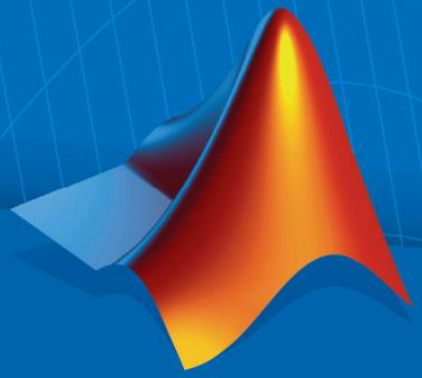
TeraSoft
www.terasoft.com.tw



Demystifying Deep Learning

"Let the computers do the hard work"

Phoebe Li, Terasoft Inc

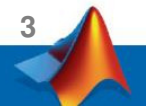
MathWorks®
Accelerating the pace of engineering and science

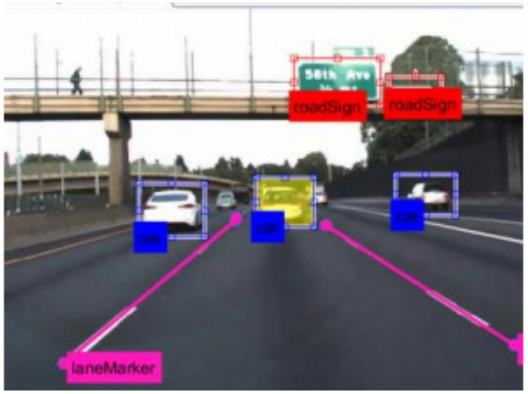
© Terasoft, Inc.

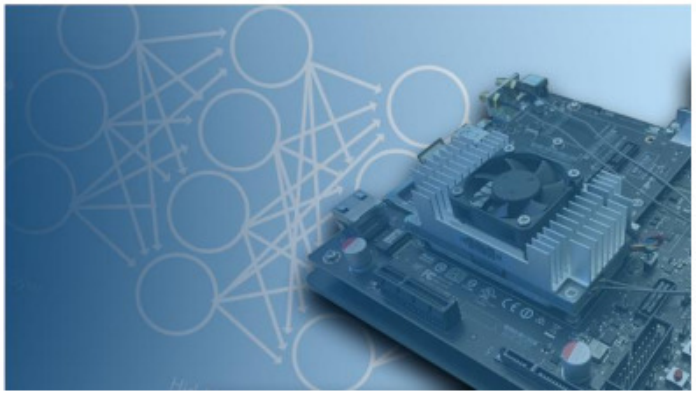
Why MATLAB for Deep Learning?

- MATLAB is Productive
- MATLAB is Fast
- MATLAB Integrates with Open Source

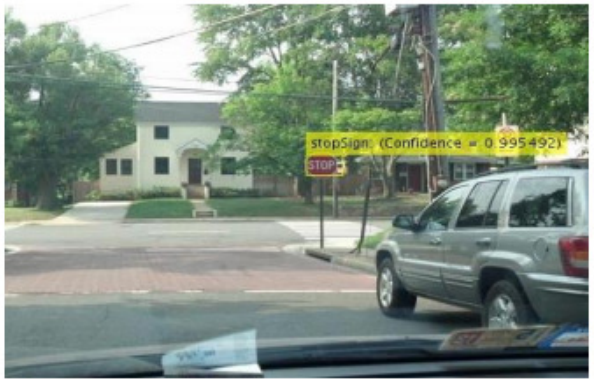
© Terasoft, Inc.

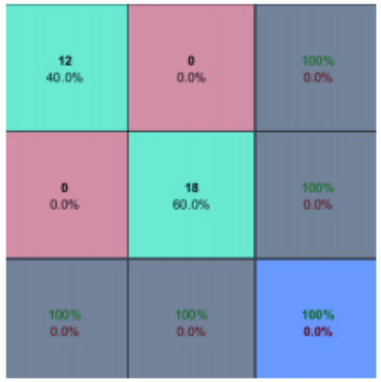






What is Deep Learning?

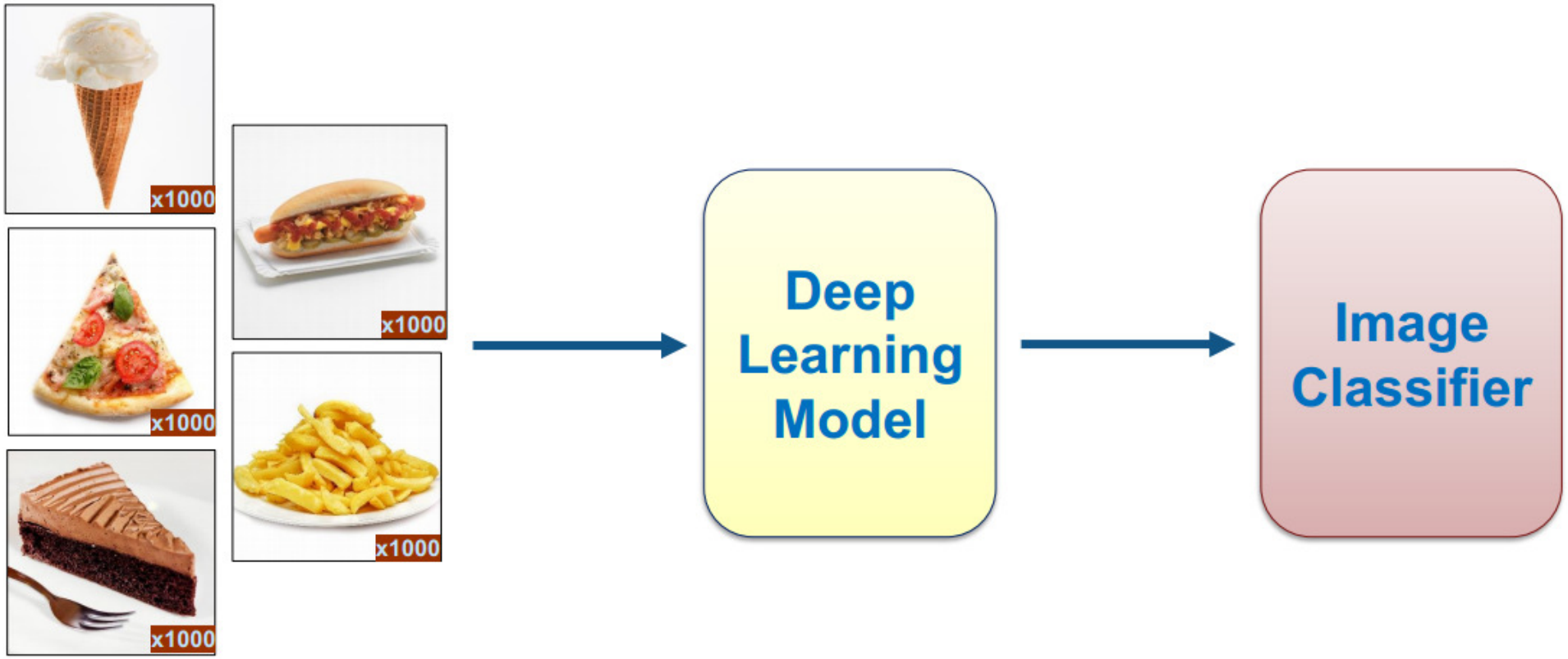




© Terasoft, Inc.

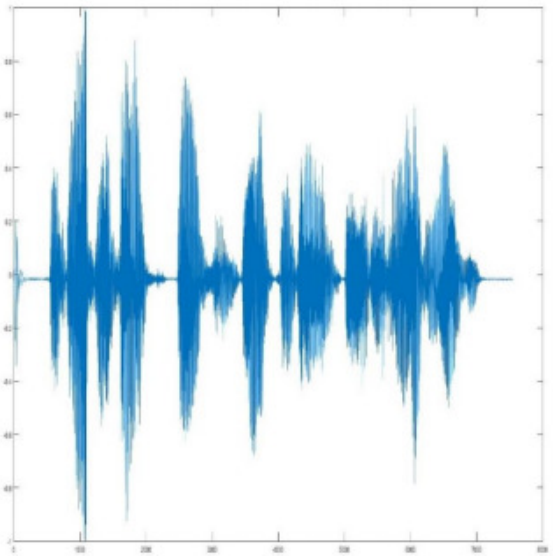
Deep Learning

Model learns to perform tasks directly from data.



© Terasoft, Inc.

Data Types for Deep Learning



Signal

Text

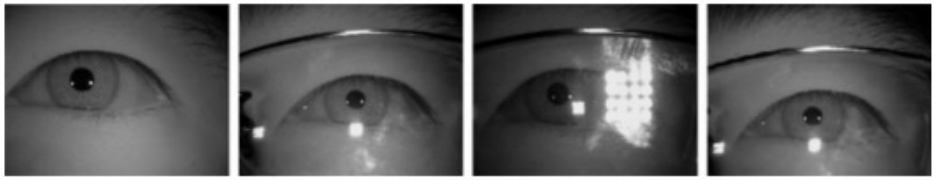
Image

© Terasoft, Inc.

