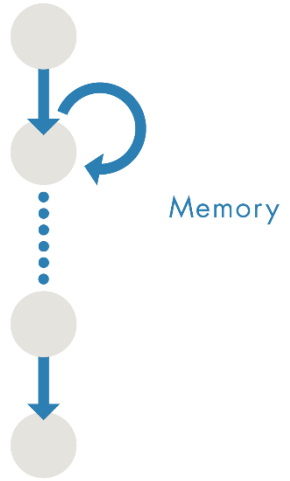
Networks:
- MNIST
- Alexnet
- VGG
- Lane detection
- Pedestrian detection

DAGNetwork



Networks:
- R-CNN (fast, faster) — *Object detection*
- GoogLeNet
- ResNet
- SegNet — *Semantic segmentation*
- FCN
- DeconvNet

Recurrent Network



Networks:
- LSTM (timeseries)

Import the Latest Models for Transfer Learning

Pretrained Models*

- AlexNet
- VGG-16
- VGG-19
- GoogLeNet
- InceptionV3
- Resnet50

Import Models from Frameworks

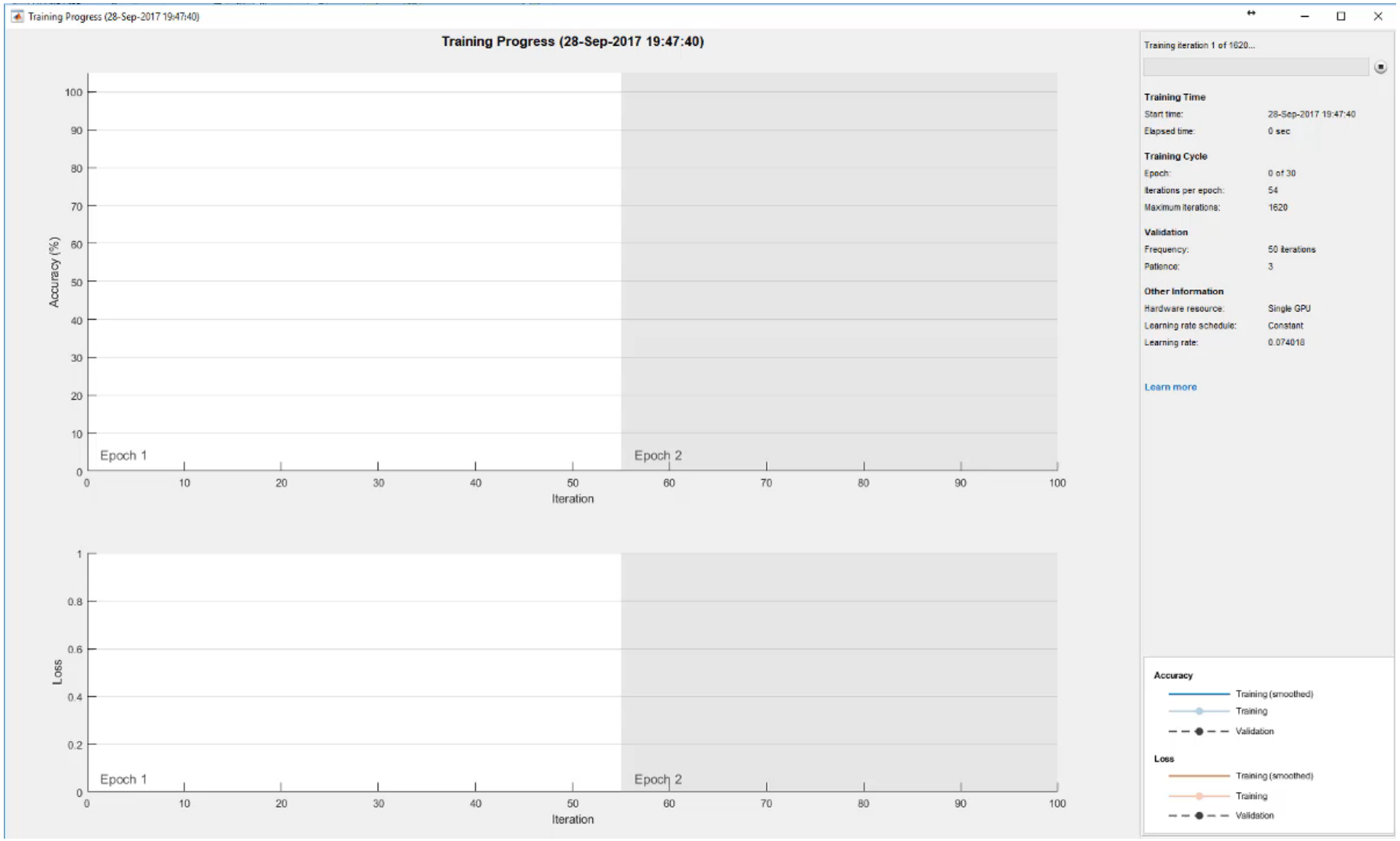
- Caffe Model Importer
- TensorFlow/Keras Model Importer

