

# MATLAB EXPO 2018

## 揭秘深度学习

Bill Chou



The background of the slide is a close-up photograph of a wooden surface, likely a table or countertop, showing natural wood grain patterns and knots. Two overlapping rectangular boxes are centered on the image. The top box is dark gray with white text, and the bottom box is white with teal text.

# Deep Learning Demo

# Image Classification

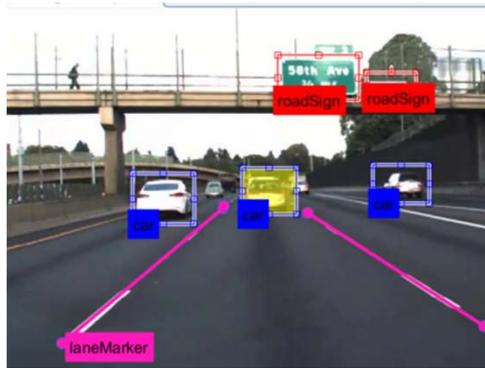
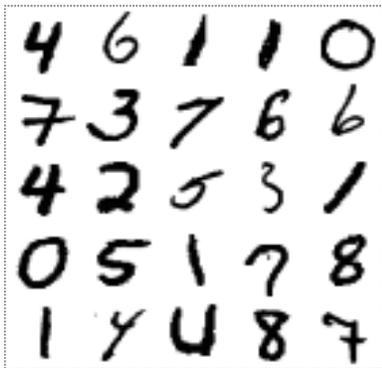
# 为什么用MATLAB要进行深度学习？

- MATLAB是高效的
- MATLAB是快速的
- MATLAB与开源软件相结合

# 深度学习应用案例

语音助理（语音到文本）  
教电子游戏人物完成游戏  
自动给黑白图像着色





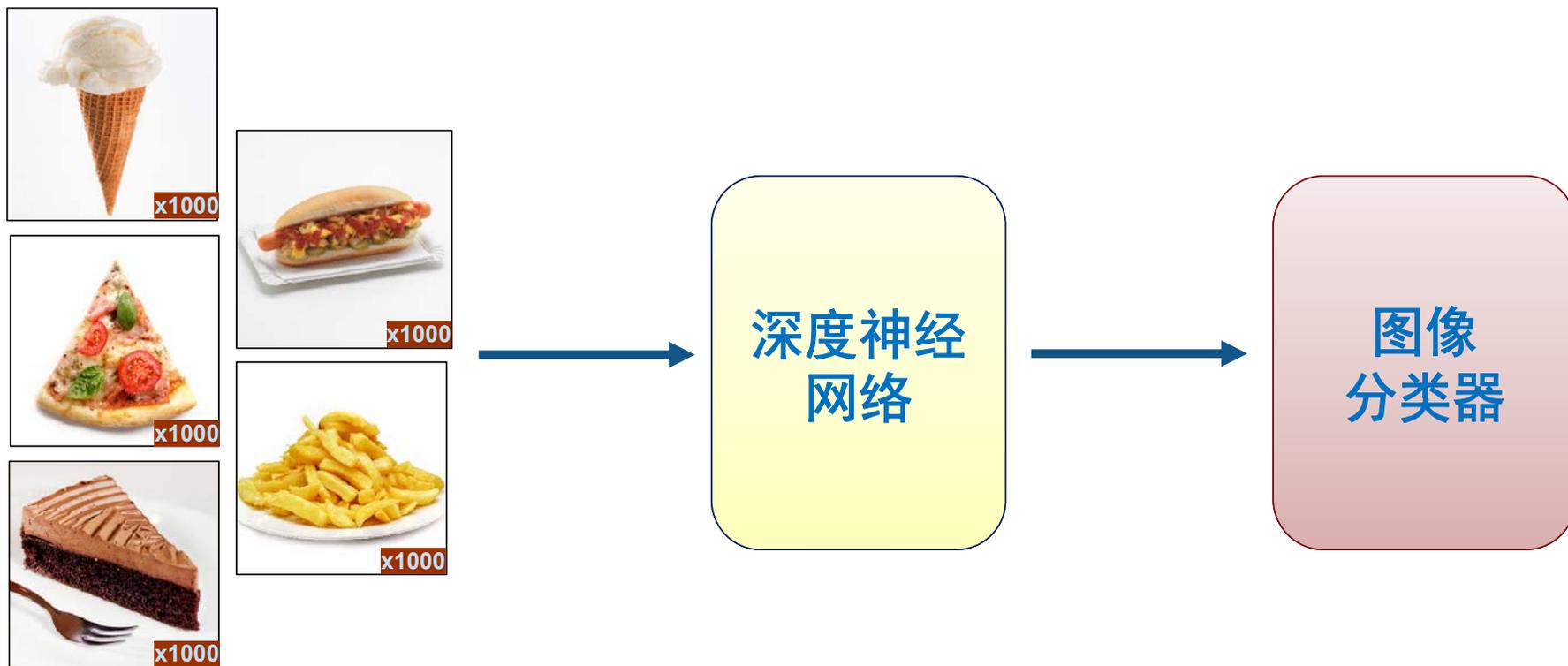
# 什么是深度学习？



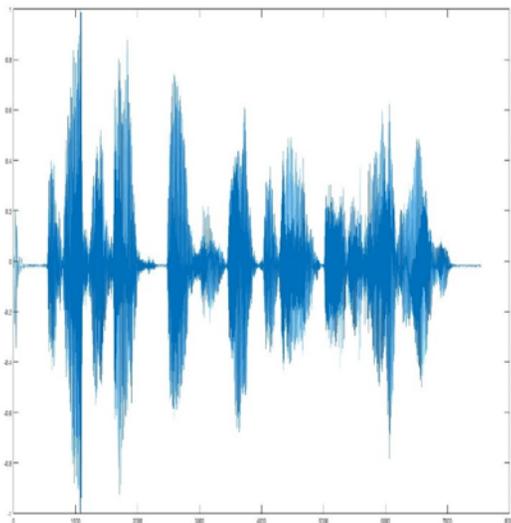
12 40.0%	0 0.0%	100% 0.0%
0 0.0%	18 60.0%	100% 0.0%
100% 0.0%	100% 0.0%	100% 0.0%

