딥러닝을 활용한 영상 인식 응용프로그램 개발 워크플로우

송완빈
Application Engineer
MathWorks Korea
Agenda

- Deep Learning Application for object recognition
- What is Deep Learning?
- Object Recognition using Deep Learning
  - Convolutional Neural Network
  - Regions with Convolutional Neural Network
  - Semantic Segmentation
- Key Takeaways
Smart Phone Application with Deep Learning

Users with Mobile Devices

Mobile Internet Browser

Mobile Camera
What is a scene where machine learning should be used?

**Image processing algorithm**

- Face
- Helmet

**Hand Written Program**

If `brightness` > 0.5
then ‘Helmet’
If `(edge_density < 4 and major_axis > 5)`
then ...

**Formula or Equation**

\[ Y_{Classify} = \beta_1 X_{edge} + \beta_2 Y_{bright} + \beta_3 Z_{intensity} + \ldots \]

**Machine Learning Approach**

- Learn classifiers using image data
- Flexible separation can be done if it works

**Model**

\[ \text{model} = \langle \text{Machine Learning Algorithm} \rangle (\text{images, label}) \]

- Specify condition by numerical value and carve
- When clear division is possible

MATLAB EXPO 2018
What is a scene where machine learning should be used?

**Image processing algorithm**

- Hand Written Program
  - If `brightness` > 0.5 then ‘Helmet’
  - If (`edge_density` < 4 and `major_axis` > 5) then ...

- Computer Program
  - ‘Face’
  - ‘Helmet’

**Formula or Equation**

\[ Y_{\text{Classify}} = \beta_1 X_{\text{edge}} + \beta_2 Y_{\text{bright}} + \beta_3 Z_{\text{intensity}} + \ldots \]

**Machine Learning Approach**

- Learn classifiers using image data
- Flexible separation can be done if it works

**Deep Learning**

- Specify condition by numerical value and carve
- When clear division is possible

MATLAB EXPO 2018
Machine learning vs Deep learning

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

Deep learning algorithms also **scale with data** – traditional machine learning **saturates**
MATLAB deep network in a nutshell

• A MATLAB deep network (**) is a MATLAB object that contains an array of trained layer objects.

• Layers array can be created, imported, edited, plotted in MATLAB

• Layers are trained with a lot of data and net = trainNetwork(..., layers), most of the time.

• MATLAB deep networks have different usages.

   **Classification**
   - Data=?
   - YPred = classify(net,X);
   - car x
dog ✓

   **Regression**
   - Data=?
   - YPred = predict(net,X);
   - 32

   **Semantic segmentation (**)**
   - Data=?
   - C = semanticseg(I,net);
   - sky
car
road

   **Object detection (**)**
   - Data=?
   - [bboxscores] = detect(net,I);
   - car

(!) Neural Network Toolbox
(*) Computer Vision Syst. Toolbox
Object Recognition using Deep Learning

- **Object recognition (whole image)**
  - CNN (Convolutional Neural Network)

- **Object detection and recognition**
  - R-CNN / Fast R-CNN / Faster R-CNN

- **Object recognition (in pixels)**
  - SegNet / FCN

**Image**

- Dog
- Cat

**Probability**

- Front of Car
- Stop Sign

**Road**

- Vehicle

MATLAB EXPO 2018
Object Recognition using Deep Learning

Object recognition (whole image)

CNN (Convolutional Neural Network)

Object detection and recognition

R-CNN / Fast R-CNN / Faster R-CNN

Object recognition (in pixels)

SegNet / FCN

Image → Probability → Dog, Cat → Front of Car → Stop Sign → Road → Vehicle
Convolutional Neural Networks

- Train “deep” neural networks on structured data (e.g. images, signals, text)
- Implements Feature Learning: Removes need for “hand crafted” features
- Be trained using GPUs for performance
Two approaches for Deep learning

**Approach 1. Train a Deep Neural Network from Scratch**

- **Convolutional Neural Network (CNN)**
  - Learned features
  - Lots of data

**Approach 2. Fine-tune a pre-trained model (Transfer learning)**

- **Pre-trained CNN**
  - Fine-tune network weights
  - Medium amounts of data

New Task

- **Car ✓**
- **Truck ✗**
- **Bicycle ✗**
**Another Deep Learning Workflows: Feature Extraction**

*Use a pretrained CNN as an automatic feature extractor*

**Recommended when:**
the accuracy is not high enough using transfer learning

<table>
<thead>
<tr>
<th>Training data</th>
<th>100s to 1000s of labeled images (small)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computation</td>
<td>Moderate computation (GPU optional)</td>
</tr>
<tr>
<td>Training Time</td>
<td>Seconds to minutes</td>
</tr>
<tr>
<td>Model accuracy</td>
<td>Good, depends on the pre-trained CNN model</td>
</tr>
</tbody>
</table>

---

**CNN Feature Extractor**

1. Traditional Feature Extraction
2. (Deep Learning)
3. Classification (Machine Learning)
   - SVM, Logistic, etc.

- **Dog ✓**
- **Cat ✗**
Deep Learning Application Approaches

- Train a Deep Neural Network from Scratch
- Fine-tune a pre-trained model (Transfer learning)
- Use a pretrained CNN as an automatic feature extractor
Deep Learning Application Approaches