* single line of code to access model

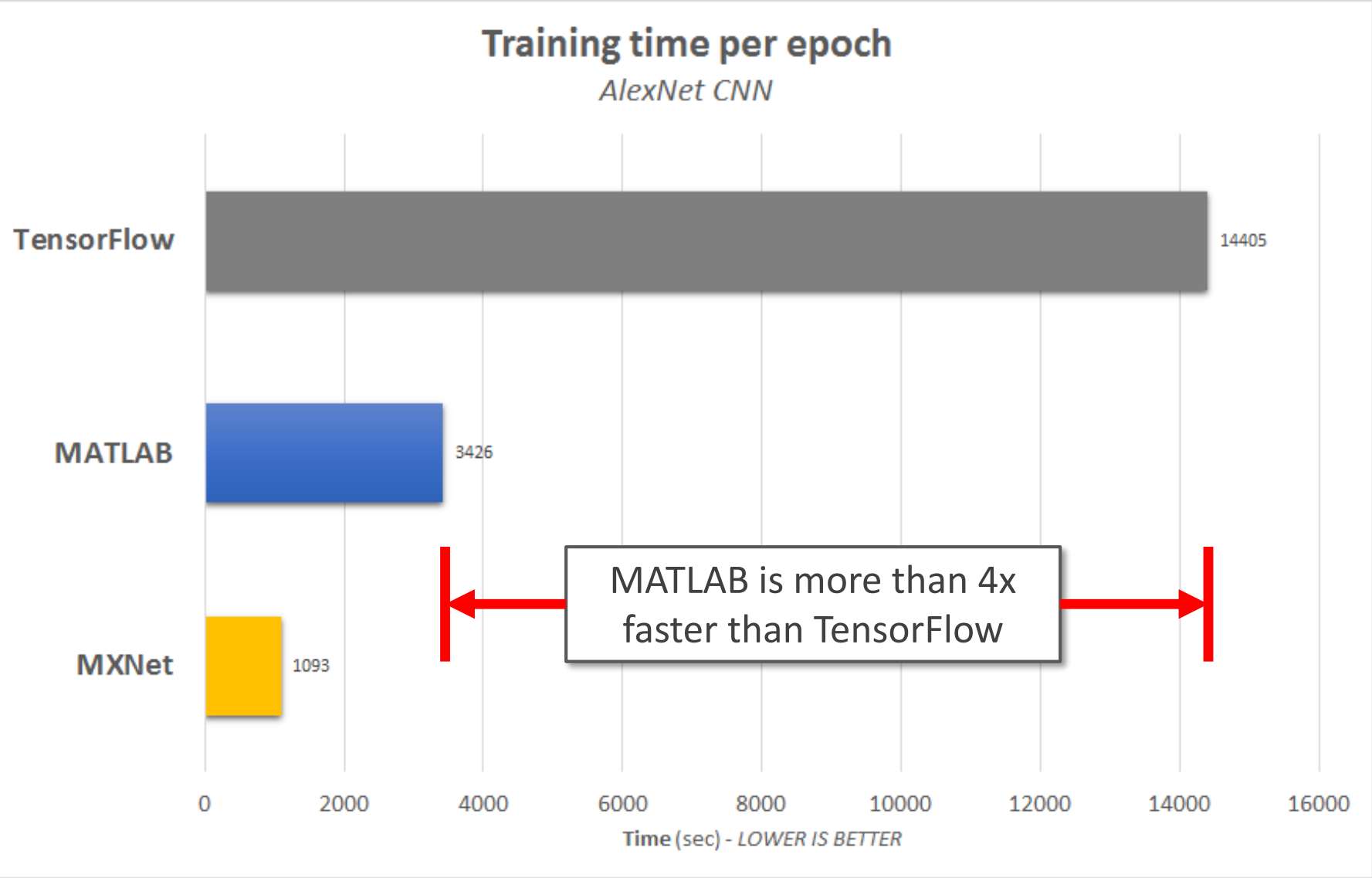
Deep Learning – In 5 lines of MATLAB code

```
load chardata

layers = createLayers;

options = trainingOptions( 'sgdm','MiniBatchSize', 8192,'Plots',
'training-progress','ExecutionEnvironment','multi-gpu');

net = trainNetwork(imgDataTrain, labelsTrain, layers, options);

predLabelsTest = net.classify(imgDataTest);
```

Live Training Progress



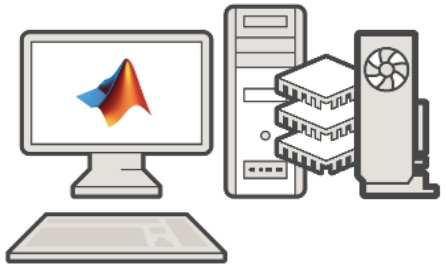
Training in MATLAB is fast



AlexNet CNN architecture trained on the ImageNet dataset, using batch size of 32, on a Windows 10 desktop with single NVIDIA GPU

Deep learning on CPU, GPU, multi-GPU and clusters

Single CPU

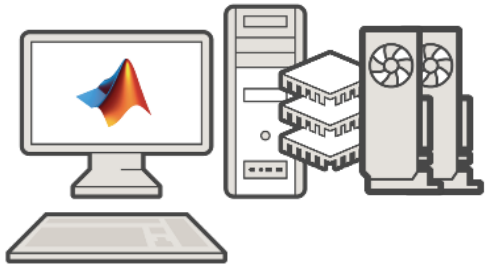


Single CPU
Single GPU

How to target?

```
opts = trainingOptions('sgdm', ...
    'MaxEpochs', 100, ...
    'MiniBatchSize', 250, ...
    'InitialLearnRate', 0.00005, ...
'ExecutionEnvironment', 'auto' );
```