# 深度学习

深度神经网络学习直接从数据执行分类任务。



# 深度学习的数据类型



信号



文本



图像

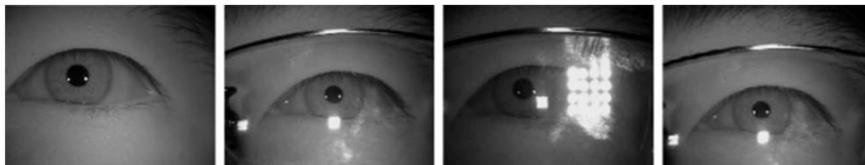
# 深度学习是多才多艺的



自动驾驶系统中的汽车和道路检测



雨水探测和清除<sup>1</sup>

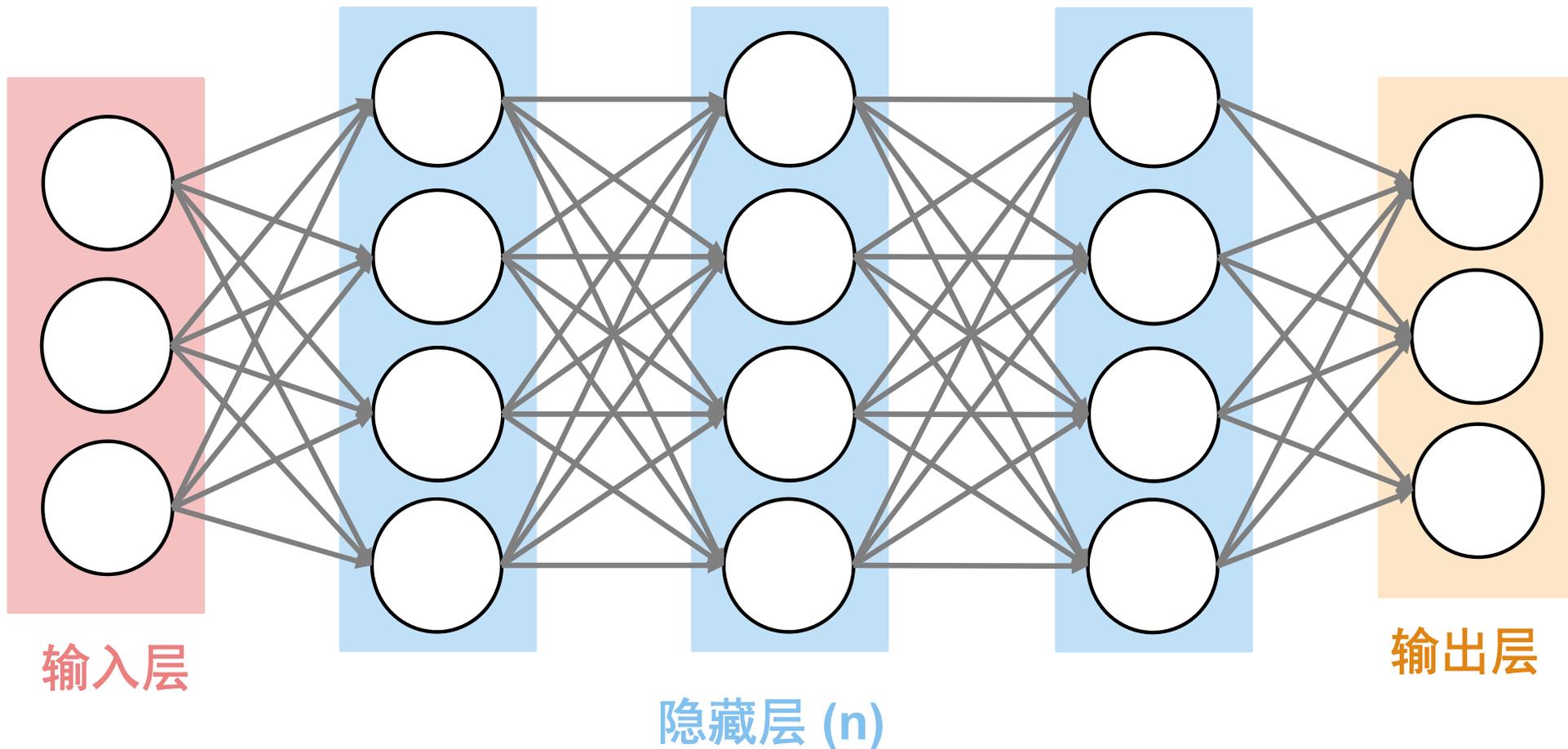


虹膜识别 - 准确率达99.4%<sup>2</sup>

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

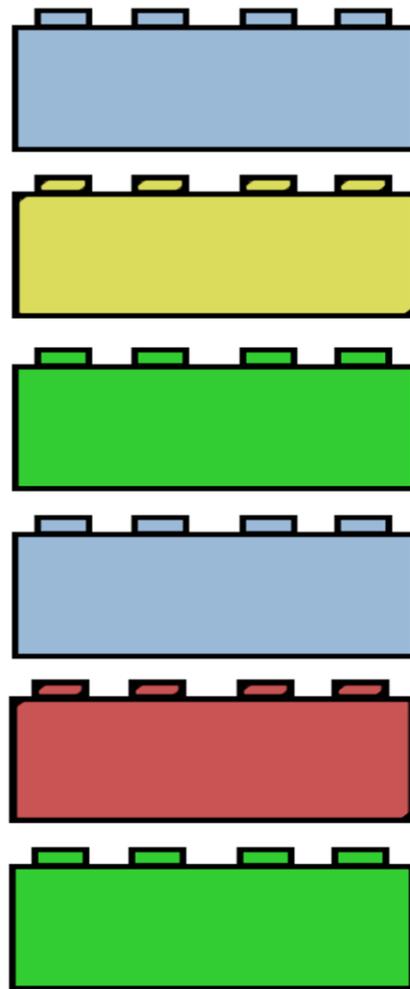
# 深度学习如何表现如此出色？

# 深度学习使用神经网络架构



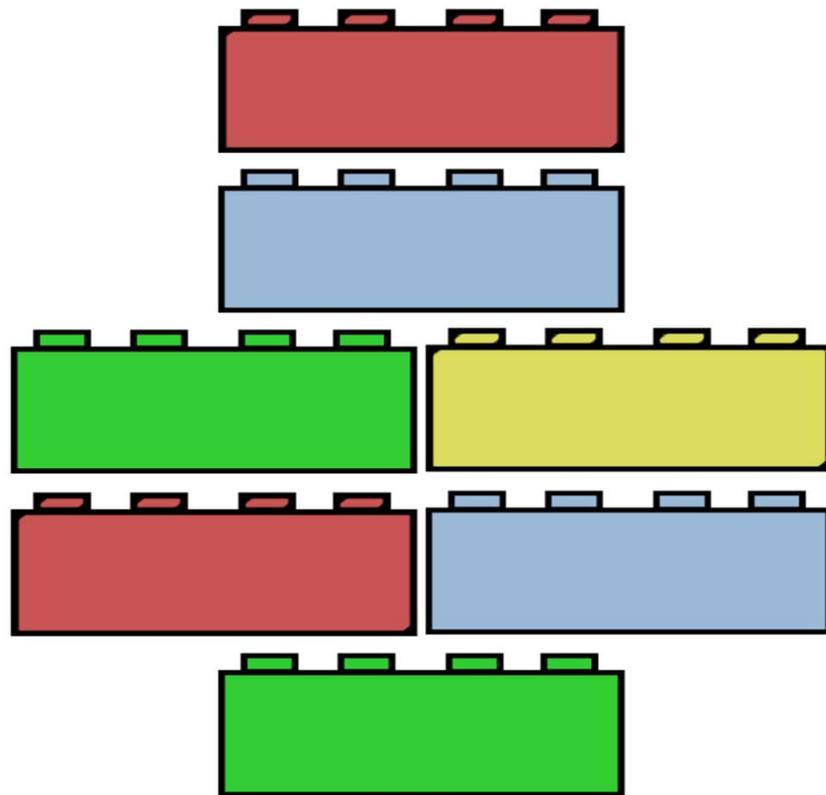
# 神经网络的层次

- 层次就像积木
  - 能堆叠起来
  - 能替换层次
- 每个隐藏层处理上一层的信息