Deep Learning is Versatile

Detection of cars and road in autonomous driving systems

Rain Detection and Removal[1]



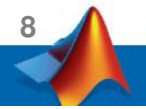
Iris Recognition – 99.4% accuracy[2]

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*

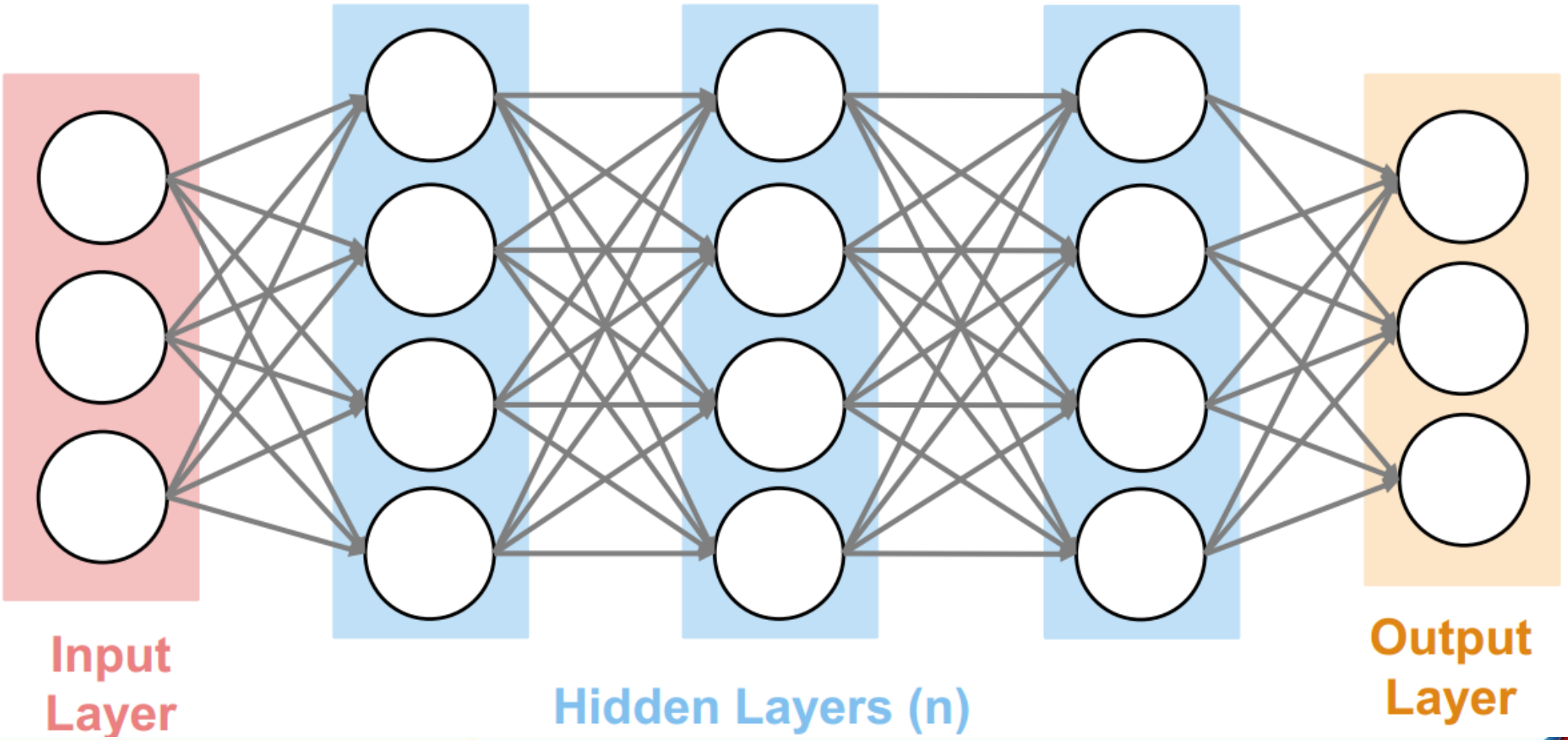
© Terasoft, Inc.

How is deep learning performing so well?

© Terasoft, Inc.

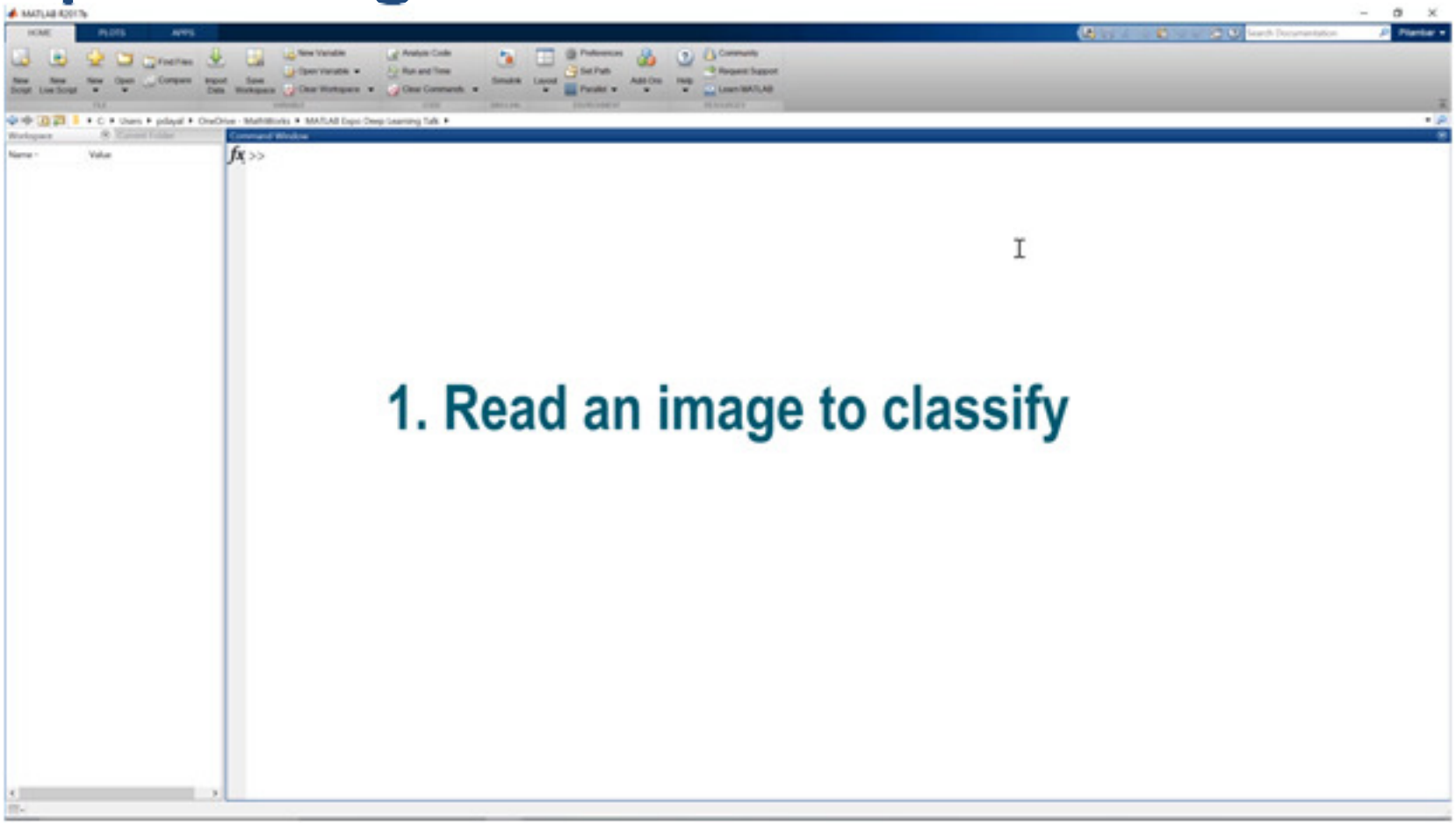


Deep Learning Uses a Neural Network Architecture



© Terasoft, Inc.

Deep Learning in 6 Lines of MATLAB Code

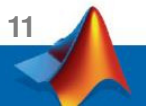


© Terasoft, Inc.

Why MATLAB for Deep Learning?

- MATLAB is Productive
- MATLAB is Fast
- MATLAB integrates with Open Source

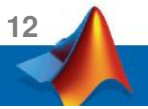
© Terasoft, Inc.



“I love to label and preprocess my data”

~ Said no engineer, ever.

© Terasoft, Inc.



Caterpillar Case Study

- World's leading manufacturer of construction and mining equipment.
- Similarity between these projects?
  - Autonomous haul trucks
  - Pedestrian detection
  - Equipment classification
  - Terrain mapping

© Terasoft, Inc.