Deep learning on CPU, GPU, multi-GPU and clusters

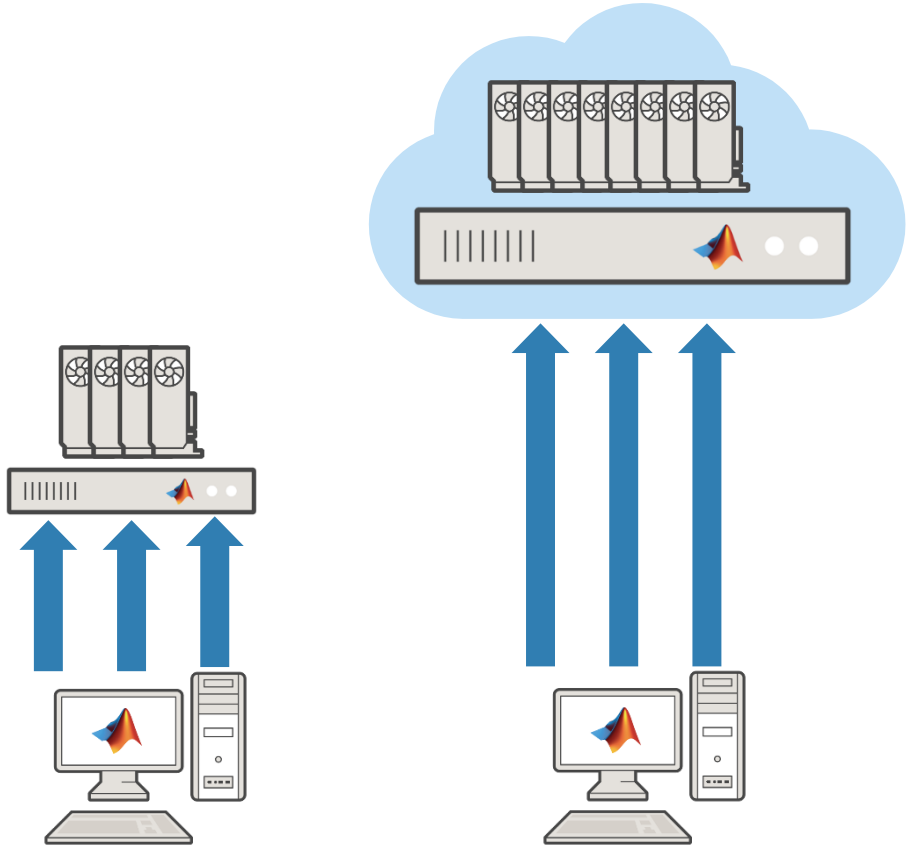


Single CPU
Multiple GPUs

HOW TO TARGET?

```
opts = trainingOptions('sgdm', ...
    'MaxEpochs', 100, ...
    'MiniBatchSize', 250, ...
    'InitialLearnRate', 0.00005, ...
'ExecutionEnvironment', 'multi-gpu' );
```

Deep learning on CPU, GPU, multi-GPU and clusters



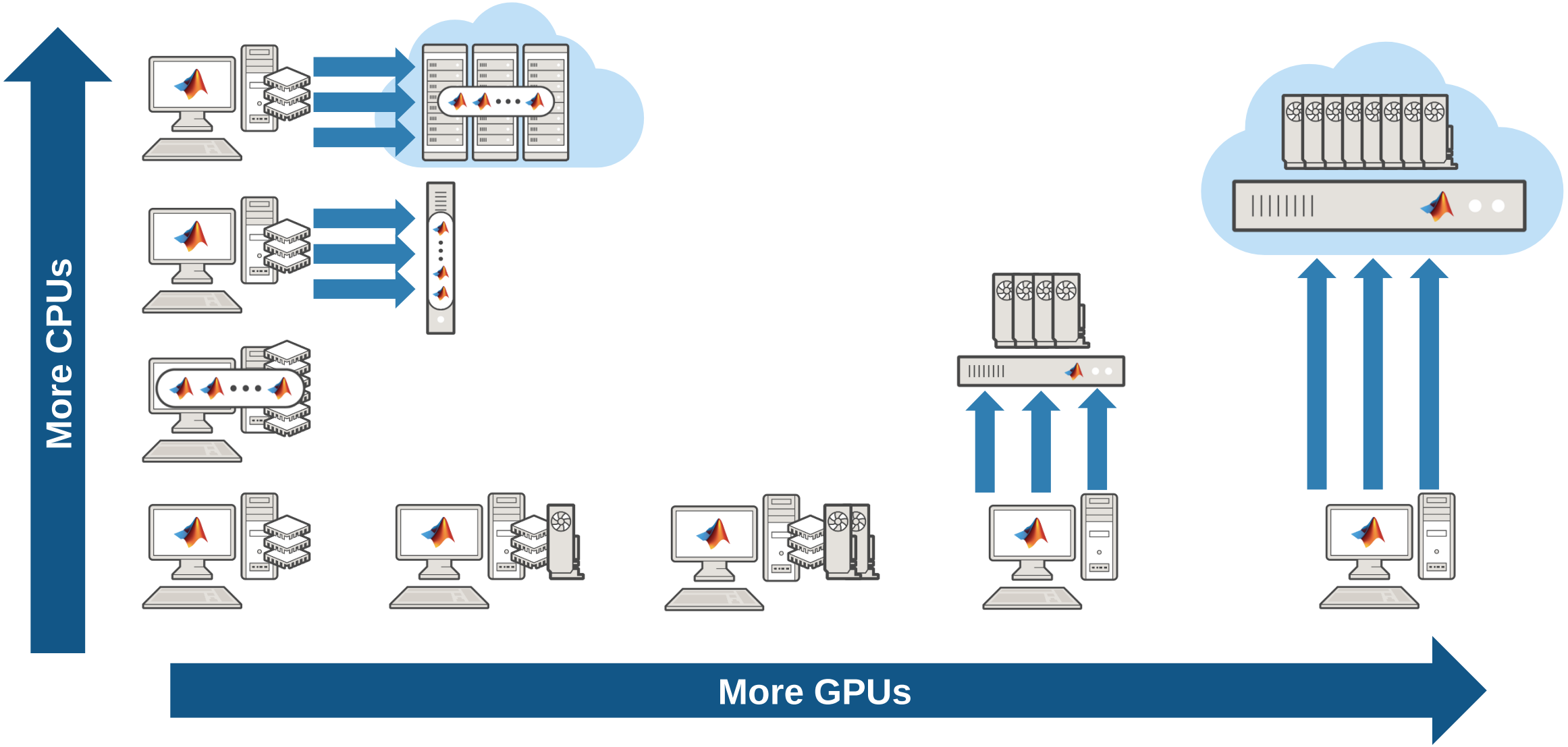
On-prem server with GPUs

Cloud GPUs (AWS, Azure, etc.)