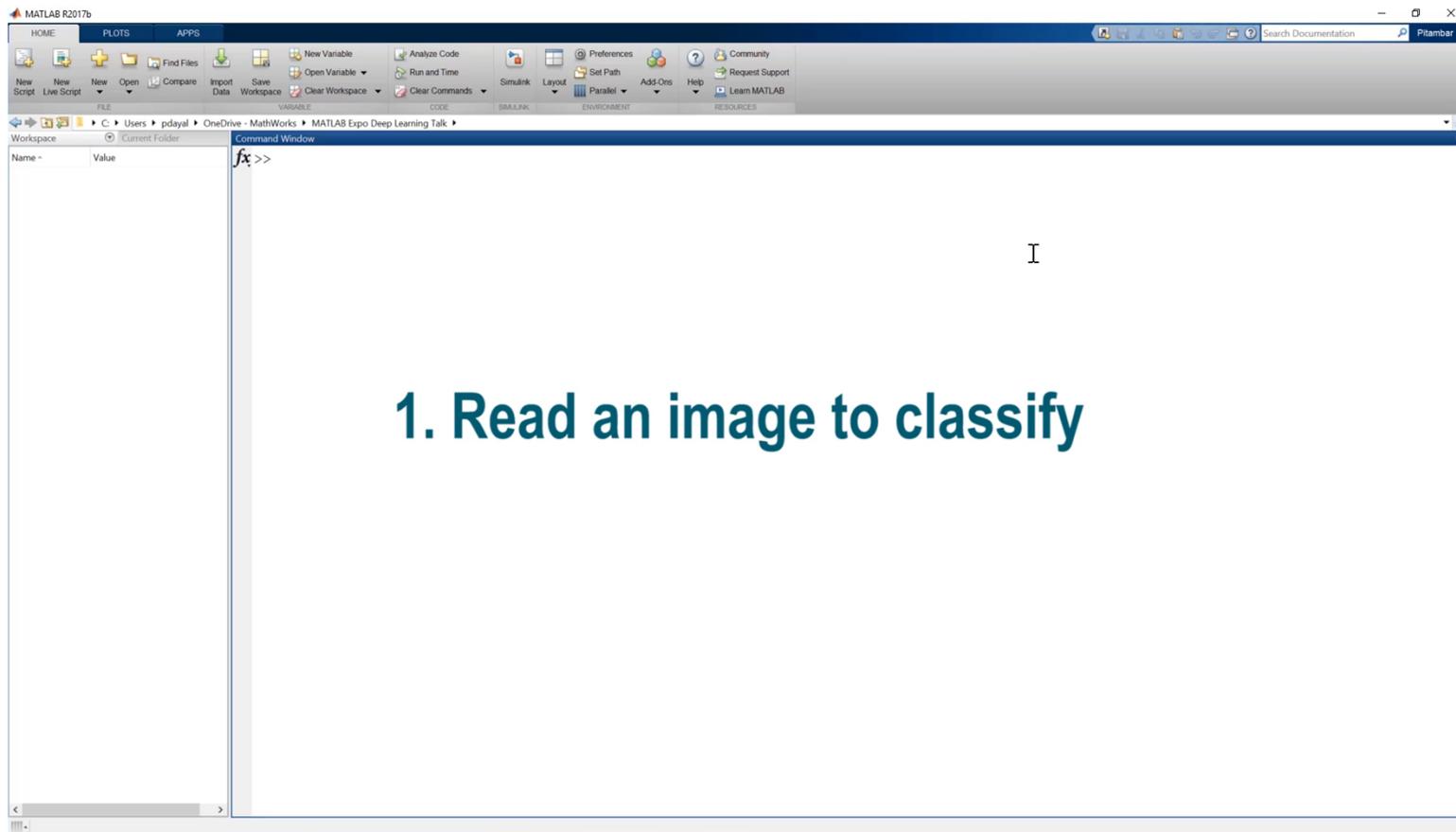


# 神经网络的层次

- 层次就像积木
  - 能堆叠起来
  - 能替换层次
- 每个隐藏层处理上一层的信息
- 可以用不同的方式排列层次



# 用6行MATLAB代码进行深入学习



# 为什么用MATLAB要进行深度学习？

- **MATLAB是高效的**
- MATLAB是快速的
- MATLAB与开源软件相结合

# “我热爱标记和预处理我的数据”

~ 没有工程师会这样说

# 卡特彼勒案例分析



- 世界领先的建筑和采矿设备制造商
- 这些项目之间的相似性？
  - 自主运输卡车
  - 行人检测
  - 设备分类
  - 地形映射

# 计算机必须从大量数据中学习

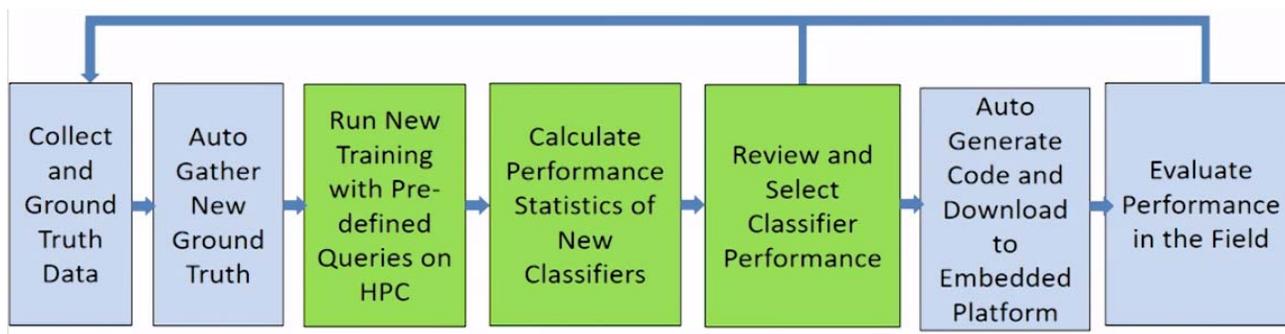
- 当创建自治系统,我们必须首先标记所有数据



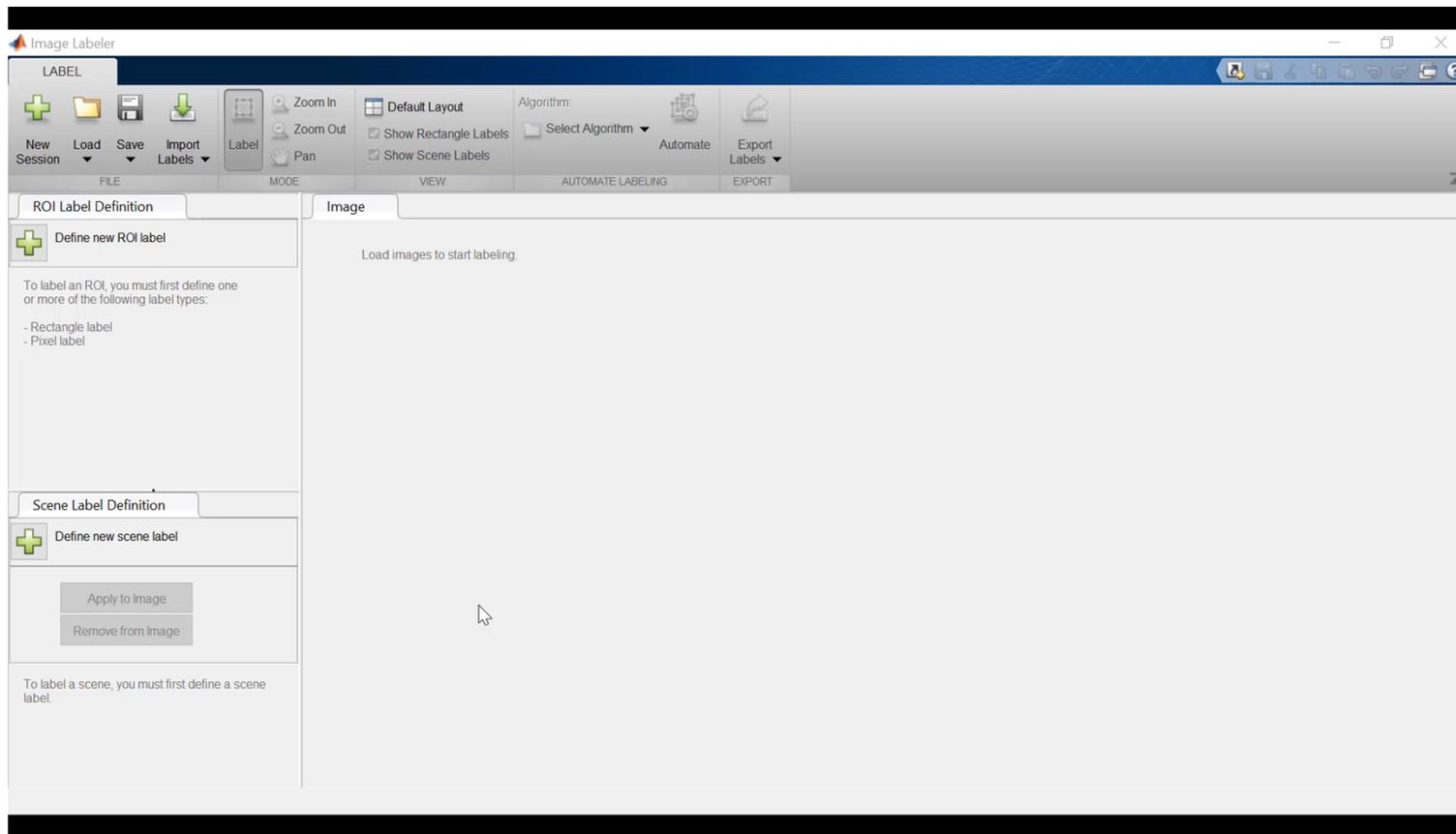
“我们之前花费太多时间来ground-truth数据”  
--Larry Mianzo, 卡特彼勒