Computer Must Learn from Lots of Data

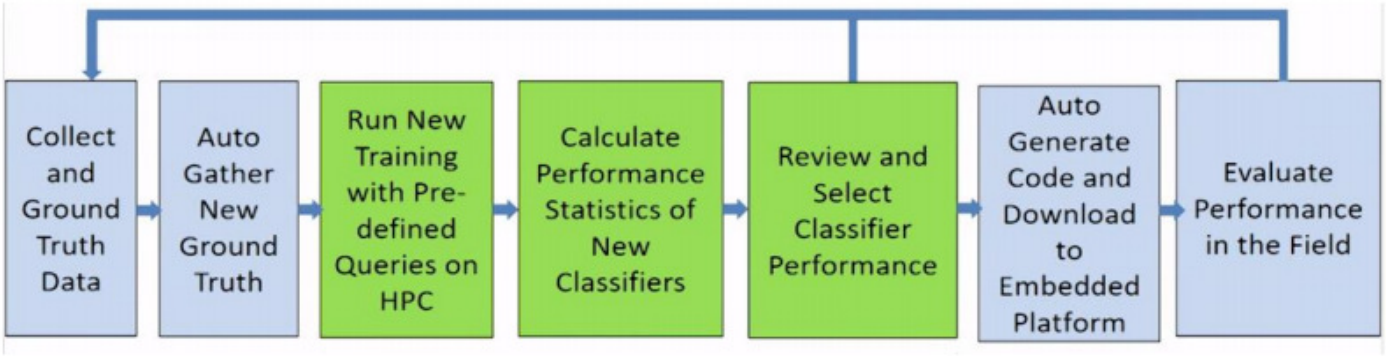
- ALL data must first be labeled to create these autonomous systems.

"We were spending way too much time ground-truthing [the data]"
--Larry Mianzo, Caterpillar

© Terasoft, Inc.

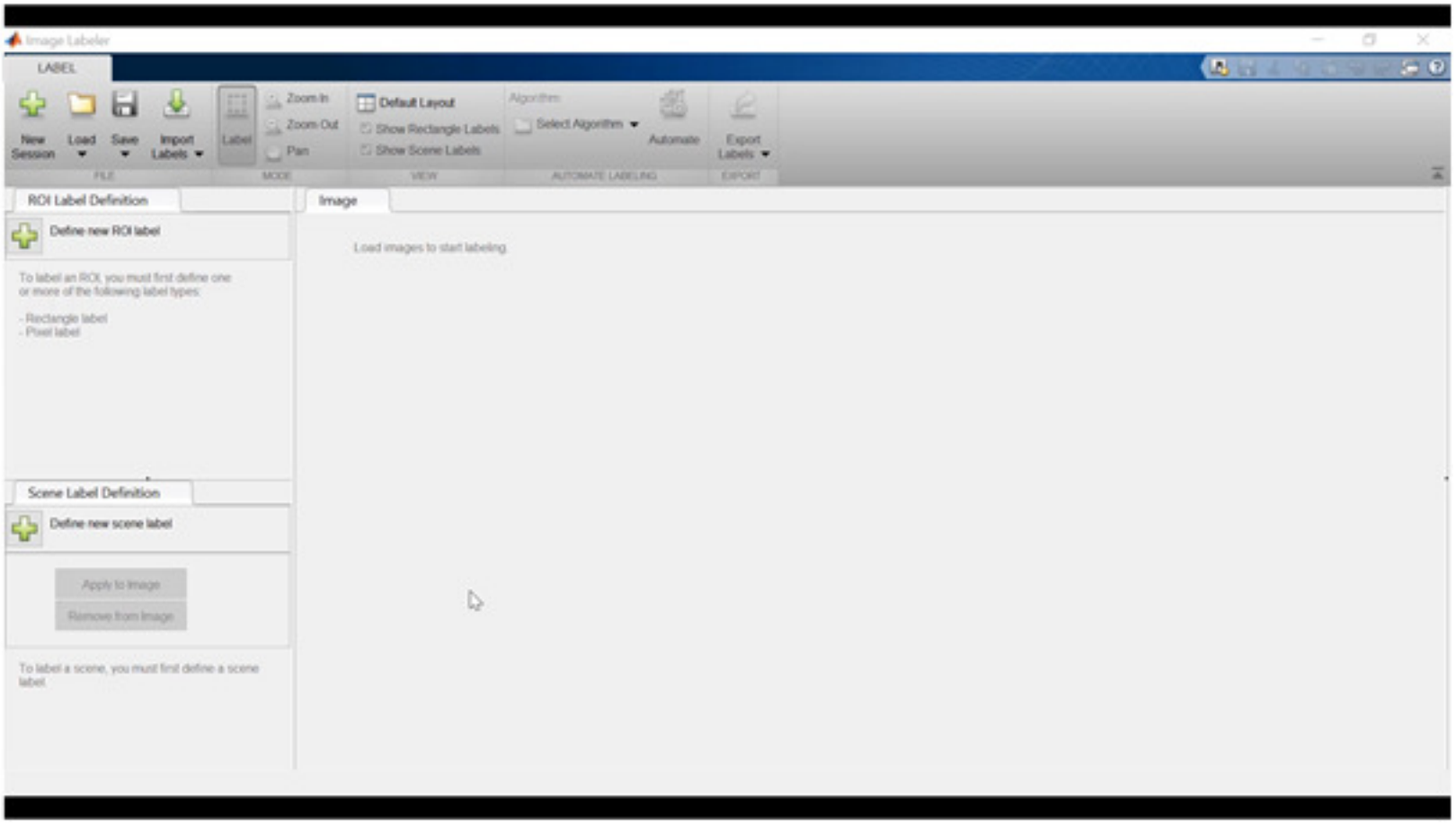
How Did Caterpillar Do with Our Tools?

- Semi-automated labeling process
  - *"We go from having to label 100 percent of our data to only having to label about 80 to 90 percent"*
- Used MATLAB for entire development workflow.
  - *"Because everything is in MATLAB, development time is short"*

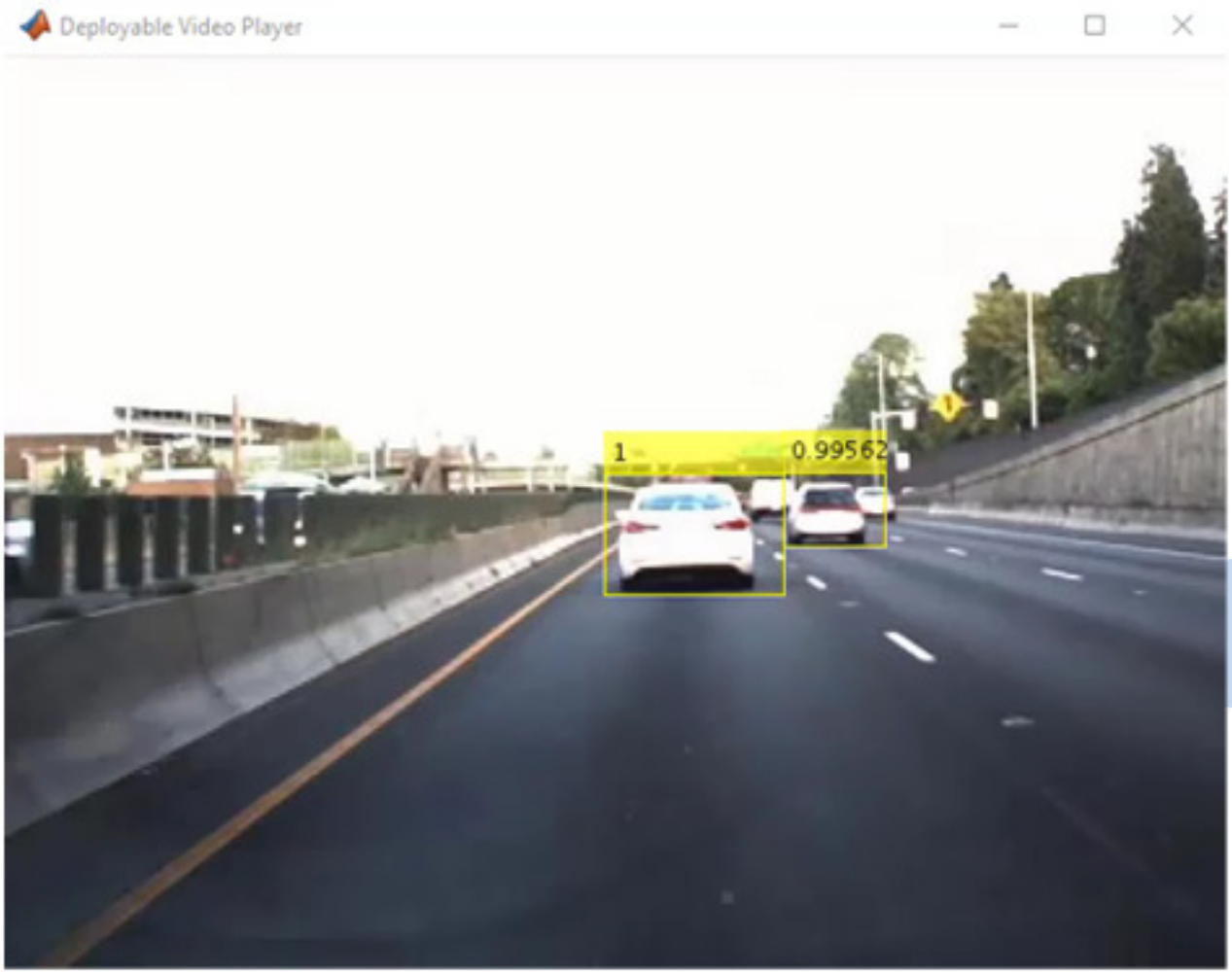


© Terasoft, Inc.

How Does MATLAB Come into Play?



© Terasoft, Inc.