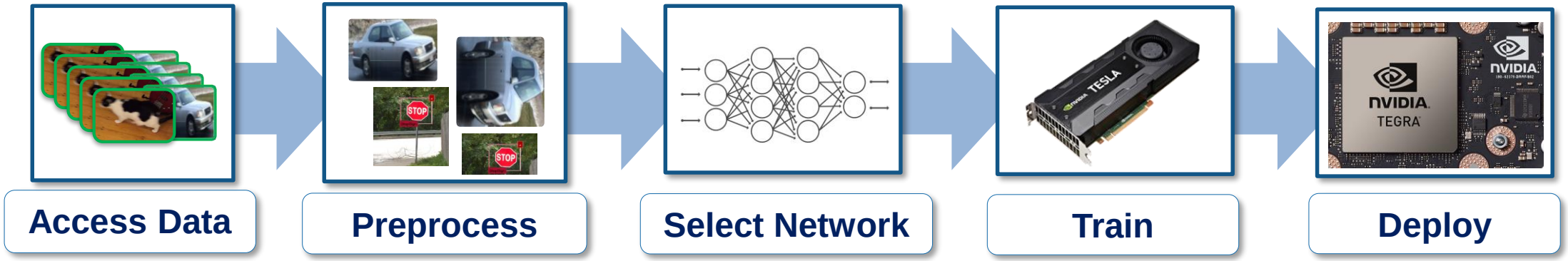
How to target?

```
opts = trainingOptions('sgdm', ...
    'MaxEpochs', 100, ...
    'MiniBatchSize', 250, ...
    'InitialLearnRate', 0.00005, ...
'ExecutionEnvironment', 'parallel' );
```

Deep learning on CPU, GPU, multi-GPU and clusters



Deep Learning Deployment Options

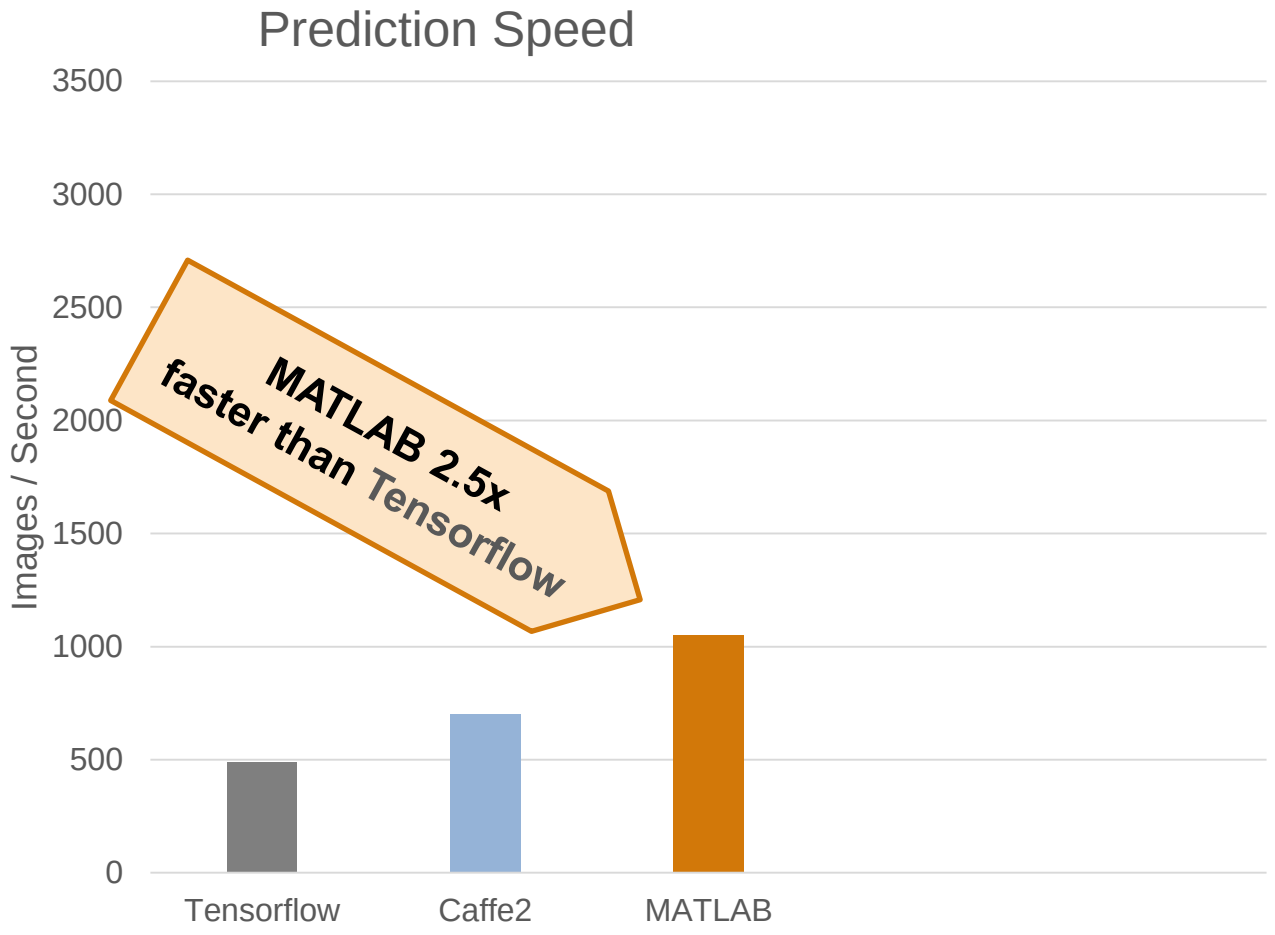


Running a Trained Model

"How do I make my model **run** faster?"