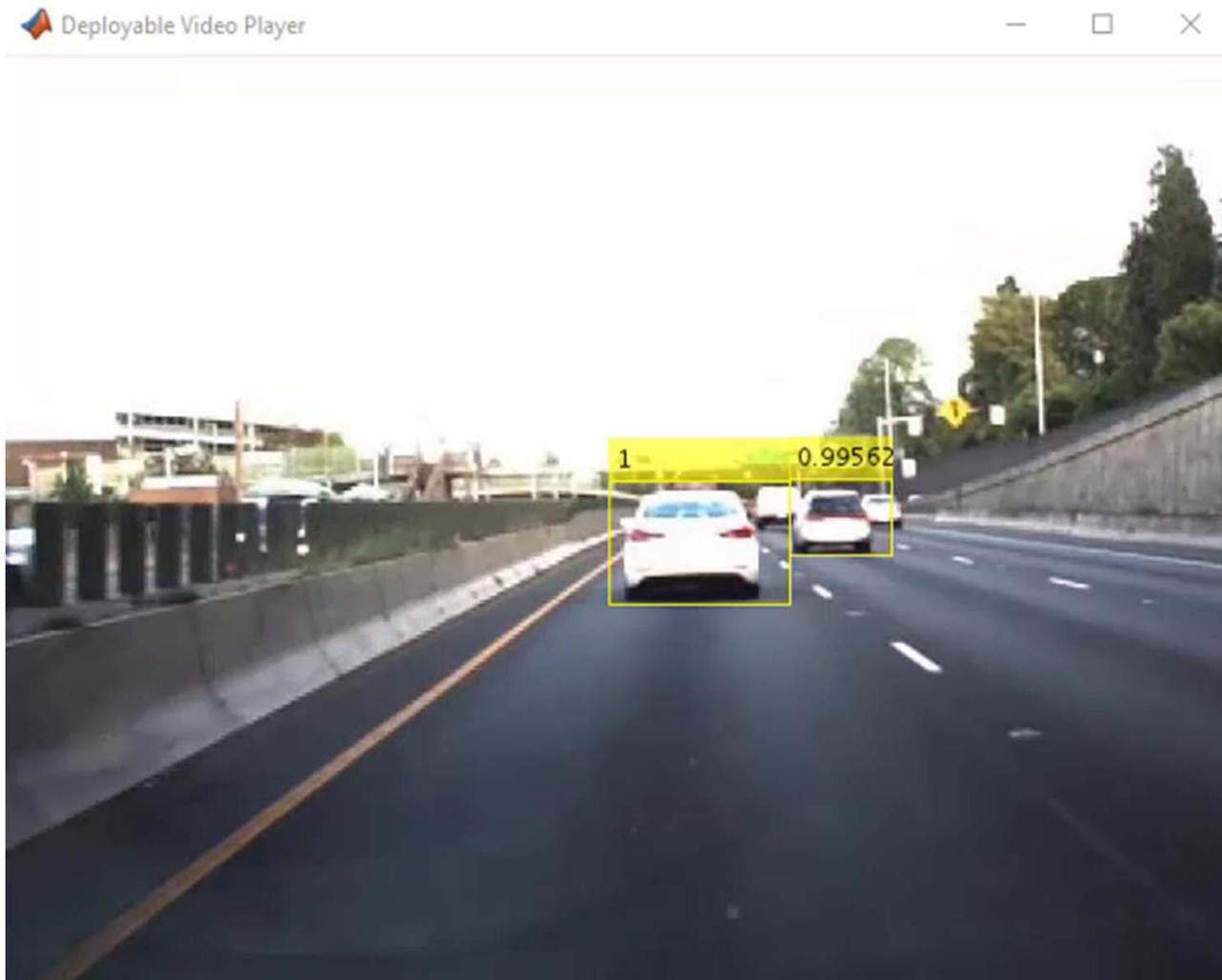
# 卡特彼勒如何使用我们的工具？

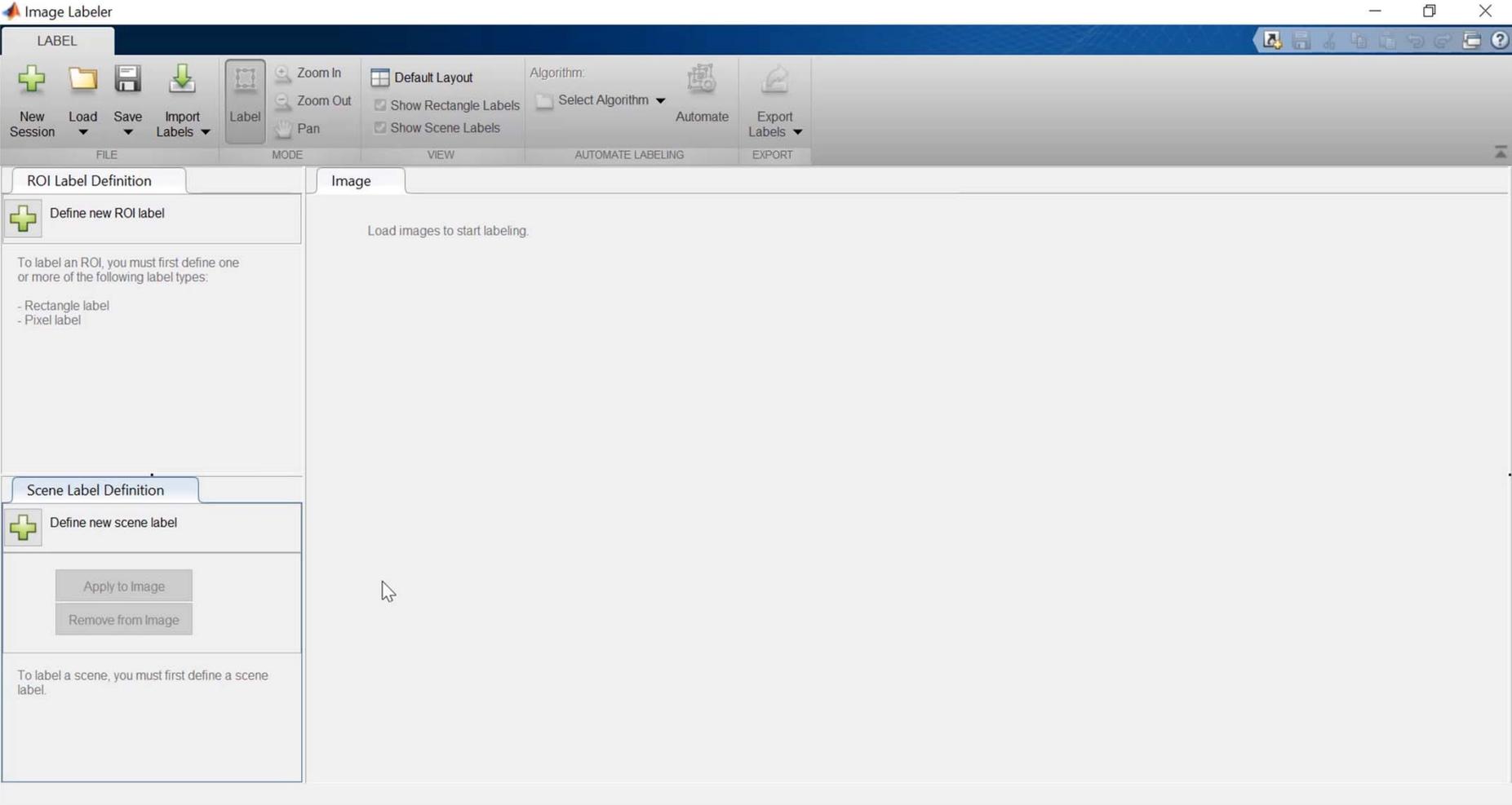
- 半自动贴标过程
  - “我们从必须标记100%的数据到必须标记大约80%到90%的数据”
- 使用MATLAB在整个开发流程
  - “因为一切都在MATLAB中，所以开发时间很短”