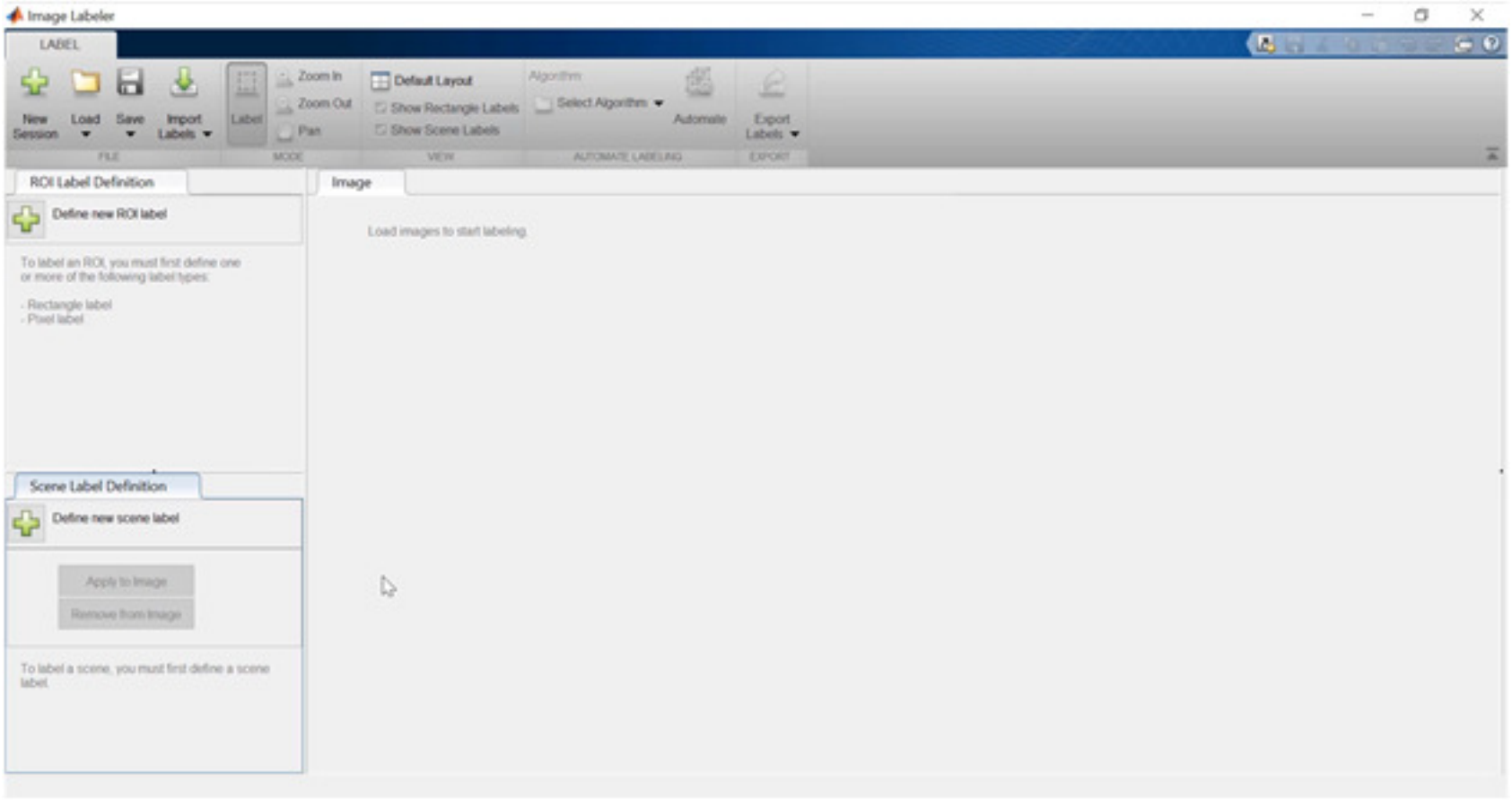
Regions with Convolutional Neural Network Features (R-CNN)

© Terasoft, Inc.

Labeling Videos: Ground Truth Labeler App

© Terasoft, Inc.

Labeling pixels



© Terasoft, Inc.

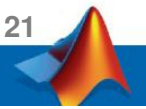
Semantic Segmentation using SegNet

© Terasoft, Inc.

MATLAB is Productive

- Image Labeler App semi-automates labeling workflow
- Improve automatic labeling by updating algorithm as you label more images correctly.
- Easy to load metadata even when labeling manually

© Terasoft, Inc.

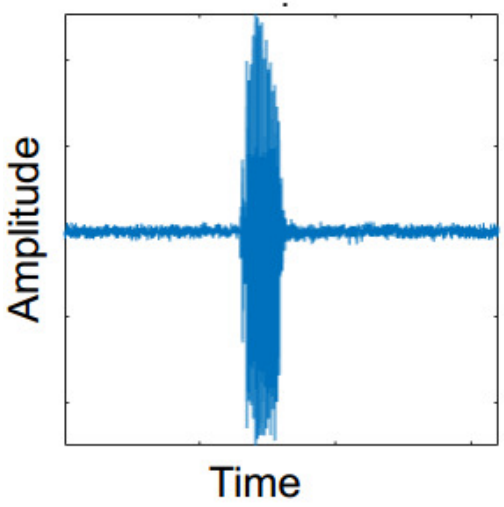


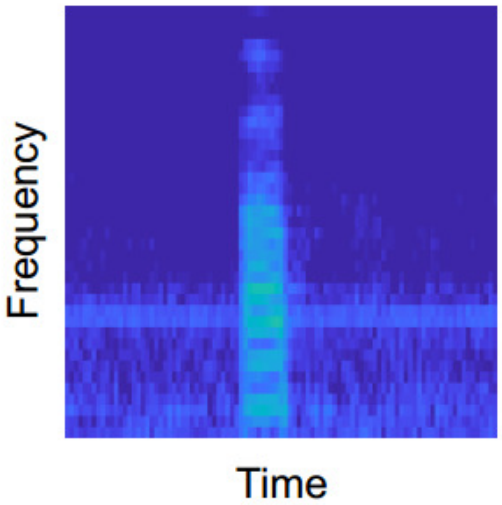
MATLAB is Fast

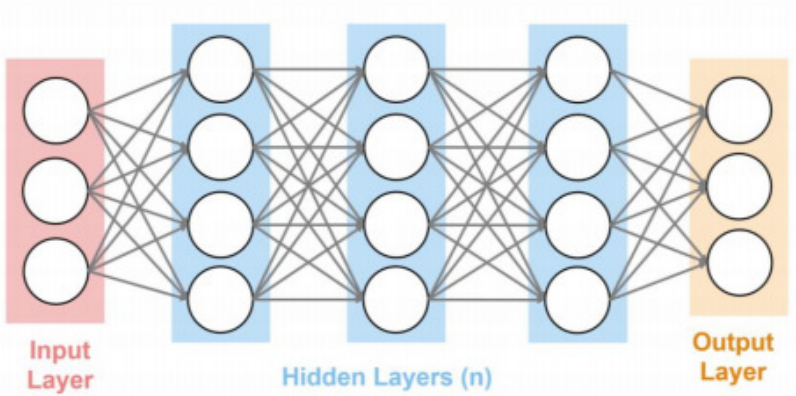
© Terasoft, Inc.

Speech Recognition Example

Audio signal → Spectrogram → Image Classification algorithm



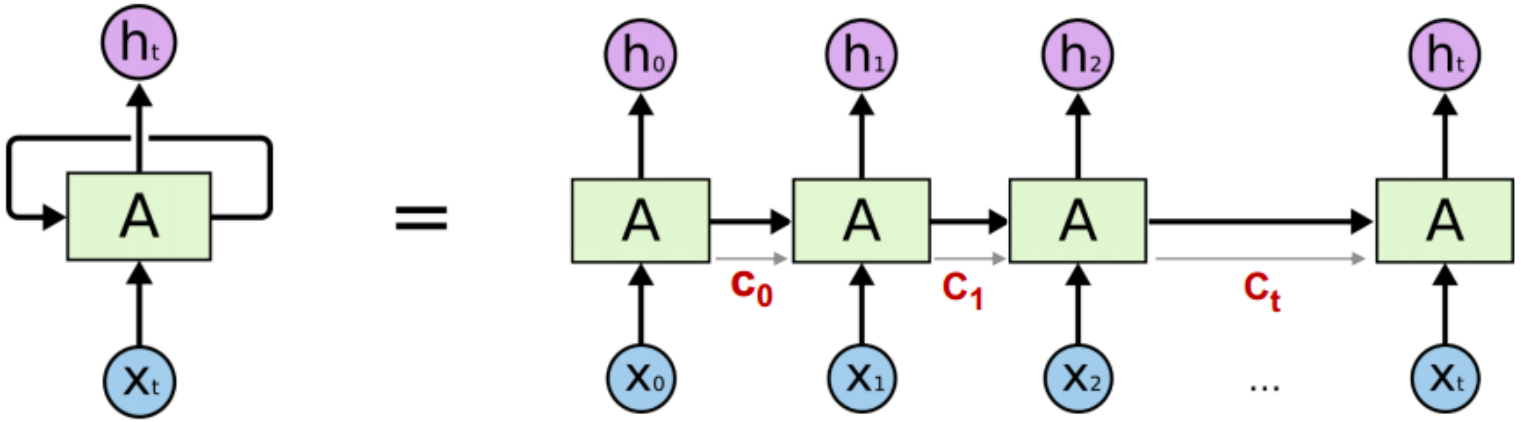




© Terasoft, Inc.

