

MATLABで試す！機械学習の応用例

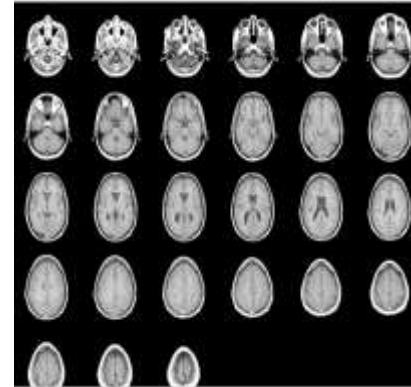
MathWorks Japan
アプリケーションエンジニアリング部
福本 拓司

一般的におこなわれる目視による評価

製造ライン



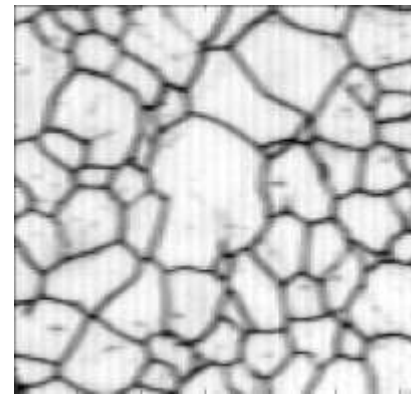
医用データ



作業現場・インフラ



研究データ



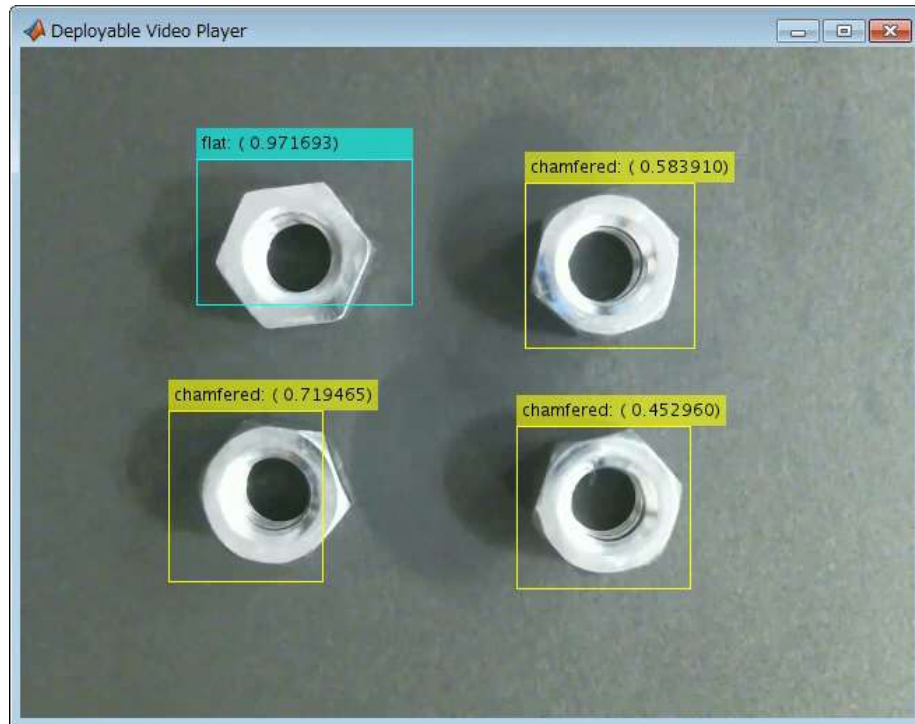
現場での目視



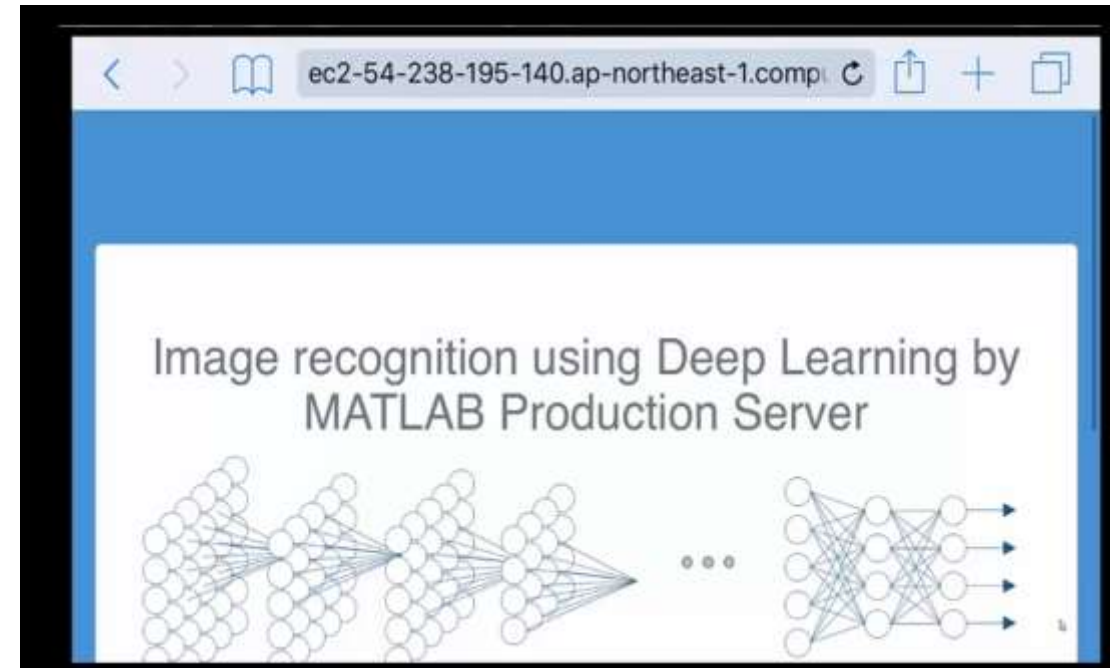
大量画像の収集
専門家によるチェック

ディープラーニング活用によって変わる働き方

Faster R-CNNにより高速に物体
の位置と状態を判断