30 fps

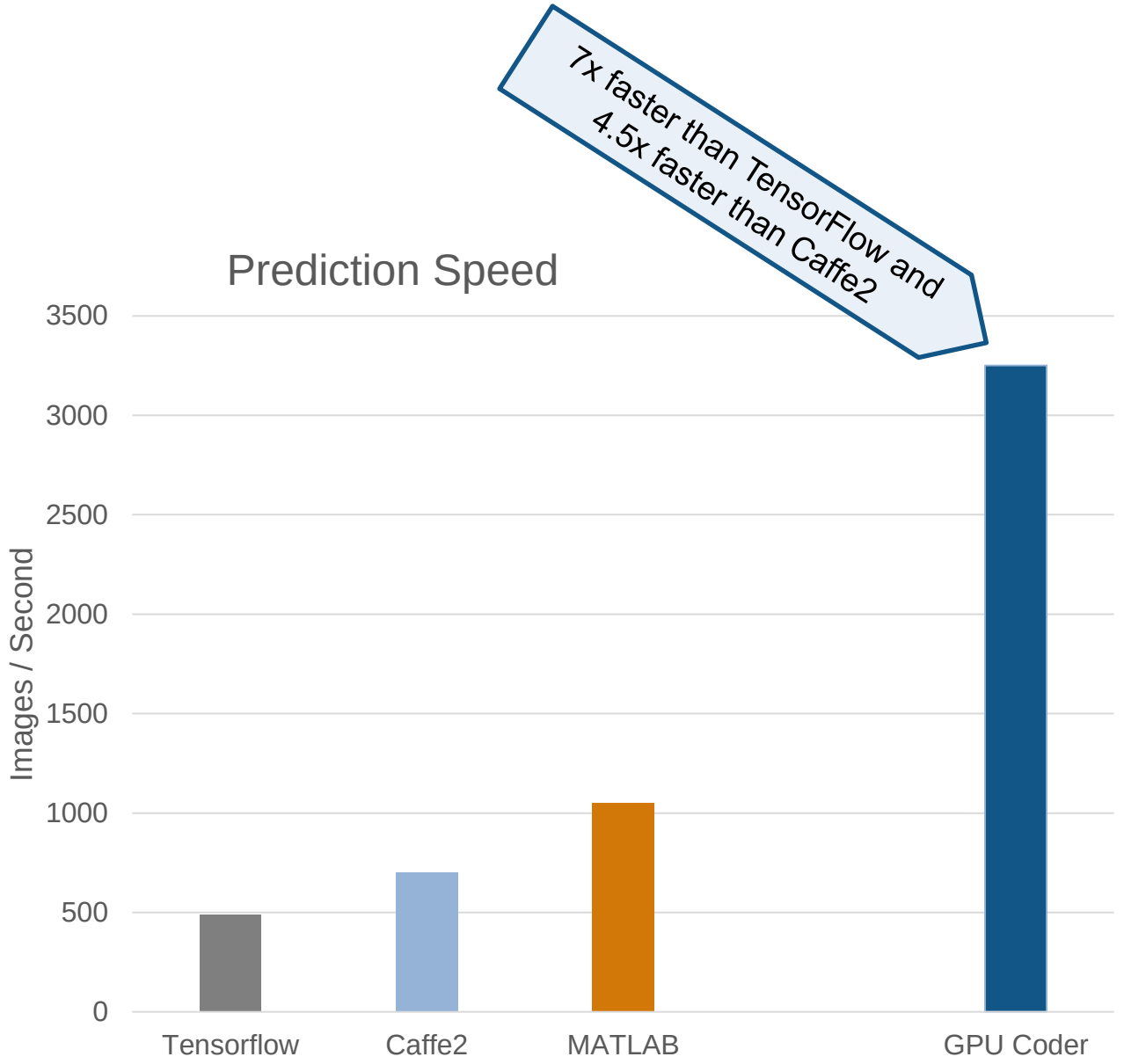
10 fps



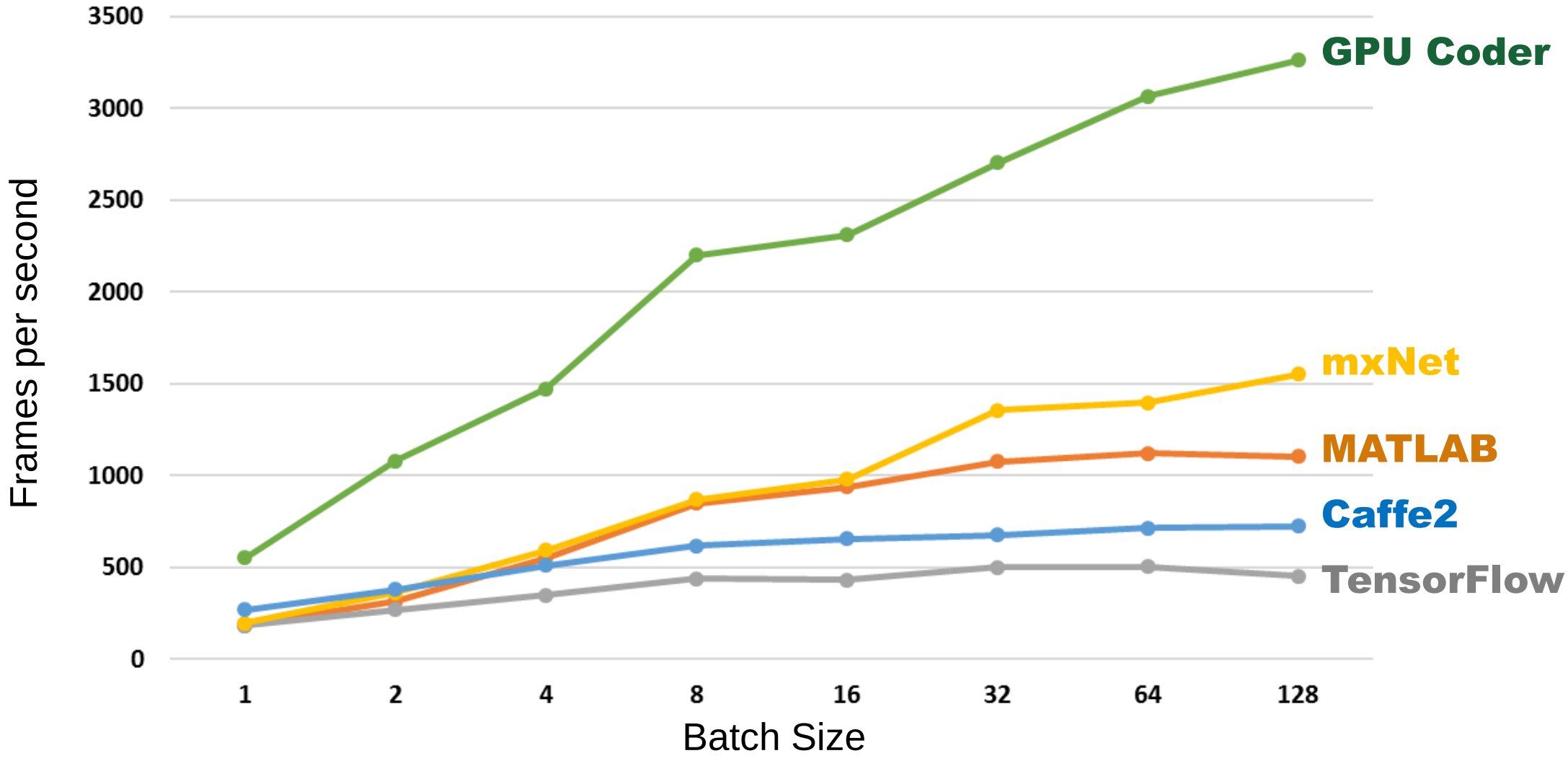
** Internal benchmarks were performed for inference performance of AlexNet using a TitanXP GPU and Intel(R) Xeon(R) CPU E5-1650 v4 @ 3.60GHz. Software versions used were MATLAB(R2017b), TensorFlow(1.2.0), and