Another Network for Signals - LSTM

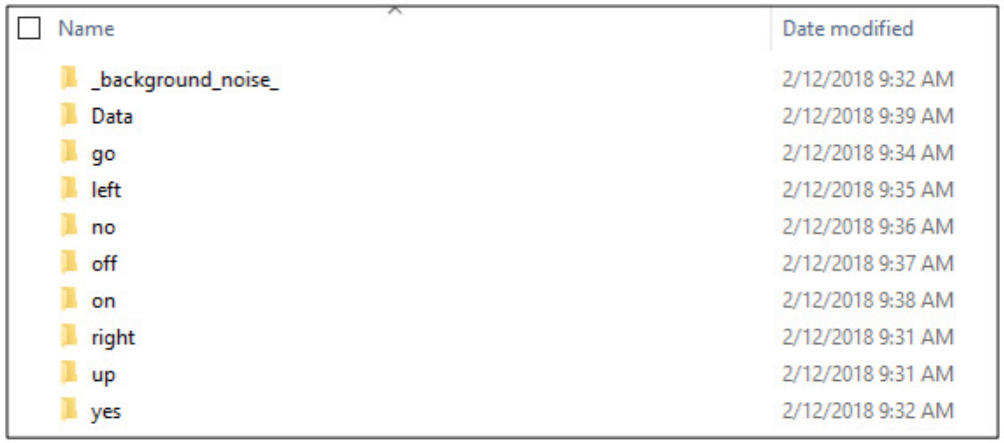
- LSTM = Long Short Term Memory (Networks)
  - Signal, text, time-series data
  - Use previous data to predict new information
- I live in France. I speak _____.



© Terasoft, Inc.

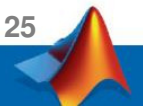
1. Create Datastore

- Datastore creates reference for data
- Do not have to load in all objects into memory

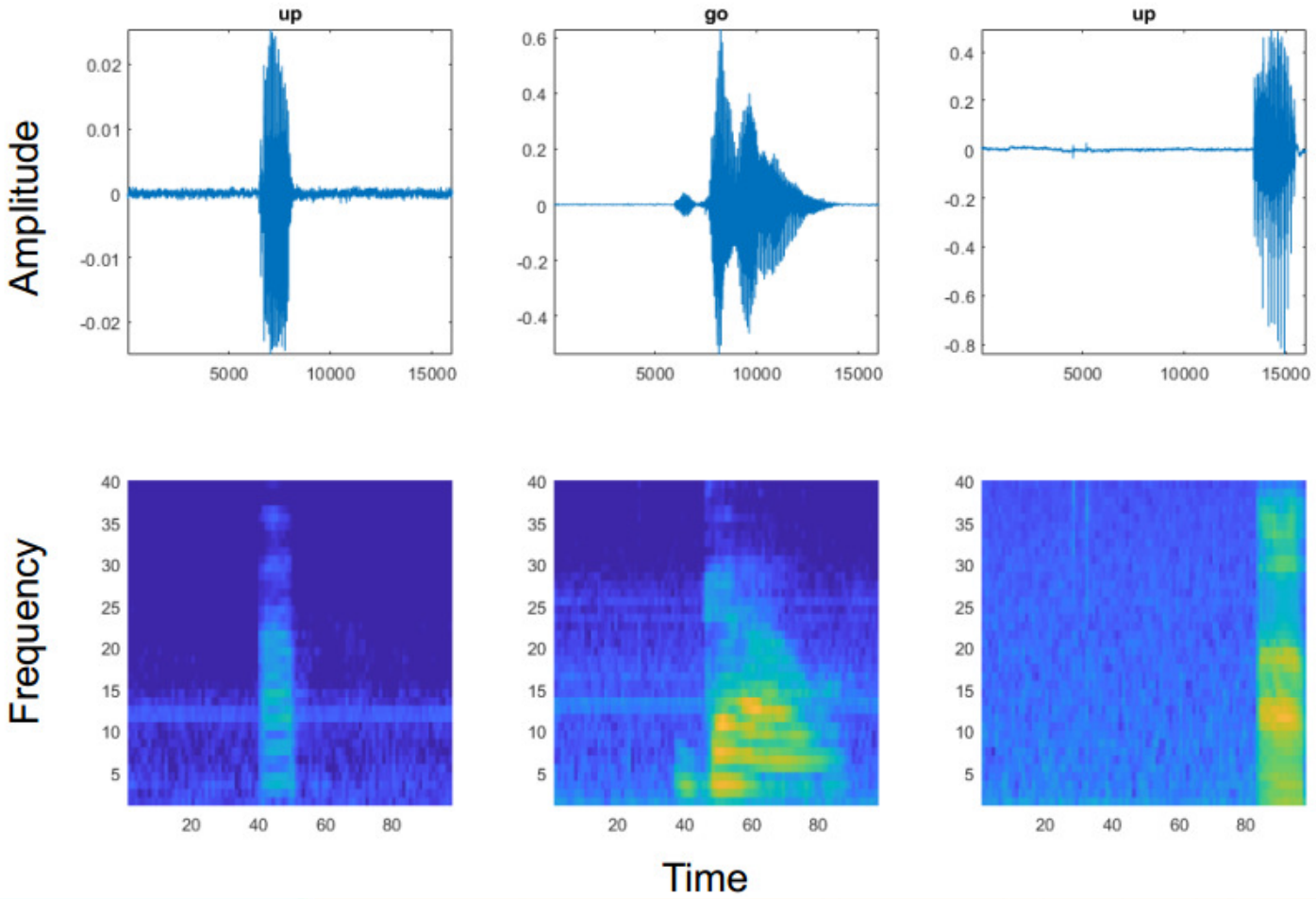



```
datafolder = fullfile(tempdir,'speech_commands_v0.01');

addpath(fullfile(matlabroot,'toolbox','audio','audiodemos'))
ads = audioexample.Datastore(datafolder, ...
    'IncludeSubfolders',true, ...
    'FileExtensions','.wav', ...
    'LabelSource','foldernames', ...
    'ReadMethod','File')
```


© Terasoft, Inc.



2. Compute Speech Spectrograms

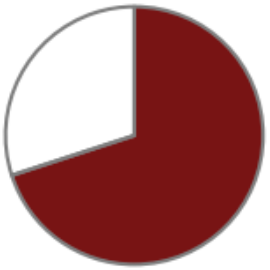


© Terasoft, Inc.

3. Split datastores

Training

70%



- Trains the model
- Computer “learns” from this data

Validation

15%

- Checks accuracy of model during training

Test

15%

- Tests model accuracy
- Not used until validation accuracy is good

© Terasoft, Inc.

4. Define Architecture and Parameters

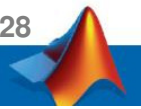
```
layers = [
    imageInputLayer(imageSize)

    convolution2dLayer(3,16,'Padding','same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(2,'Stride',2)

    convolution2dLayer(3,32,'Padding','same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(2,'Stride',2,'Padding',[0,1])

    dropoutLayer(dropoutProb)
    convolution2dLayer(3,64,'Padding','same')
    batchNormalizationLayer
    reluLayer

    dropoutLayer(dropoutProb)
    convolution2dLayer(3,64,'Padding','same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(2,'Stride',2,'Padding',[0,1])

    dropoutLayer(dropoutProb)
    convolution2dLayer(3,64,'Padding','same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer([1 13])

    fullyConnectedLayer(numClasses)
    softmaxLayer
    weightedCrossEntropyLayer(classNames,classWeights)];
```

Neural Network Architecture

```
miniBatchSize = 128;
validationFrequency = floor(numel(YTrain)/miniBatchSize);
options = trainingOptions('adam', ...
    'InitialLearnRate',5e-4, ...
    'MaxEpochs',25, ...
    'MiniBatchSize',miniBatchSize, ...
    'Shuffle','every-epoch', ...
    'Plots','training-progress', ...
    'Verbose',false, ...
    'ValidationData',{XValidation,YValidation}, ...
    'ValidationFrequency',validationFrequency, ...
    'ValidationPatience',Inf, ...
    'LearnRateSchedule','piecewise', ...
    'LearnRateDropFactor',0.1, ...
    'LearnRateDropPeriod',20);
```

Training Parameters

© Terasoft, Inc.