# 如何使用MATLAB贴标？









# MATLAB是高效的

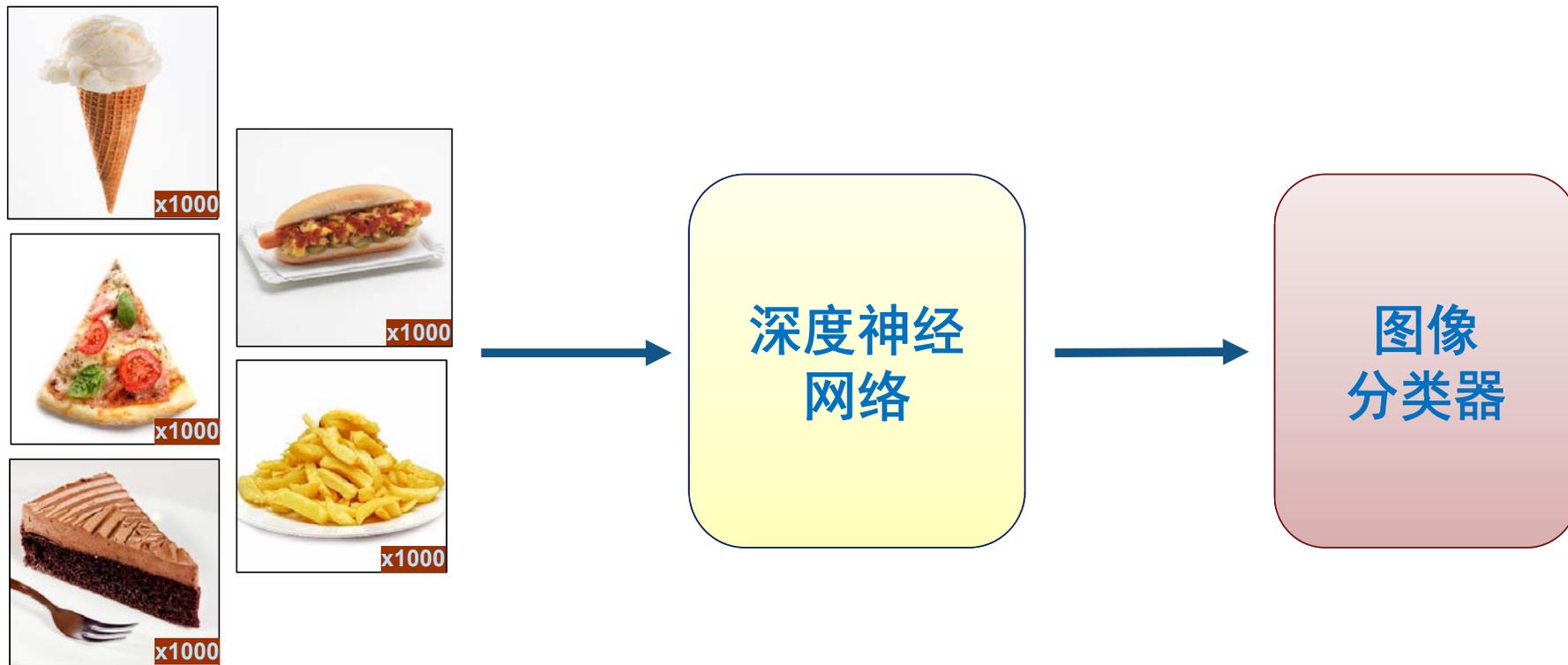
- Image Labeler App 半自动化标签流程
- Bootstrapping
  - 当您正确标记更多图像时，它会更新算法来改进自动标记
- 即使手动标记也易于加载元数据

# MATLAB是快速的



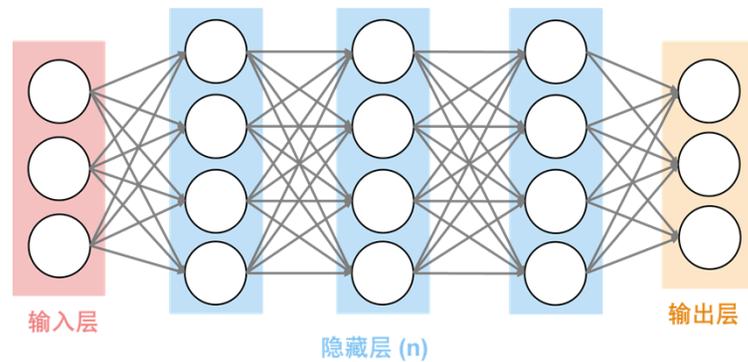
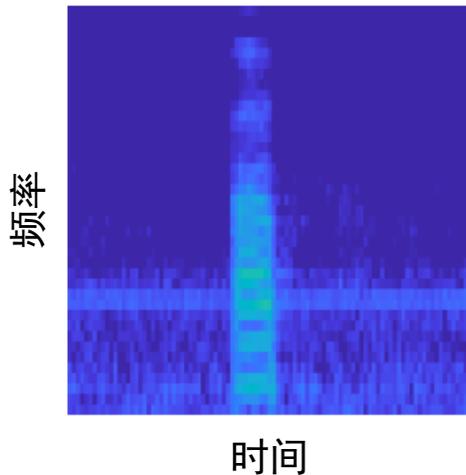
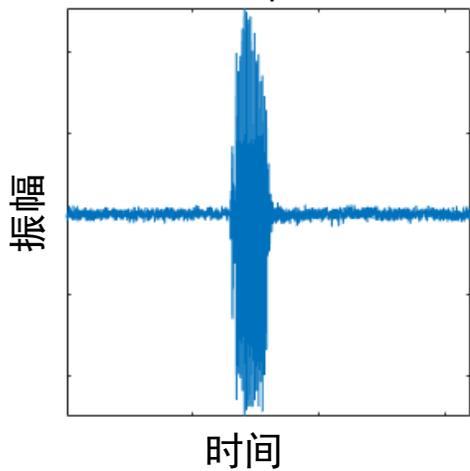
# 深度学习训练是什麼？

将标记的数据输入神经网络



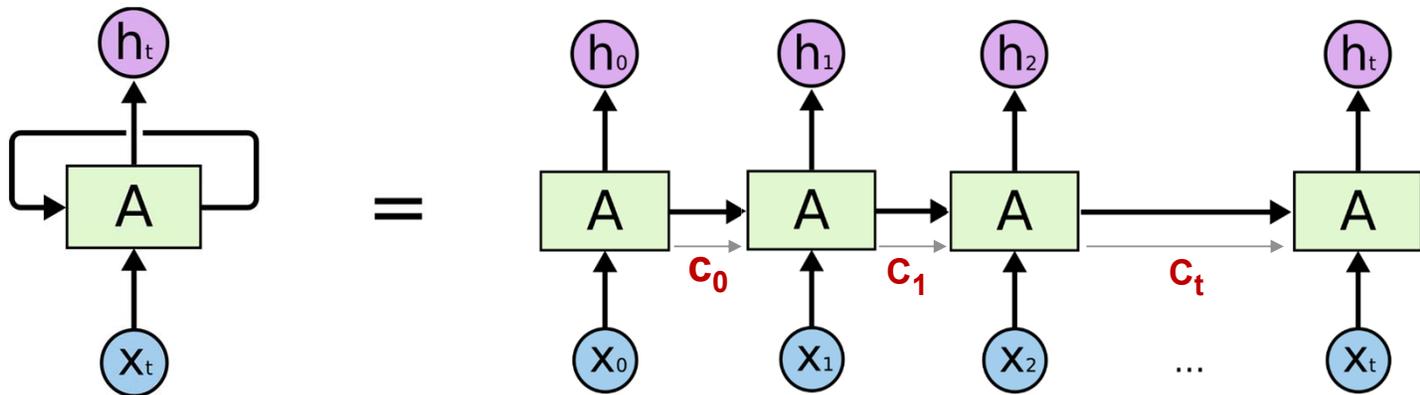
# 语音识别示例

音频信号 → 频谱图 → 图像分类算法



# 另一个信号神经网络 – LSTM网络

- LSTM = Long Short Term Memory (长短期记忆网络)
  - 信号, 文本, 时间序列数据
  - 使用以前的数据来预测新的信息
- 我住在法国, 我说 \_\_\_\_\_