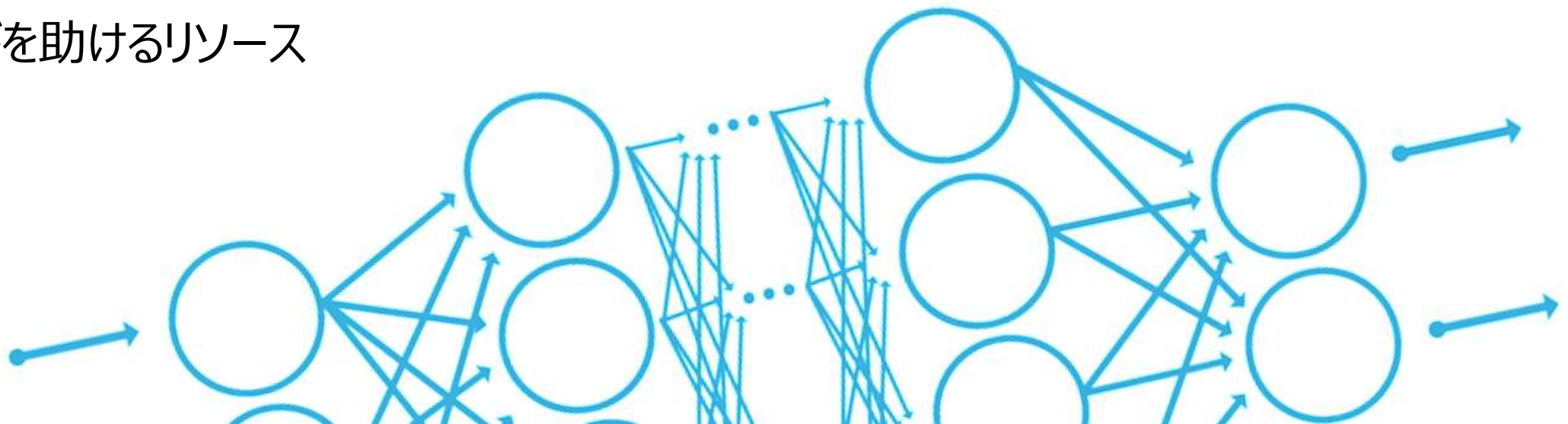
スマホでとった画像をAIで判定



個人差や人為的なミスの影響を受けない判断が可能
⇒ 自動化による人件費削減
⇒ 特定分野の知識の教育が不要

Agenda

- これからはじめるディープラーニング活用ワークフロー
- ディープラーニングの異常検知への適用
- 物体検出と領域の切り出し
- 学びを助けるリソース



機械学習を利用すべき場面とは？

画像処理のアルゴリズム



手書き
プログラム

'Face'

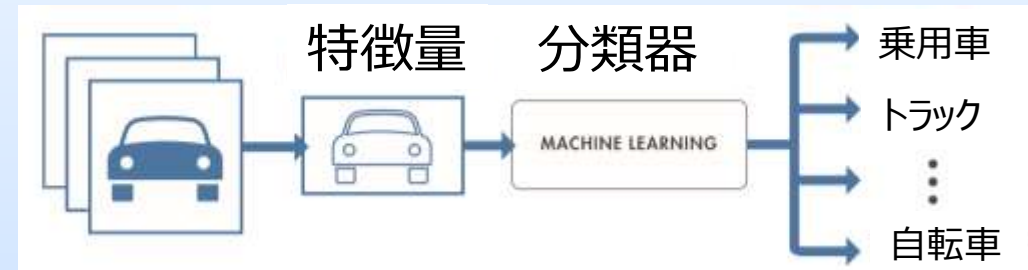
'Helmet'

```
If brightness > 0.5  
    then 'Helmet'  
If edge_density < 4 and major_axis > 5  
    then    ...
```