4. How to choose the right structure to start from?

Pretrained Models for Transfer learning

- Access pretrained models from top researchers with a single line of code

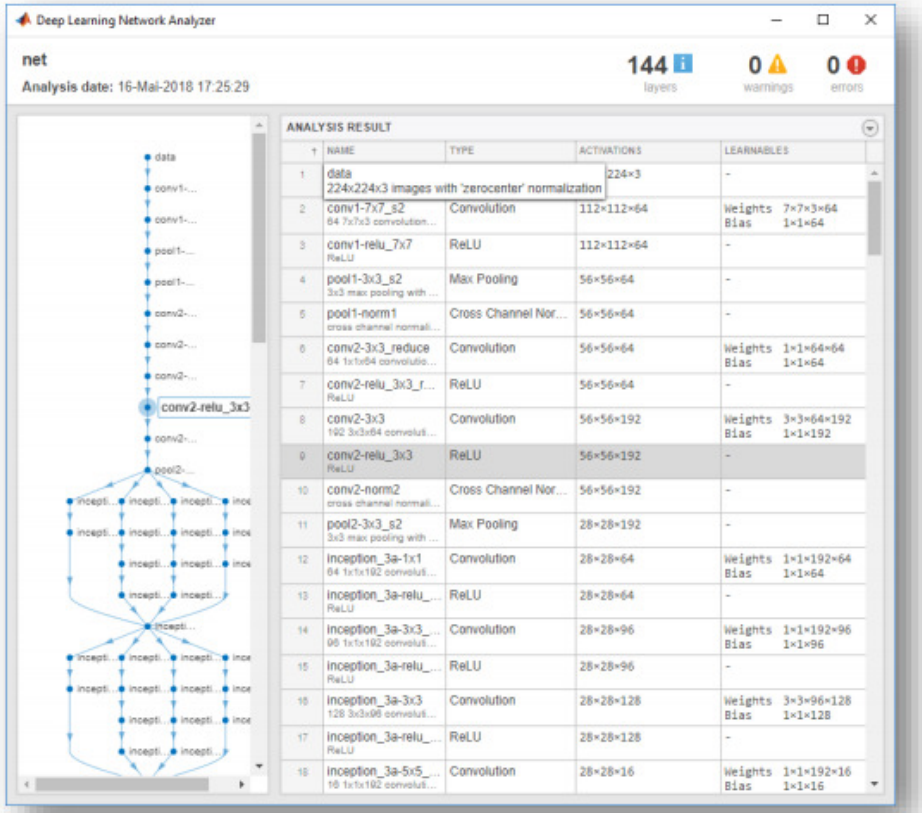
```
net = alexnet;
net = vgg16;
net = inceptionv3;
net = googlenet;
...
```

Import Models from OSS Frameworks

- Caffe Model Importer
- TensorFlow/Keras Model Importer

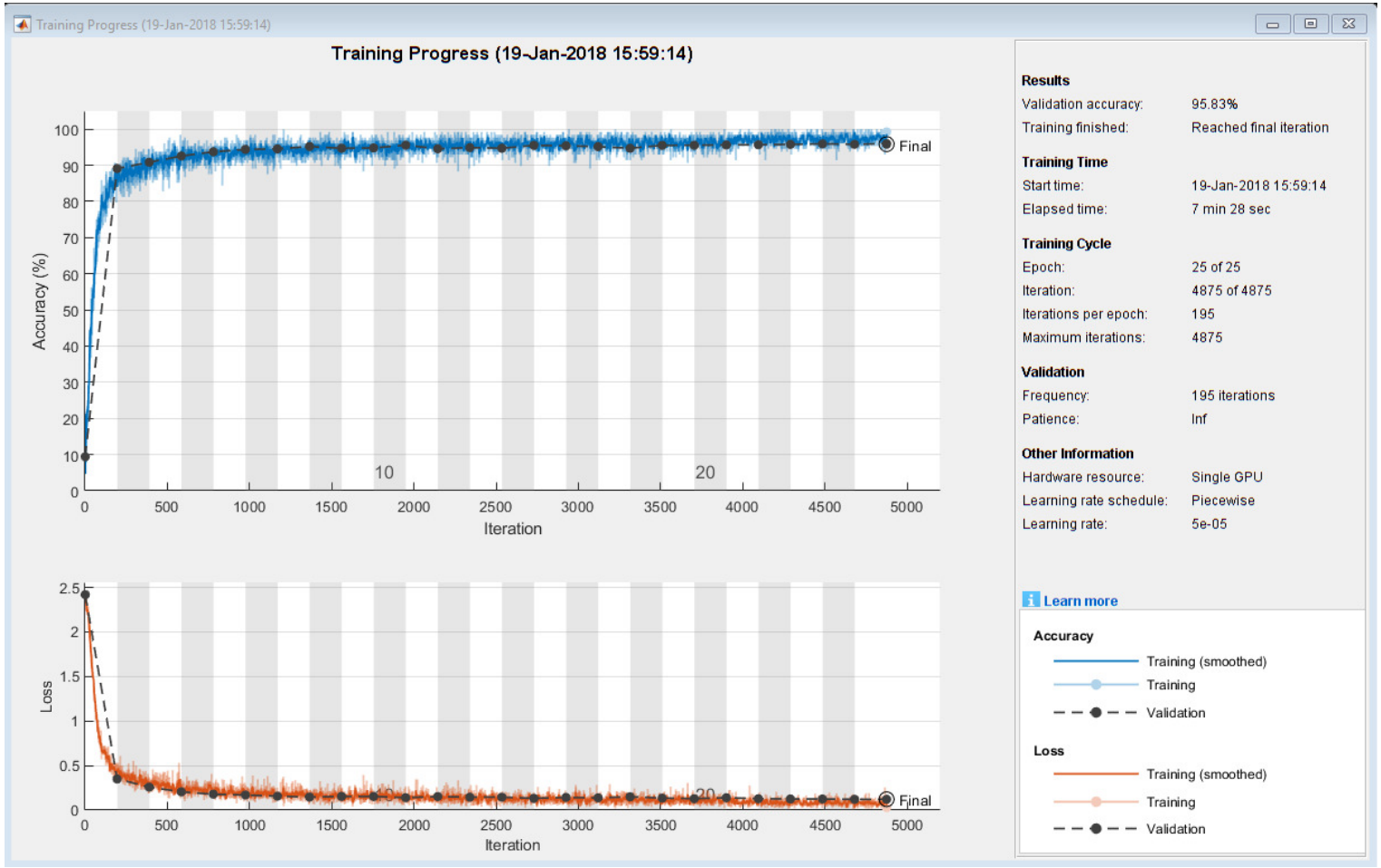
Start from scratch

- Verify structure with Network Analyzer App



© Terasoft, Inc.

5. Train Network



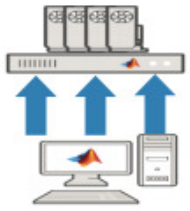
© Terasoft, Inc.

Deep Learning on CPU, GPU, Multi-GPU and Clusters

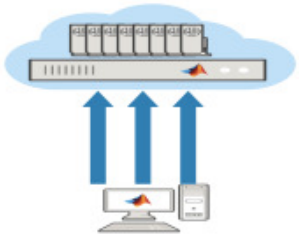
Single CPU

Single CPU
Single GPU

Single CPU, Multiple GPUs



On-prem server with GPUs



Cloud GPUs (AWS)

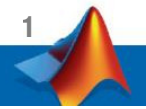
How to target?

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

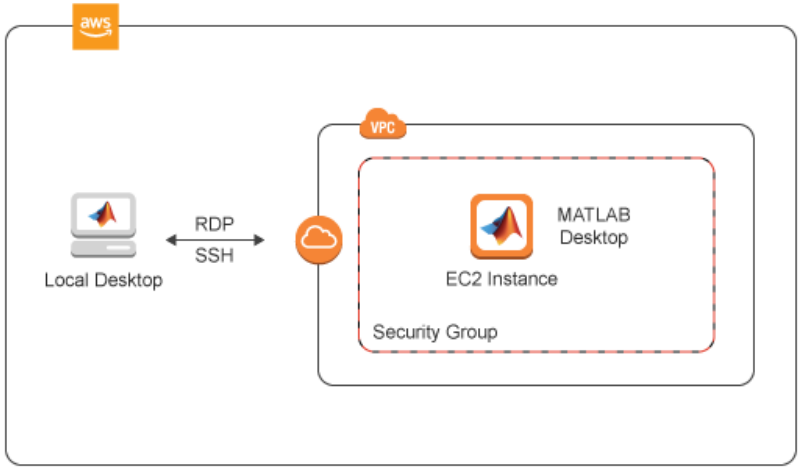
```
opts = trainingOptions('sgdm', ...
    'MaxEpochs', 100, ...
    'MiniBatchSize', 250, ...
    'InitialLearnRate', 0.00005, ...
'ExecutionEnvironment', 'multi-gpu' );
```

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

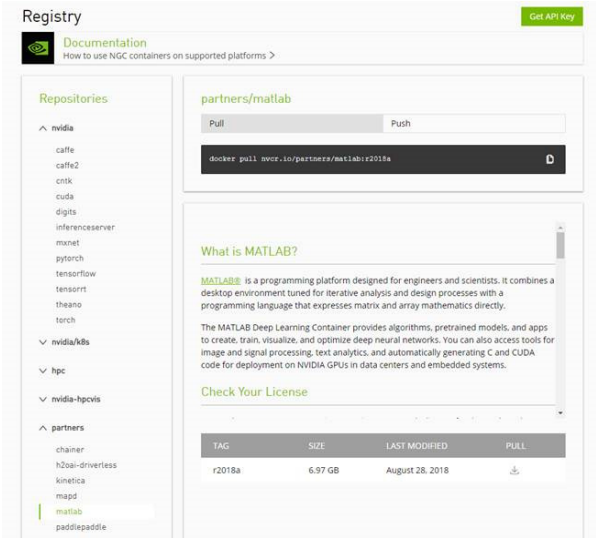

© Terasoft, Inc.