Caffe2(0.8.1). The GPU accelerated versions of each software were used for benchmarks. All tests were run on Windows 10.*

Running a Trained Model

GPU Coder- Convert to NVIDIA CUDA code



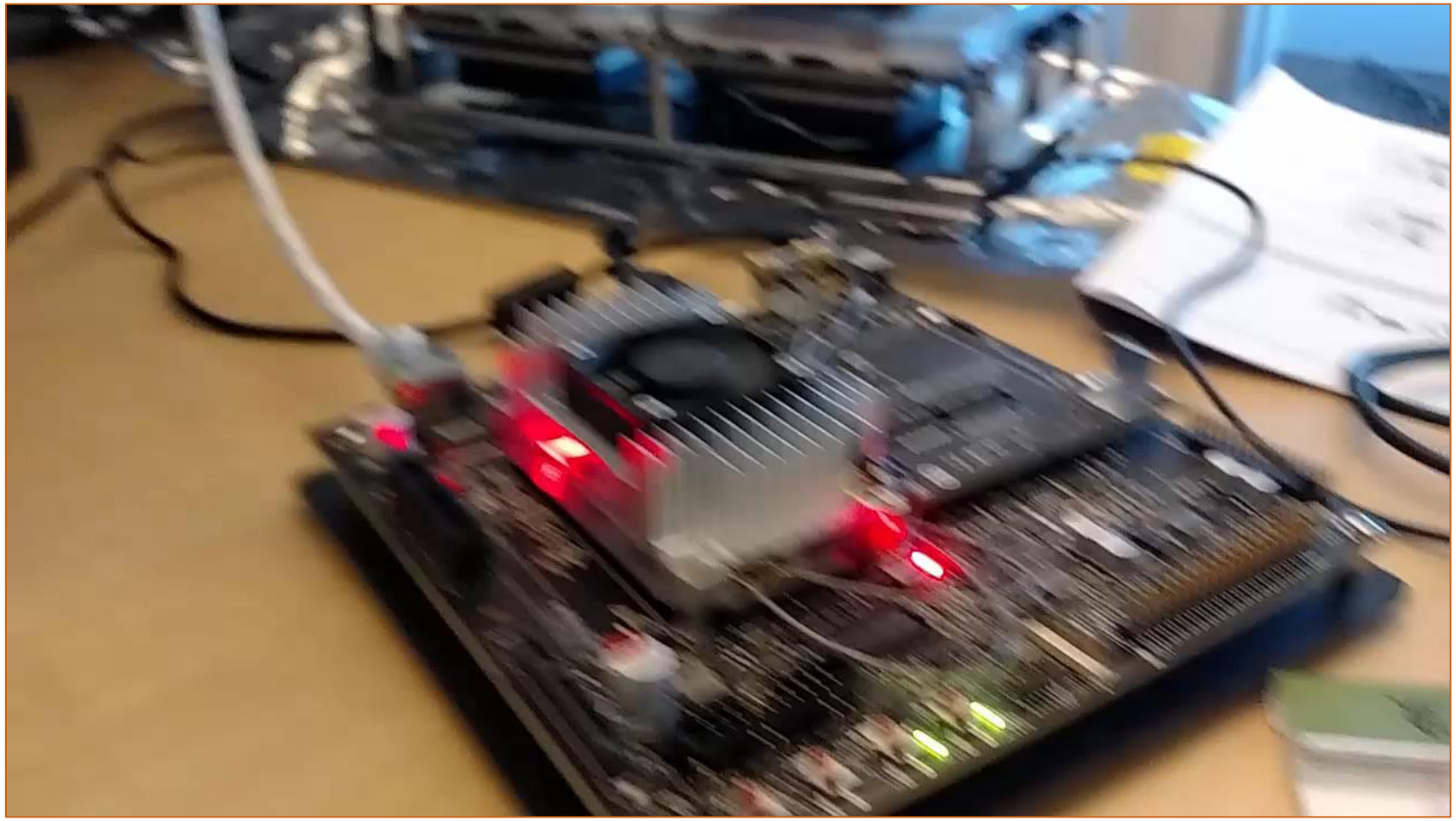
AlexNet Inference on NVIDIA Titan XP GPU