数値で条件を指定し切り分け

明確な切り分けが可能の場合に○

機械学習



分類器 = <機械学習アルゴリズム>(学習データ, ラベル)

画像データを使って分類器を学習

うまくくと柔軟な切り分けができる

機械学習を利用すべき場面とは？

画像処理のアルゴリズム



手書き
プログラム

'Face'

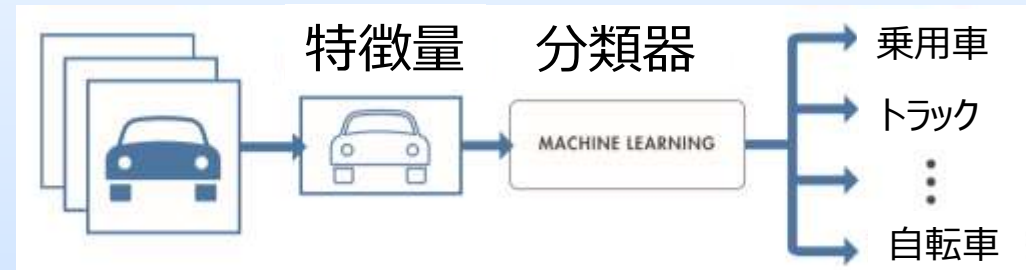
'Helmet'

```
If brightness > 0.5  
    then 'Helmet'  
If edge_density < 4 and major_axis > 5  
    then ...
```

数値で条件を指定し切り分け

明確な切り分けが可能の場合に○

機械学習



Deep Learning



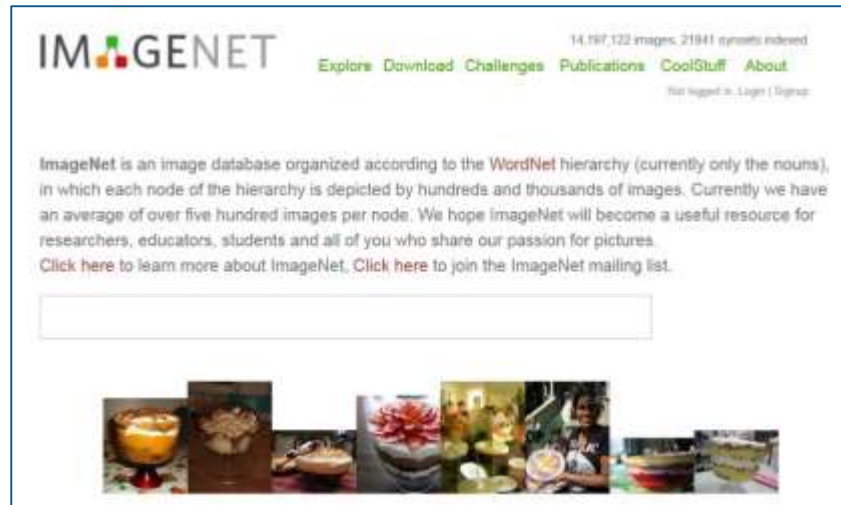
画像データを使って分類器を学習

うまくくと柔軟な切り分けができる

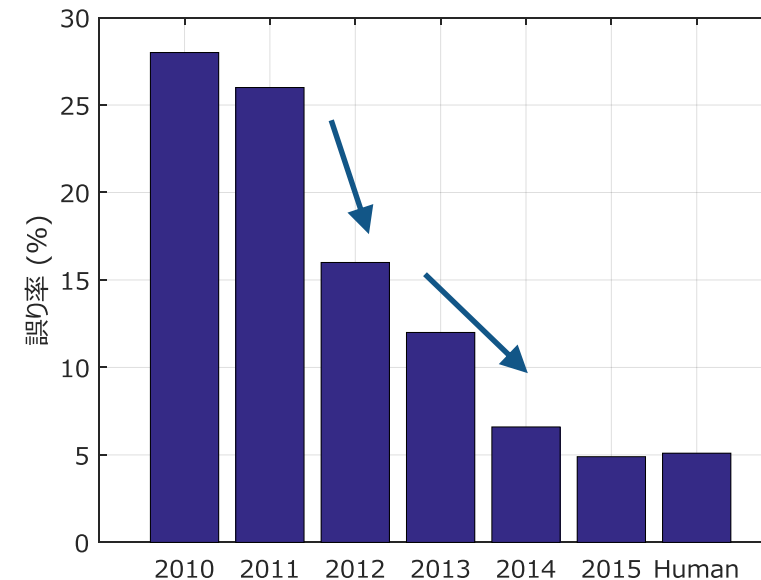
ILSVRC (ImageNet Large Scale Visual Recognition Challenge)

ImageNet とは ?