R2018b: Easier to Get Started on Cloud and HPC



Reference Architectures



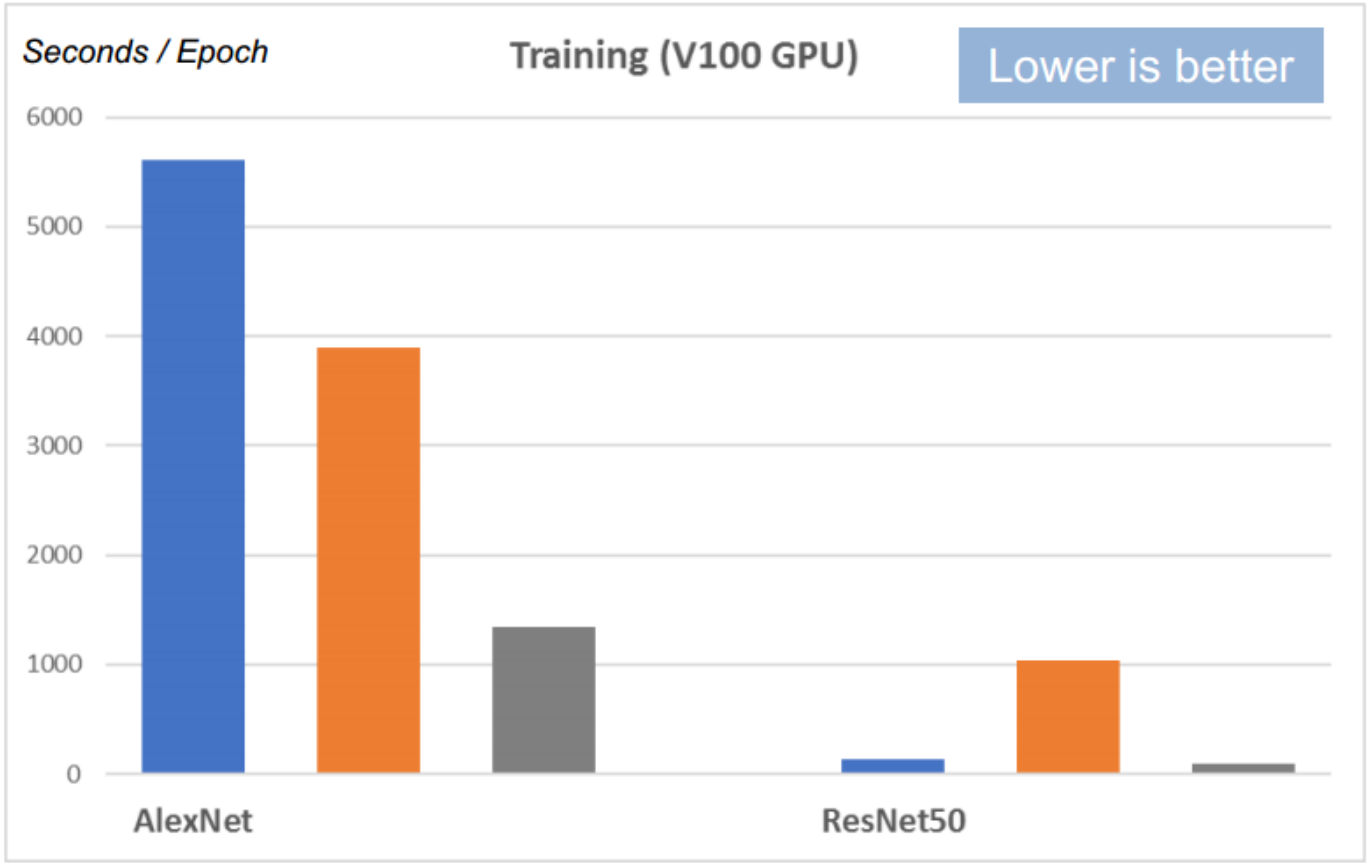
Containers for Nvidia GPU Cloud (NGC).
Also supports NVIDIA DGX

Why is this important?

1. Customers want to get started in cloud instead of buying hardware (Reference Architectures)
2. Customers want to use NVIDIA DGX super computer (NGC, DGX) to accelerate training

© Terasoft, Inc.

Training Performance



TensorFlow

MATLAB

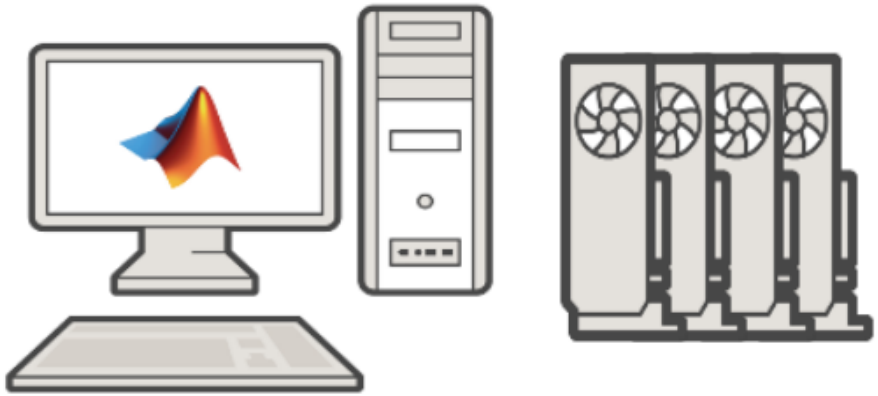
MXNet

Batch size 32

© Terasoft, Inc.

MATLAB is Fast for Deployment

- Target a GPU for optimal performance
- NVIDIA GPUs use CUDA code
- We only have MATLAB code. Can we translate this?



© Terasoft, Inc.

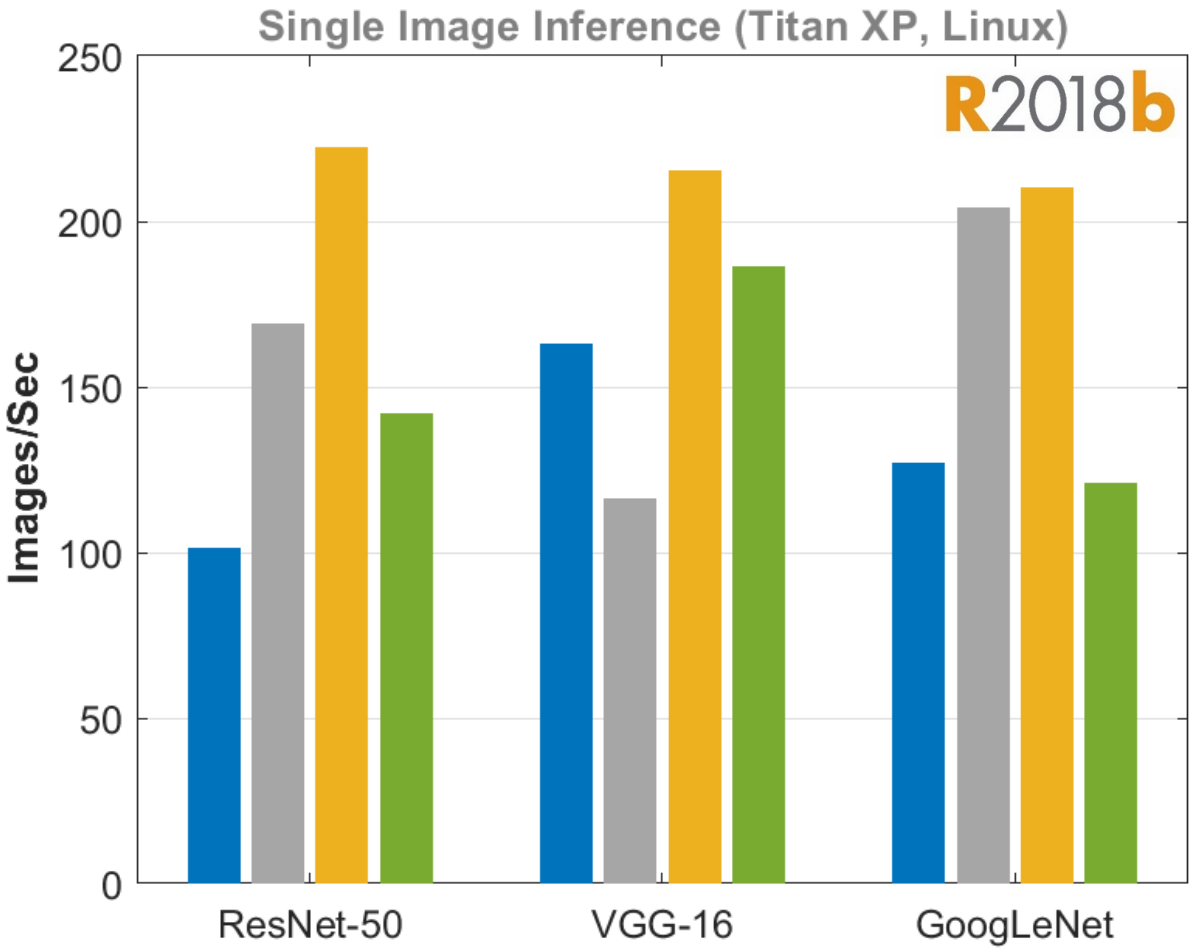
GPU Coder

- Automatically generates **CUDA** Code from MATLAB Code
  - can be used on NVIDIA GPUs

- CUDA extends C/C++ code with constructs for parallel computing

© Terasoft, Inc.

With GPU Coder, MATLAB is fast



Faster than TensorFlow, MXNet, and PyTorch

Intel® Xeon® CPU 3.6 GHz - NVIDIA libraries: CUDA9 - cuDNN 7 - Frameworks: TensorFlow 1.8.0, MXNet 1.2.1, PyTorch 0.3.1

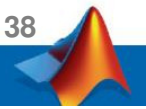
© Terasoft, Inc.