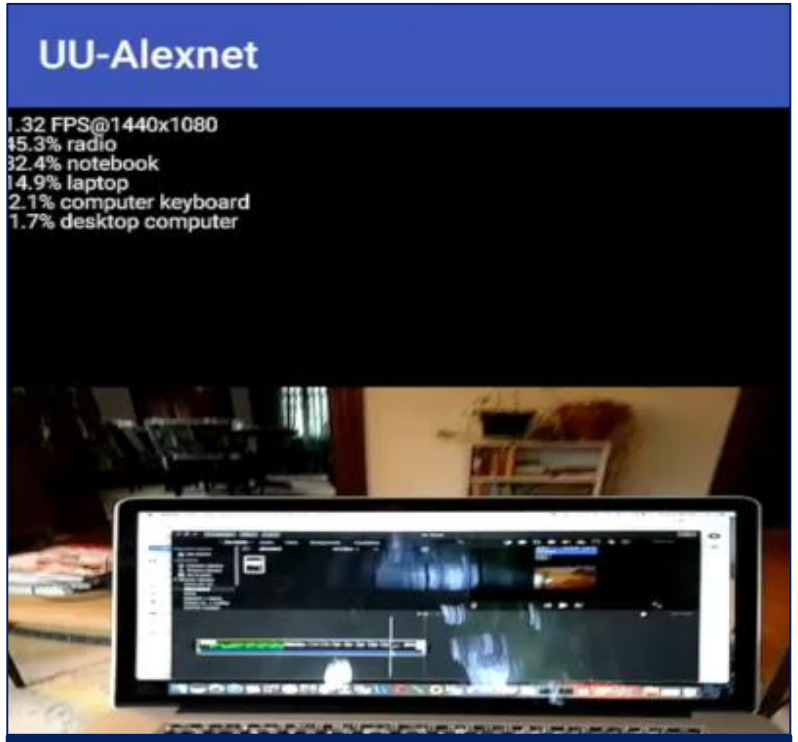
Testing platform

CPU	Intel(R) Xeon(R) CPU E5-1650 v3 @ 3.50GHz
GPU	Pascal TitanXP

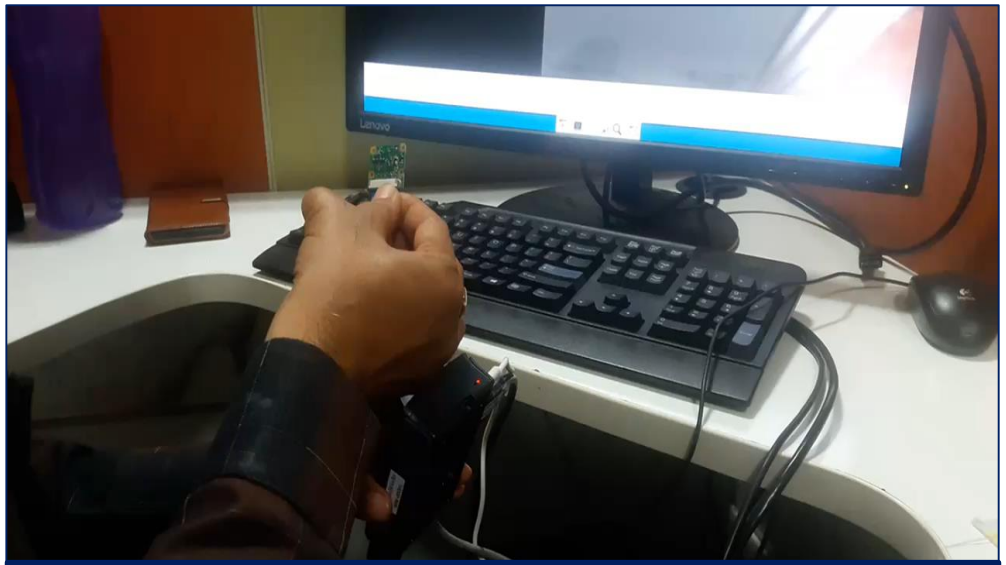
MATLAB GPU Coder



More Deployment options coming



Android phone



Raspberry Pi board

Deep Learning - end to end product development

Access and Explore Data

Files

Databases

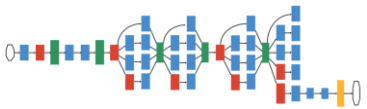
Sensors

Label and Preprocess Data

Clean Data

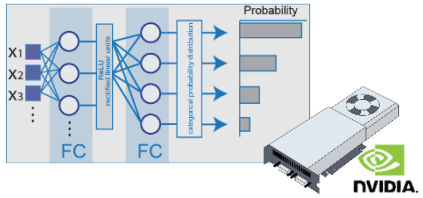
Feature Engineering

Select Model



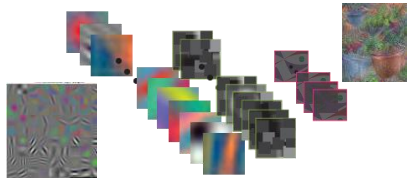
Develop Predictive Models

Train



Tune

Visualise

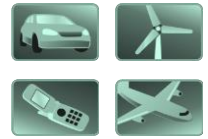