→画像認識の研究のための大規模な画像データベース
→1000のカテゴリを持ち、カテゴリ毎に1000枚の画像



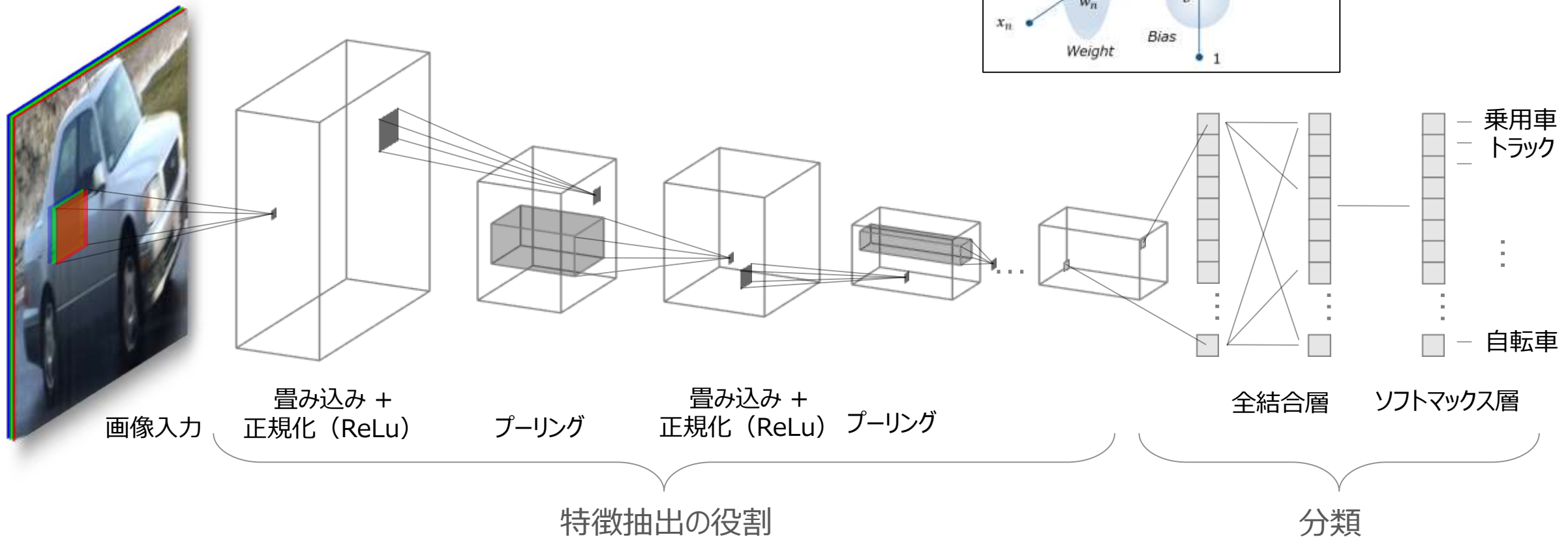
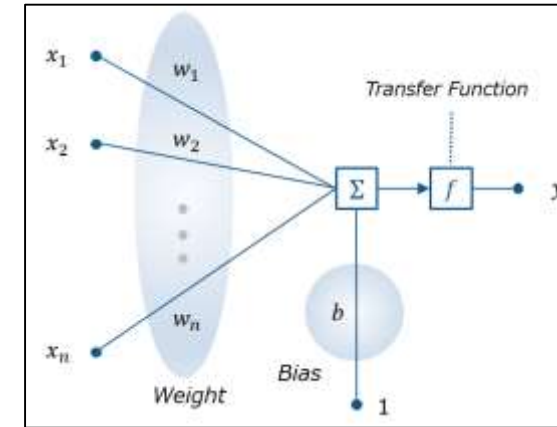
<http://www.image-net.org/>



CNNの登場によって10%以上の性能向上(2012)
GoogLeNet, VGG等の深いCNNが登場(2014)

畳み込みニューラルネットワーク(Convolutional Neural Network)