Fine-tune a pre-trained model (Transfer learning)

Use a pre-trained CNN as an automatic feature extractor
Deep Learning Workflow from Scratch

Repeat these steps until network reaches desired level of accuracy

- Preprocess Images
- Define Layers in CNN
- Set training options
- Train the network
- Test trained network
- Deploy trained network

MATLAB EXPO 2018
Example: MNIST, The “Hello, World!” of computer vision

<table>
<thead>
<tr>
<th>What?</th>
<th>A set of handwritten digits from 0-9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Why?</td>
<td>An easy task for machine learning beginners</td>
</tr>
</tbody>
</table>
| How many?   | 60,000 training images  
                10,000 test images |
| Best results? | 99.79% accuracy |

Sources:  
http://yann.lecun.com/exdb/mnist/  
https://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results
Example: MNIST, The “Hello, World!” of computer vision

Handwritten Character
28 x 28 pixel

- Typical CNN Architecture
- Learning is performed by error back propagation method

Number (0-9)
Train a Deep Neural Network from Scratch

- Example of network construction with 28x28 sized image

```matlab
layers = [...
    imageInputLayer([28 28 1], 'Normalization', 'none');
    convolution2dLayer(5, 20);
    reluLayer();
    maxPooling2dLayer(2, 'Stride', 2);
    fullyConnectedLayer(10);
    softmaxLayer();
    classificationLayer();
];

opts = trainingOptions('sgdm', 'MaxEpochs', 50);
net = trainNetwork(XTrain, TTrain, layers, opts);
```

Define Layers as in order
Define learning rate and maximum iteration number
Call learning functions

Automatically judge the presence or absence of GPU. If there is GPU, learning with CPU if not.
Demo
Hurdle to create a network from scratch

- What kind of learning does the network with high precision do?

- A hurdle when creating a network with scratch
  - Knowledge of network construction
  - Large number of image sets
  - Extensive calculation cost

**AlexNet**
- 5 to 6 days of learning with 2 NVIDIA® GeForce® GTX 580

**VGG Net**
- 2 to 3 weeks study with NVIDIA® GeForce® TITAN Black 4 aircraft

Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton "ImageNet Classification with Deep Convolutional Neural Networks" In NIPS, pp.1106-1114, 2012

Source: ILSVRC Top-5 Error on ImageNet
Deep Learning Application Approaches

Train a Deep Neural Network from Scratch

Fine-tune a pre-trained model (Transfer learning)

Use a pretrained CNN as an automatic feature extractor
Deep Learning Application Approaches

1. Train a Deep Neural Network from Scratch
   - Use a pretrained CNN as an automatic feature extractor

2. Fine-tune a pre-trained model (Transfer learning)
   - Medium amounts of data

Fine-tune network weights

New Task

Car ✓
Truck ✗
Transfer Learning in CNN: Replace final layers
Transfer Learning Workflow

Load pretrained network

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

Last layers learn task-specific features

1 million images
1000s classes
Transfer Learning Workflow

Load pretrained network

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

1 million images
1000s classes

Replace final layers

New layers learn features specific to your data

Fewer classes
Learn faster

Last layers that learned task specific features
Transfer Learning Workflow

Load pretrained network
- Early layers that learned low-level features (edges, blobs, colors)
- Last layers that learned task-specific features

Replace final layers
- Fewer classes
- Learn faster

Train network
- 1 million images, 1000s classes
- 100s images, 10s classes

Training images

Training options
Transfer Learning Workflow

Load pretrained network
- Early layers that learned low-level features (edges, blobs, colors)
- 1 million images, 1000s classes

Replace final layers
- Fewer classes
- Learn faster
- New layers to learn features specific to your data
- 100s images, 10s classes

Predict and assess network accuracy
- Test images

New Trained Network

Train network

Predict and assess network accuracy

Test images

New Trained Network
Transfer Learning Workflow

Load pretrained network
- Early layers that learned low-level features (edges, blobs, colors)
- 1 million images, 1000s classes

Replace final layers
- Fewer classes
- Learn faster
- New layers to learn specific features to your data

Deploy results

Train
- Probability
- Boat
- Plane
- Car

Test images
- Trained Network

Predict and assess network accuracy

MATLAB EXPO 2018
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
- Fewer classes, learn faster
- New layers to learn features specific to your data
- Fewer classes, learn faster

Train network
- Training images
- Training options
- 100s images, 10s classes

Predict and assess network accuracy
- New Trained Network
- Test images

Deploy results
- Probability
- Boat
- Plane
- Car
- Train
- Predict and assess network accuracy
Example: Transfer Learning

Validation Accuracy : 100%

MathWorks Logo

Laptop

Smart Phone

Clip
Deep Learning Application Approaches

Train a Deep Neural Network from Scratch

Fine-tune a pre-trained model (Transfer learning)

Use a pretrained CNN as an automatic feature extractor
Deep Learning Application Approaches

Train a Deep Neural Network from Scratch

Fine-tune a pre-trained model (Transfer learning)
Transfer Learning: Feature extraction

Input → Convolution + ReLu → Pooling → Convolution + ReLu → Pooling → Feature Learning → SVM, Logistic, etc.