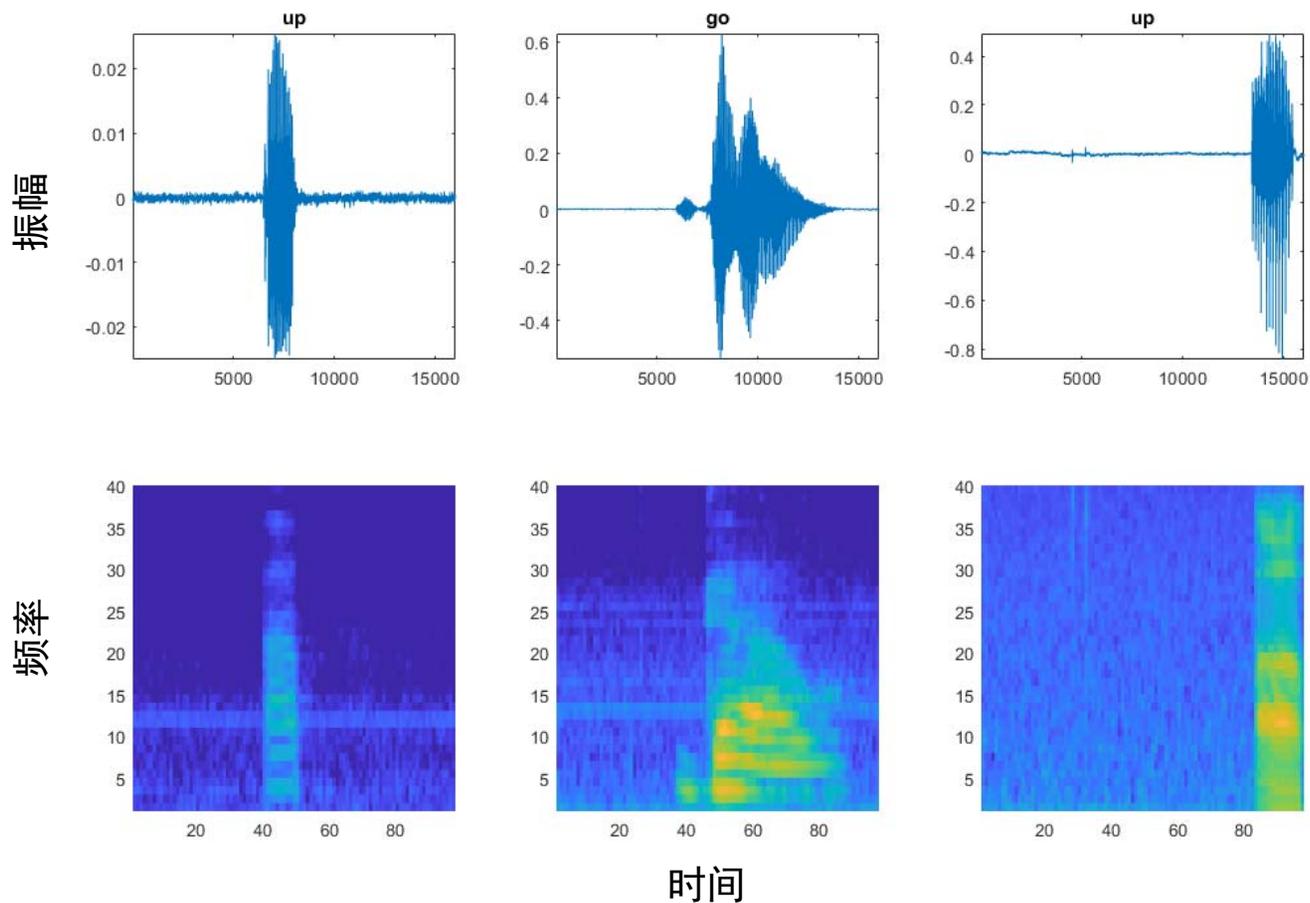
# 1. 创建数据存储

- 数据存储创建数据 reference
- 不必将所有 object 加载到内存中

<input type="checkbox"/> Name	Date modified
_background_noise_	2/12/2018 9:32 AM
Data	2/12/2018 9:39 AM
go	2/12/2018 9:34 AM
left	2/12/2018 9:35 AM
no	2/12/2018 9:36 AM
off	2/12/2018 9:37 AM
on	2/12/2018 9:38 AM
right	2/12/2018 9:31 AM
up	2/12/2018 9:31 AM
yes	2/12/2018 9:32 AM

```
datafolder = fullfile(tempdir,'speech_commands_v0.01');  
  
addpath(fullfile(matlabroot,'toolbox','audio','audiodem'))  
ads = audioexample.Datastore(datafolder, ...  
    'IncludeSubfolders',true, ...  
    'FileExtensions','.wav', ...  
    'LabelSource','foldernames', ...  
    'ReadMethod','File')
```

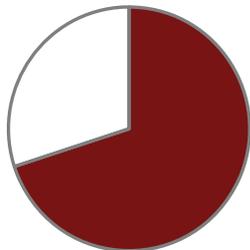
## 2. 计算语音频谱图



### 3. 拆分数据存储

#### 训练

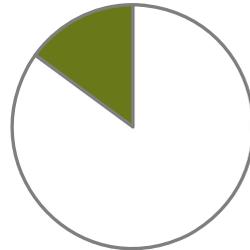
70%



- 训练神经网络
- 计算机从这些数据中“学习”

#### 验证

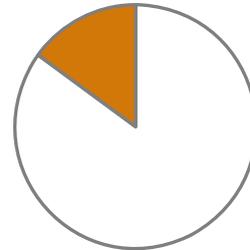
15%



- 在训练期间检查神经网络的准确性

#### 测试

15%



- 测试神经网络的准确性
- 在验证准确性达到标准之后使用

## 4. 定义网络结构和参数

```
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

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

maxPooling2dLayer([1 13])

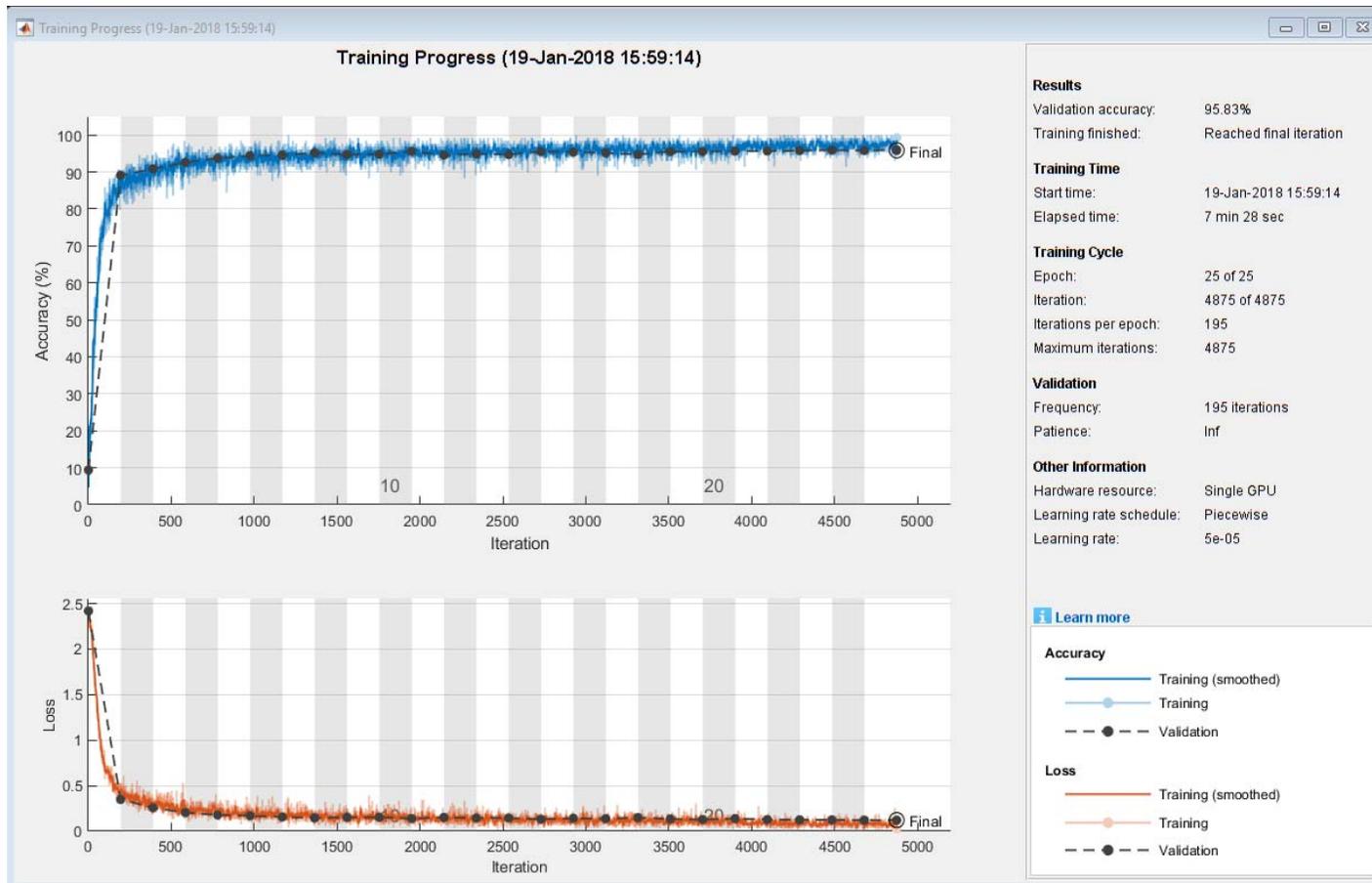
fullyConnectedLayer(numClasses)
softmaxLayer
weightedCrossEntropyLayer(classNames,classWeights)];
```