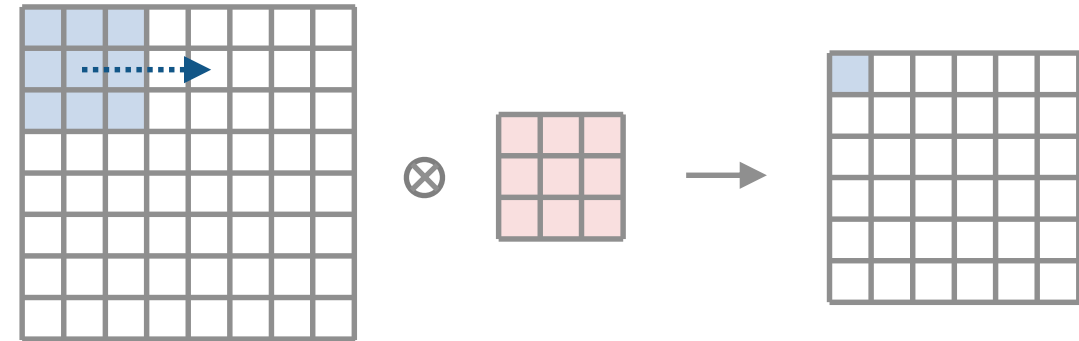
畳み込み層・プーリング層・正規化層などを
積み重ねて作られた多層のニューラルネットワーク



Convolution Layer (畳み込み層) / Pooling Layer (プーリング層)

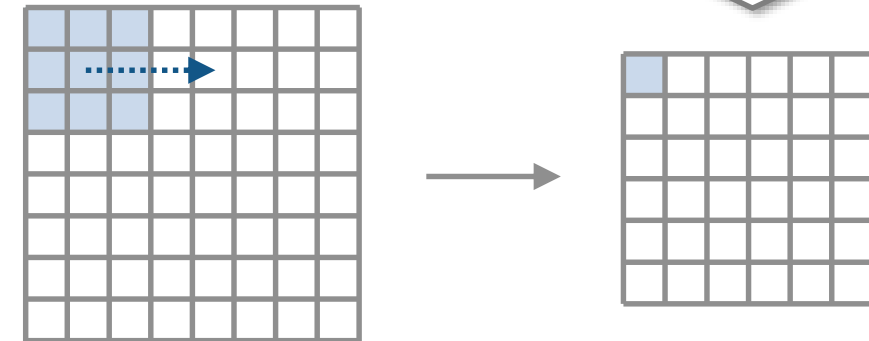
Convolution Layer (畳み込み層)

- 画像のフィルタ処理に相当する処理