Machine learning
What is Feature Extraction?

- Image Pixels
- Color
- Histogram of Oriented Gradients
- Edges

Dense → Sparse
Example: Image anomaly Detection Task

100 hexagon nuts

- **Task**: Find 4 defective units in 100 test images.

- **Challenge**
  - Number of defective units is very small.
  - Difficult to apply supervised learning to this task.
Example: Image anomaly Detection Task

Load pretrained network
- Early layers that learned low-level features (edges, blobs, colors)
- Last layers that learned task specific features
- Load pretrained AlexNet

Extract features and train 1-class SVM
- Extract features with pre-trained model (AlexNet)
- Train 1-class SVM with 100 images
- Unsupervised training

Predict with Test images
- Predict
- List sorted by predicted score
- Find 4 defective units from 100 test images
Demo
Object Recognition using Deep Learning

- **Object recognition (whole image)**
  - CNN (Convolutional Neural Network)

- **Object detection and recognition**
  - R-CNN / Fast R-CNN / Faster R-CNN

- **Object recognition (in pixels)**
  - SegNet / FCN

**Image**
- **Probability**
  - Dog
  - Cat
  - Front of Car
  - Stop Sign
  - Road
  - Vehicle

MATLAB EXPO 2018
R-CNNs: Regions with Convolutional Neural Networks

- Object detection and identification method combining CNN with computer vision method

Detection example when R-CNN learns car front and stop signs

MATLAB EXPO 2018
R-CNNs: Regions with Convolutional Neural Networks

- Take a neural network trained for **image classification** and modify it for **object detection**.

1. For a given input image, region proposals (ROIs) are generated
2. Each region proposal is then independently send through the CNN to compute features
3. The neural network makes a classification for each region proposal

The differences between the R-CNN methods occur at the region proposal stage.
R-CNNs Learning

- \( \text{>> detector} = \text{trainRCNN} \text{ObjectDetector(groundTruth, network, options)} \)
  
  FasterRCNN
  
- \( \text{>> [bbox, score, label]} = \text{detect(detector, image)}; \)
Ground truth Labeling

“How do I label my data?”

New App for Ground Truth Labeling
Label pixels and regions for semantic segmentation

Data
Example: Detect and Classify hexagon nut using Faster RCNN

- Detect and classify hexagon nut into Top/Bottom using Faster R-CNN.
- Features of Top / bottom are Same shape, Same color, Same size and Tiny difference in texture
Demo
Example: Detect and Classify hexagon nut using Faster R-CNN

Control manipulator robot using this object recognition.
Object Recognition using Deep Learning

Object recognition (whole image)

CNN (Convolutional Neural Network)

Object detection and recognition

R-CNN / Fast R-CNN / Faster R-CNN

Object recognition (in pixels)

SegNet / FCN

Image

Probability

Front of Car

Stop Sign

Road

Vehicle

MATLAB EXPO 2018
Semantic Segmentation

Original Image

ROI detection

Pixel classification
Semantic Segmentation Network

- Boat
- Airplane
- Other classes
Semantic Segmentation Network
Semantic Segmentation

CamVid Dataset
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
Semantic Segmentation

- A method of categorizing each pixel based on its meaning.

- Distinguish between sidewalks and roadways
- It is not just looking at colors
Example : Semantic Segmentation for Free road detection

- Use Semantic Segmentation to detect the free space on the road as well as lanes and pavements
MATLAB deep network in a nutshell

- A MATLAB deep network (**) is a MATLAB object that contains an array of trained layer objects.

- Layers array can be created, imported, edited, plotted in MATLAB.

- Layers are trained with a lot of data and 
  \[
  \text{net} = \text{trainNetwork}(\ldots, \text{layers}), \text{most of the time.}
  \]

- MATLAB deep networks have different usages.

  **Classification**
  - Data=?
  - car \(\times\) dog \(\checkmark\)
  - \(\text{YPred} = \text{classify}(\text{net},X)\);

  **Regression**
  - Data=?
  - 32
  - \(\text{YPred} = \text{predict}(\text{net},X)\);

  **Semantic segmentation (*)**
  - Sky
  - Car
  - Road
  - \(\text{C} = \text{semanticseg}(I,\text{net})\);

  **Object detection (*)**
  - [bboxscores] = detect(\text{net},I);

---

(**) Neural Network Toolbox

(\text{(*) \ Computer Vision Syst. Toolbox})
MathWorks can help you do Deep Learning

Free resources

- Guided evaluations with a MathWorks deep learning engineer
- Proof-of-concept projects
- Deep learning hands-on workshop
- Seminars and technical deep dives
- Deep learning onramp course

More options

- Consulting services
- Training courses
- Technical support
- Advanced customer support
- Installation, enterprise, and cloud deployment
- MATLAB for Deep Learning