Integrate Models with Systems

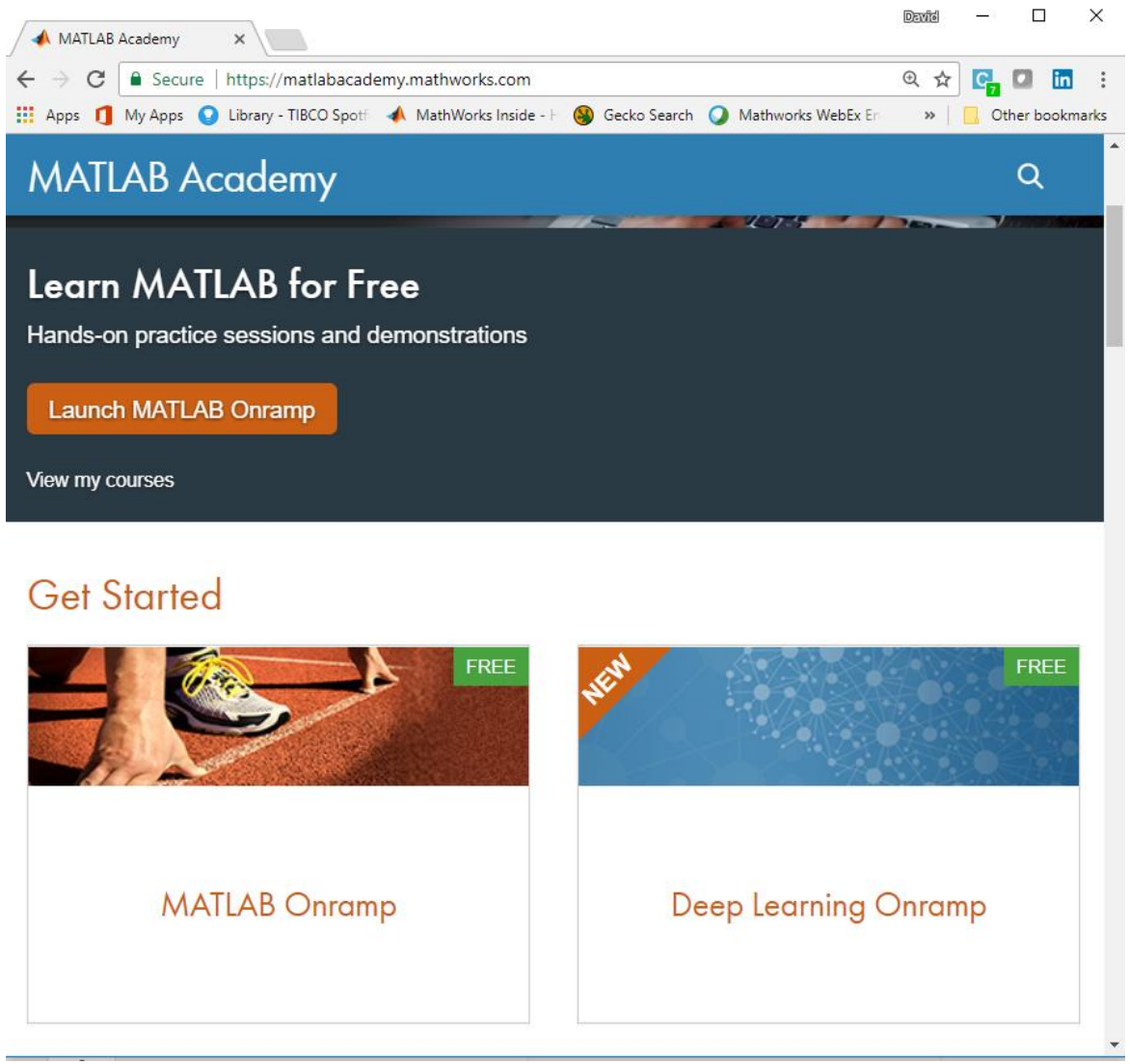
Desktop Apps

Web Based Systems

On Device



More Resources - Learn Deep Learning for free



MathWorks® can help you do Deep Learning

Free resources

- **Guided evaluations with a MathWorks deep learning engineer**
- Proof-of-concept projects
- **FREE online Deep learning ramp up**
- Seminars and technical deep dives
- Documentation and user community

Other options

- Consulting services
- Training courses
- Technical support
- Advanced customer support
- Installation, enterprise, and cloud deployment
- Deep Learning: Top 7 Ways to Get Started with MATLAB