- 特徴抽出器としての役割



Pooling Layer (プーリング層)

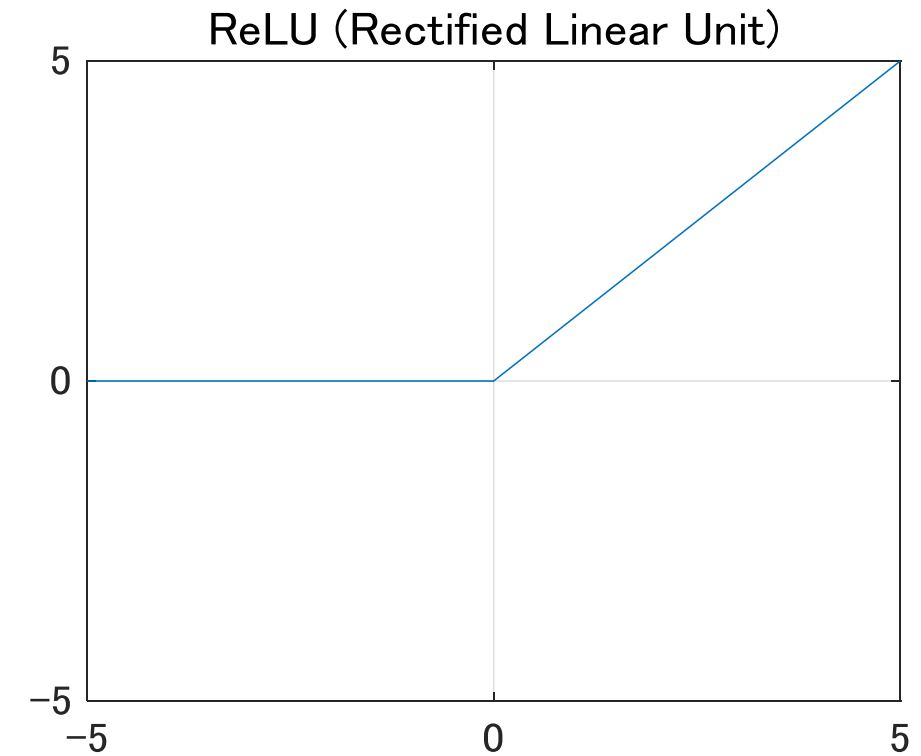
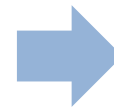
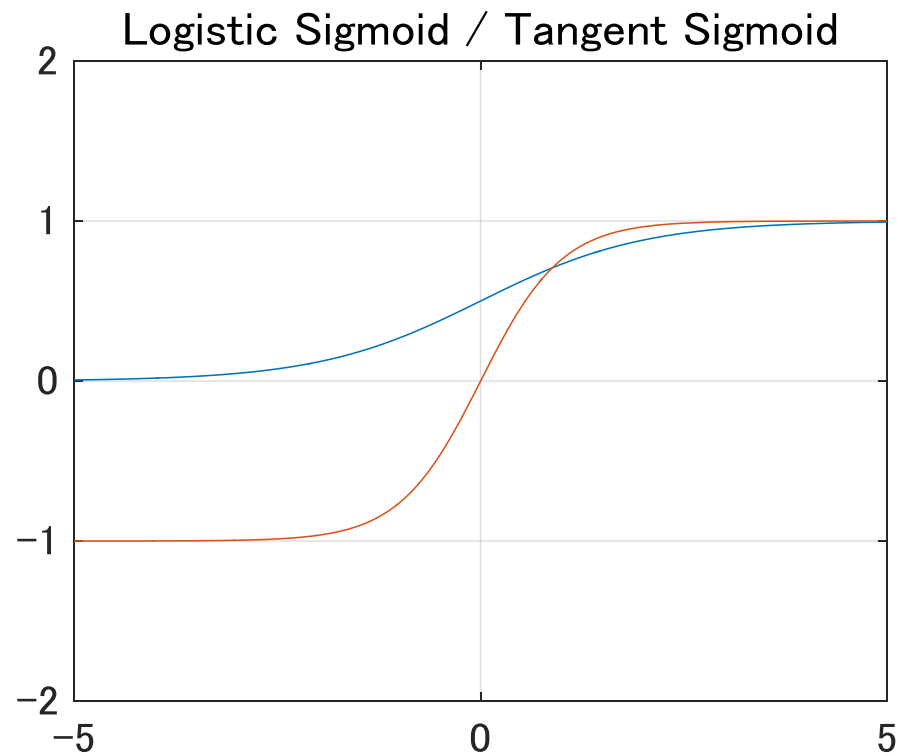
- 領域内の最大値または平均値を出力
- 平行移動等に対するロバスト性に関係
- スライドと呼ばれる間引きを行うこともある