神经网络结构

```
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);
```

训练参数

# 5. 训练神经网络



# 深度学习在CPU, GPU, 多GPU, 和HPC集群

## 训练参数



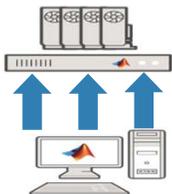
单CPU



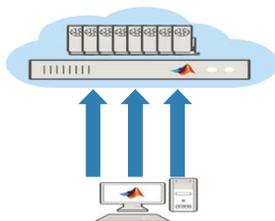
单CPU  
单GPU



单CPU, 多GPU



带GPU的内部服务器



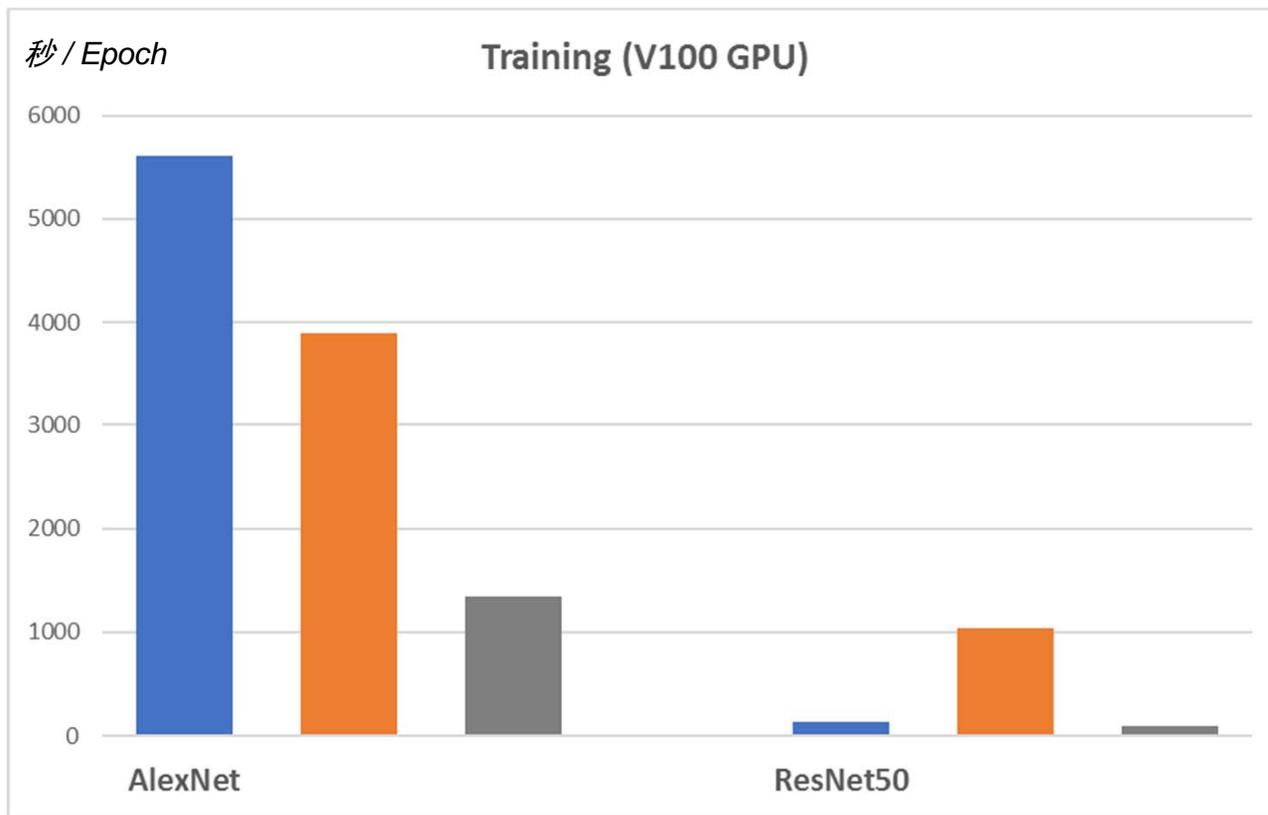
GPU云  
(AWS)

```
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' );
```

# 训练性能



TensorFlow

MATLAB

MXNet

批量 32

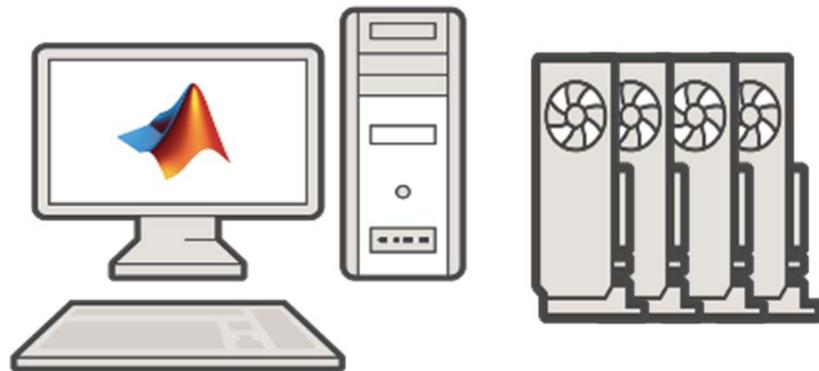
# 培训是一个迭代过程

```
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);
```