GPU Coder Performance

© Terasoft, Inc.

Why MATLAB?

- MATLAB is Productive
- MATLAB is Fast
- MATLAB Integrates with Open Source



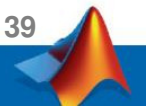
© Terasoft, Inc.

Used MATLAB and Open Source Together

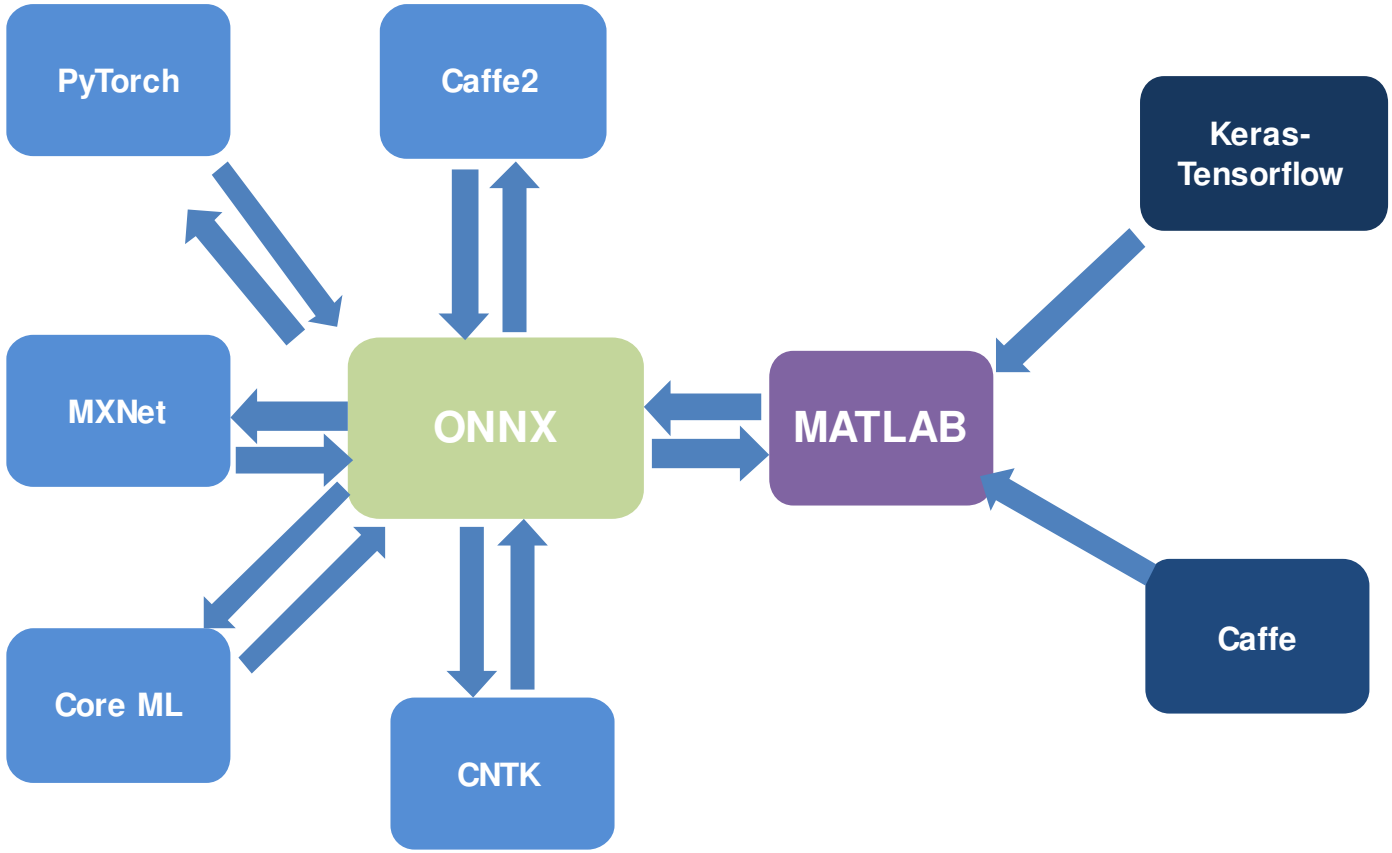
1. Deep Joint Rain Detection and Removal from a Single Image" Wenhan Yang, Robby T. Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan

- Used Caffe and MATLAB together
- Achieved significantly better results than an engineered rain model.
- Use our tools where it makes your workflow easier!

© Terasoft, Inc.



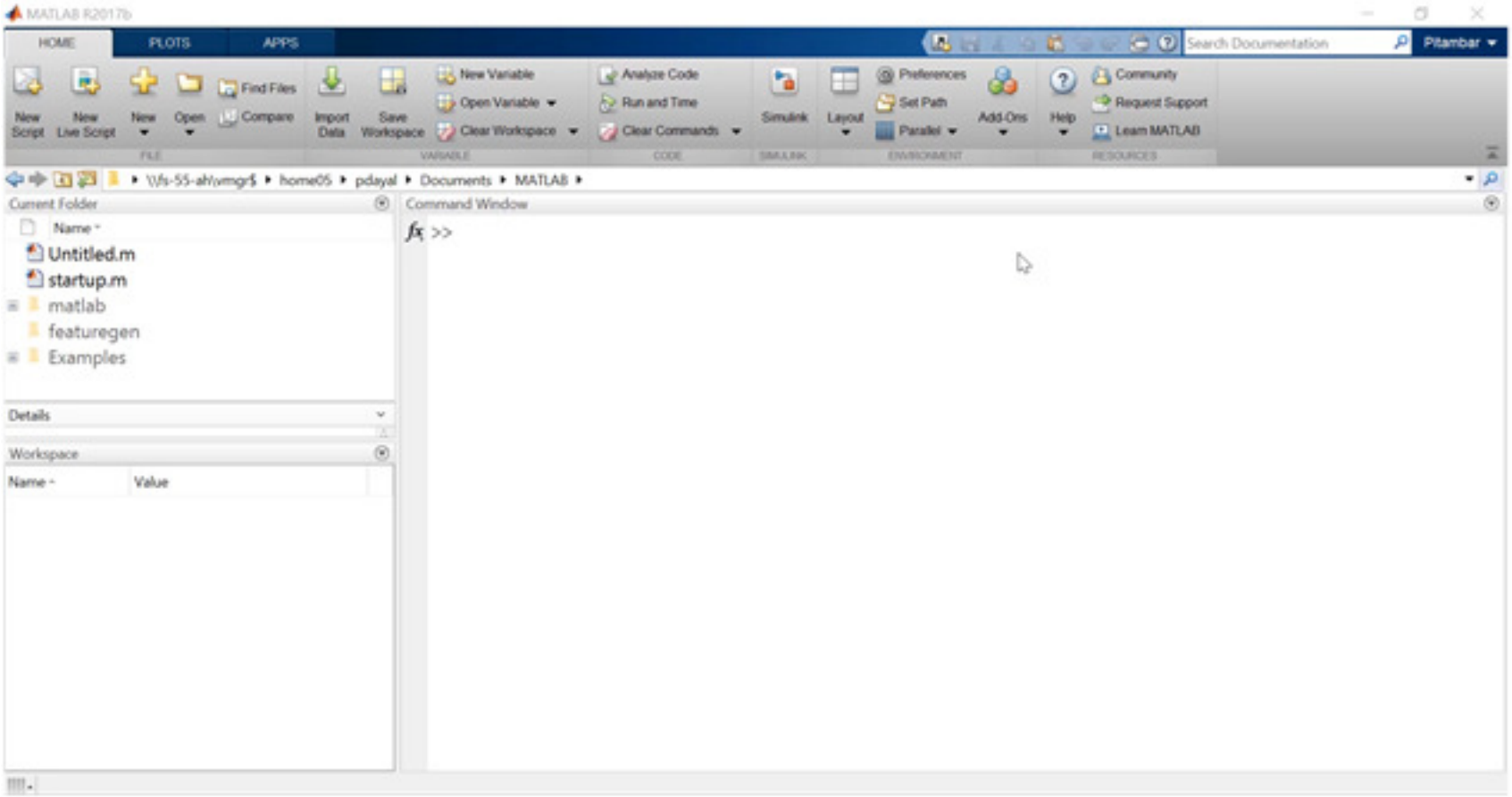
MATLAB Integrates with Open Source Frameworks



- ONNX: Input and export to several frameworks
- Model importers
  - Caffe
  - TensorFlow-Keras

© Terasoft, Inc.

Keras-Tensorflow Importer

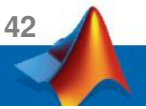


© Terasoft, Inc.

MATLAB Integrates with Open Source Frameworks

- MATLAB supports entire deep learning workflow
  - Use when it is convenient for your workflow
- Access to latest models
- Improved collaboration with other users

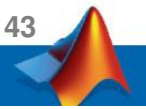
© Terasoft, Inc.



MATLAB Integrates with Open Source Frameworks

- MATLAB is Productive
- MATLAB is Fast (*Performance*)
- MATLAB Integrates with Open Source *(Frameworks)*

© Terasoft, Inc.



© Terasoft, Inc.