最大値を出力する場合 : Max Pooling
平均値を出力する場合 : Average Pooling

ReLU層 (Rectified Linear Unit)

Logistic Sigmoid等の値が飽和する関数より、ReLUの方が学習が早く進むことがわかってきた



画像分野におけるディープラーニングの新展開

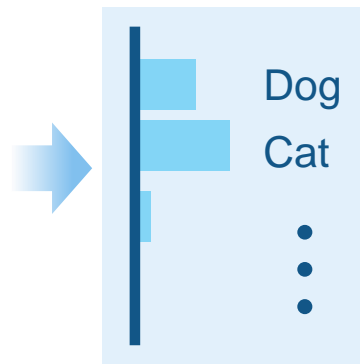
R2017a
R2017b

画像分類（画像全体）

CNN (Convolutional Neural Network)



画像



確率値

物体の検出

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



自動車の前面

停止標識

物体の検出と領域の切り出し

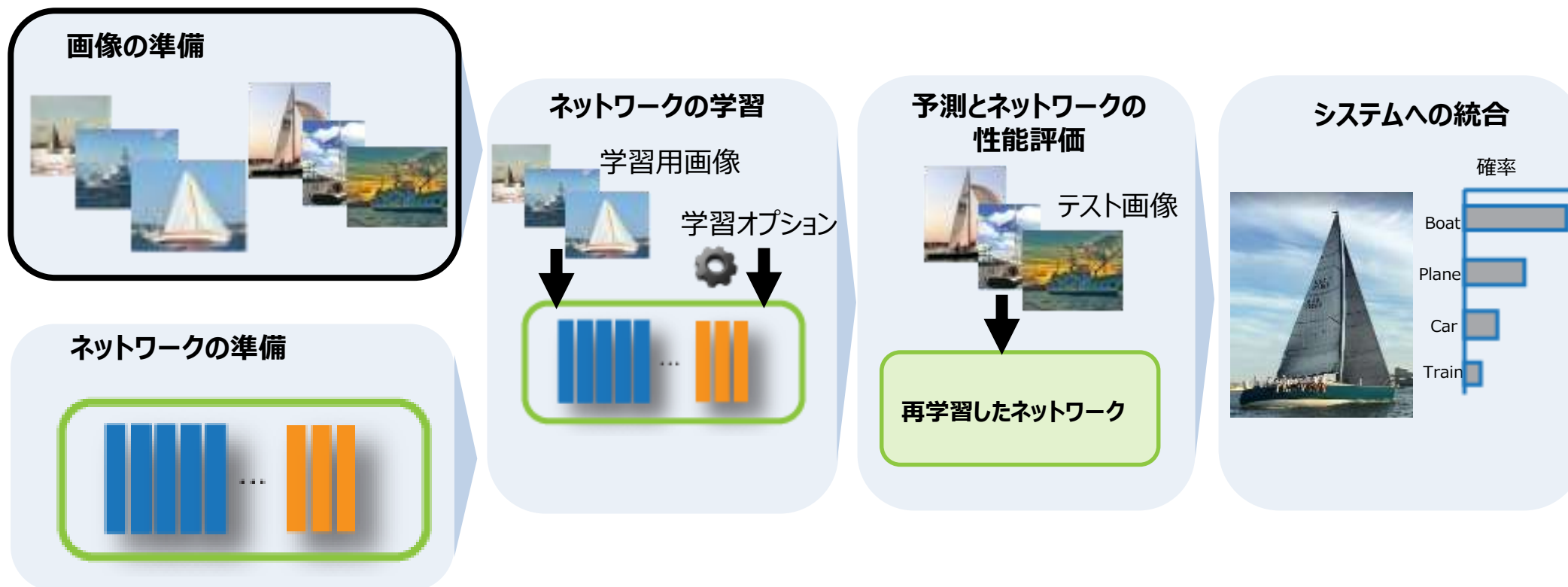
SegNet / FCN



車道

自動車

MATLABによるディープラーニングワークフロー



効率よい学習と精度向上のためのデータ拡張

- 大量の画像セットにメモリ効率の良いアクセス