*参数根据性能进行调整*

# MATLAB部署是快速的

- 部署到GPU以获得最佳性能
- NVIDIA GPU使用CUDA代码
- 我们只有MATLAB代码。我们可以翻译成CUDA代码吗？



# GPU Coder

- 从MATLAB代码自动生成 **CUDA** 代码
  - 可在NVIDIA GPU上跑

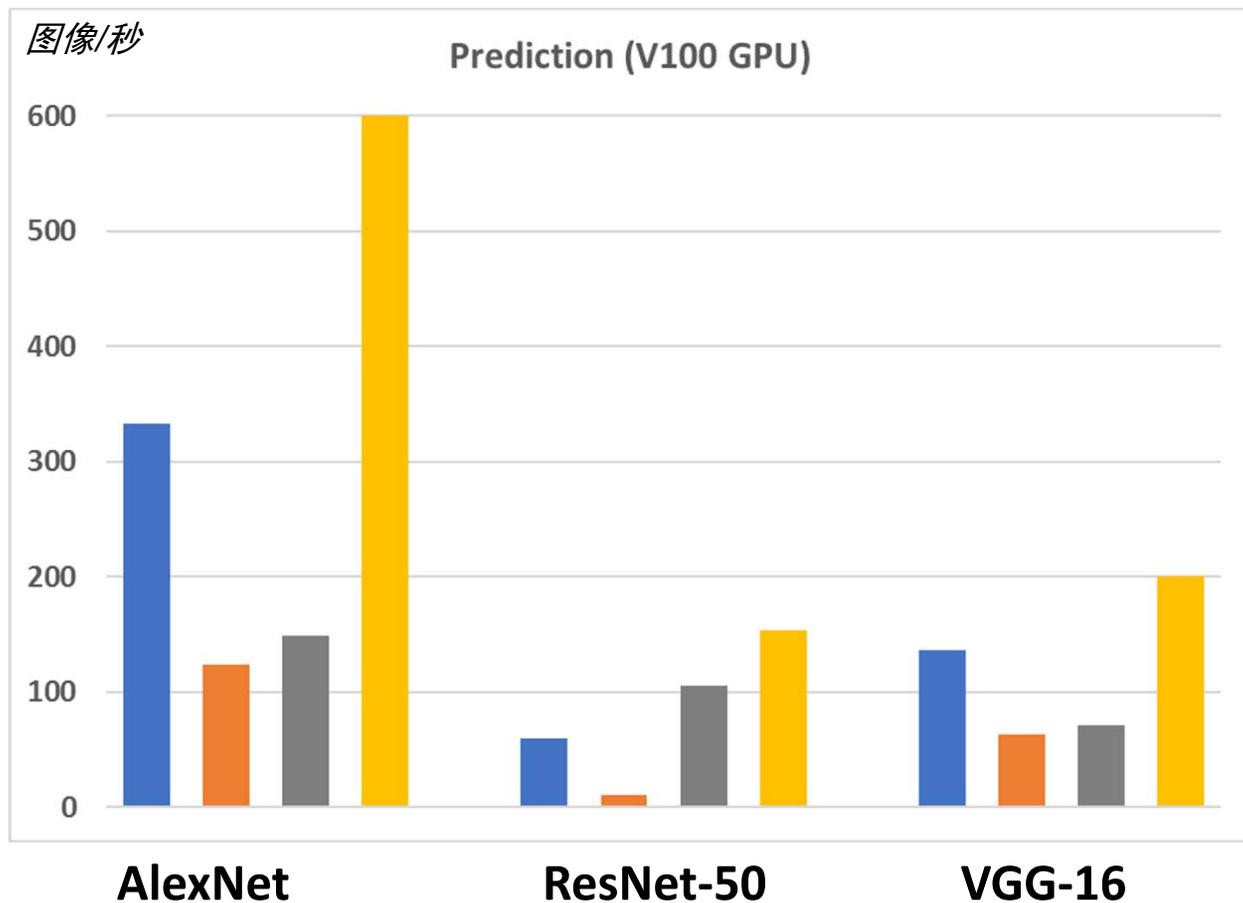


- CUDA 将 C / C ++ 代码扩展为并行计算的结构

# GPU Coder 性能

**Inference with MATLAB**

## GPU Coder深度学习推理性能强



为什么GPU Coder推理性能强？

- 分析和优化网络结构
- 代码生成科技投入了15+年时间

TensorFlow

MATLAB

MXNet

GPU Coder

# 为什么用MATLAB要进行深度学习？

- MATLAB是高效的
- MATLAB是快速的
- **MATLAB与开源软件相结合**

# MATLAB与开源软件一起使用



- Caffe和MATLAB一起使用
- 比专门设计的降雨模型取得更好的效果
- 使用MATLAB，让您的工作流程更轻松！

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

# MATLAB与开源软件相结合

- 可以使用 add-on 获得许多 pretrained 神经网络
- 用户想要导入最新的神经网络
- 直接从 Tensorflow 或 Caffe 导入神经网络
  - 提高与其他用户的协作

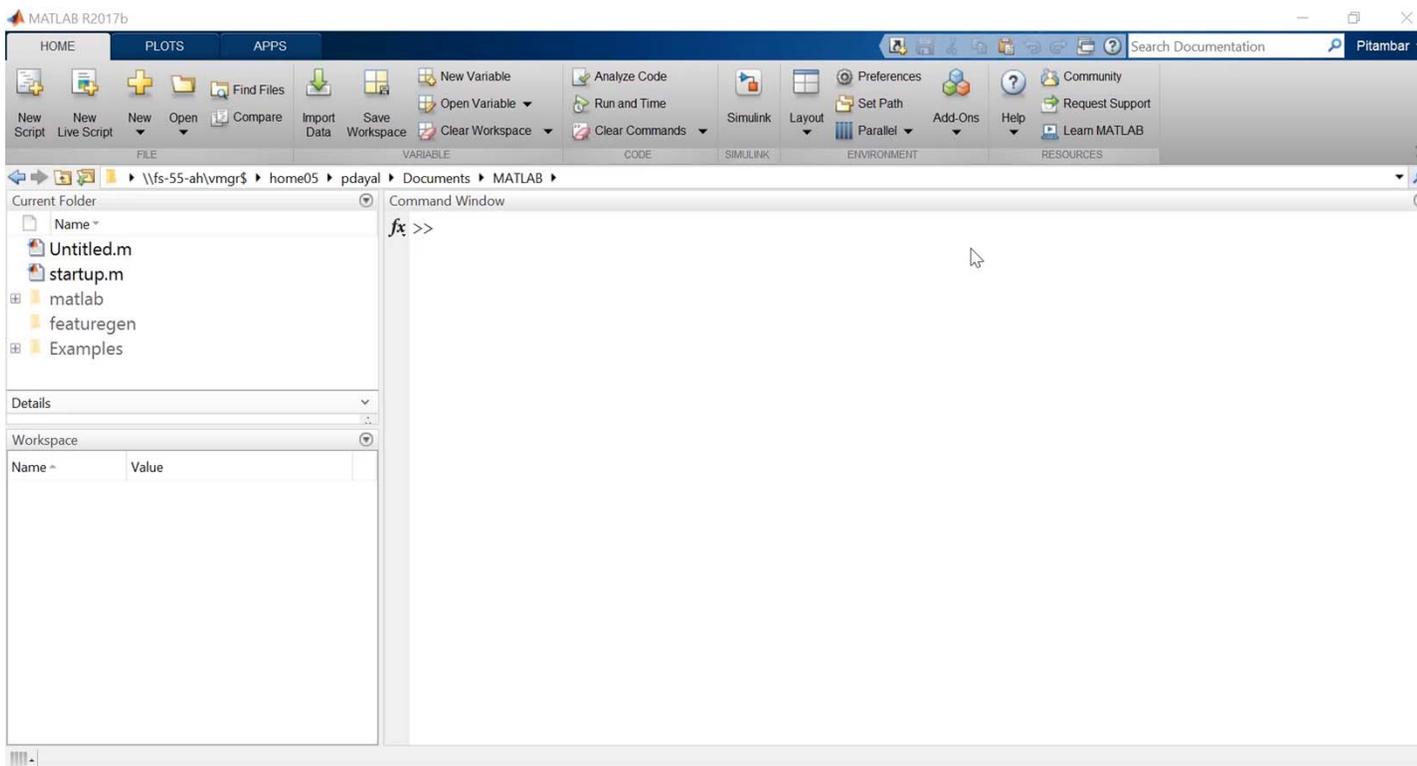
## KERAS IMPORTER

Importer for TensorFlow-Keras Models

## Caffe

MODELS

# Keras-Tensorflow 导入器



# MATLAB与开源神经架构相结合

- MATLAB支持整个深度学习工作流程
  - 在方便您的工作流程时使用
- 导入最新的神经网络
- 提高与其他用户的协作

# 为什么用MATLAB要进行深度学习？

- MATLAB是高效的
- MATLAB是快速的 (*性能强*)
- MATLAB与开源软件相结合 (*神经架构*)