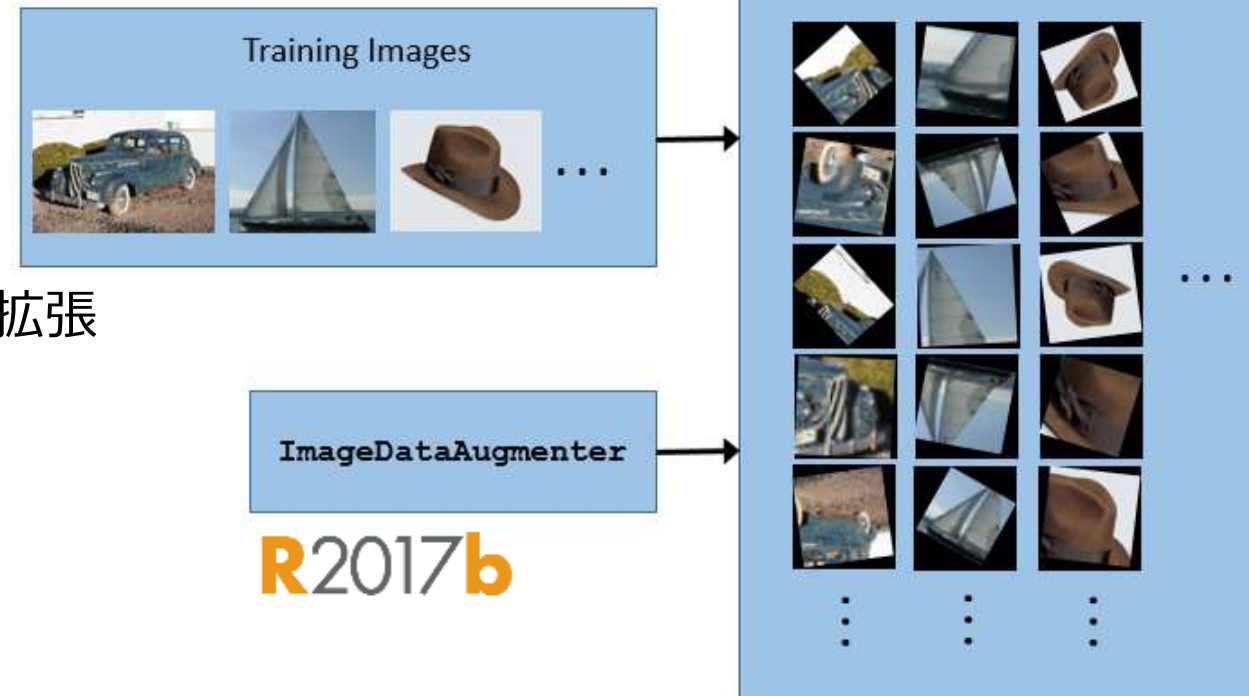
- imageDatastore

- データ拡張

- imageDataAugmenterによる画像データの拡張

- スケール
- せん断
- 回転

等の制限をかけながらデータを拡張



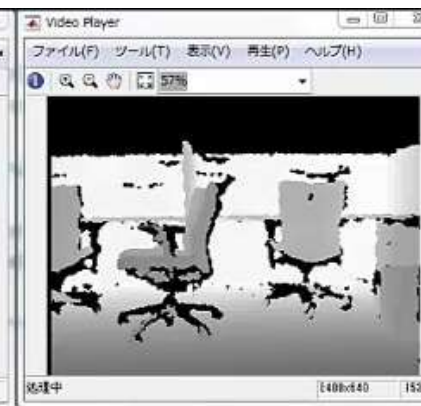
各種カメラデバイスからのデータ取り込み

Image Acquisition Toolbox

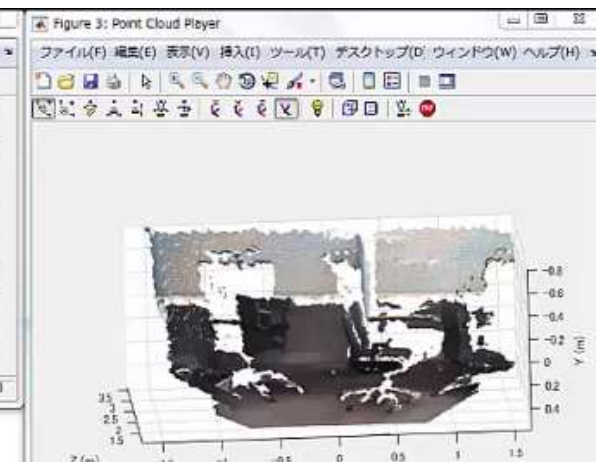
- 業界標準のHWからの動画像取込み機能を提供
 - フレームグラバ
 - Analog
 - Camera Link
 - DCAM 互換 FireWire (I IDC 1394)
 - GigE Vision
 - USB3 Vision
 - IPカメラ
- Microsoft Kinect



RGB

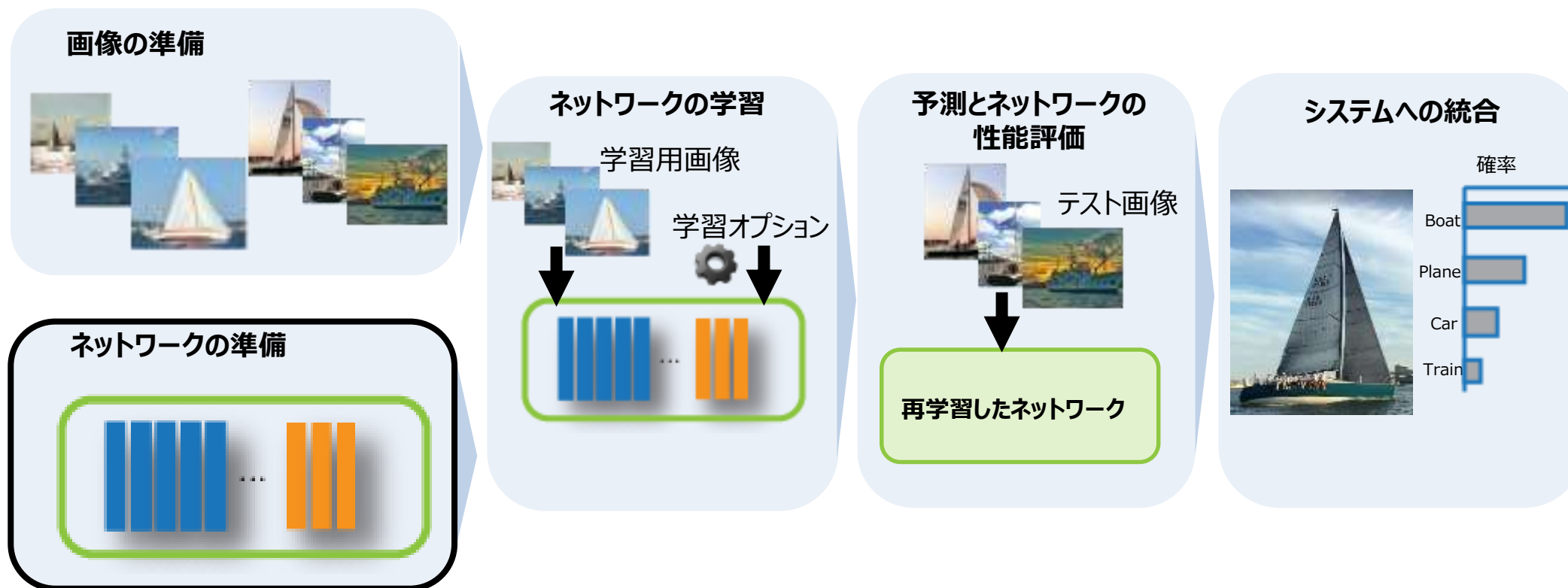


Depth



ポイントクラウド

MATLABによるディープラーニングワークフロー



畳み込みニューラルネットの構築と学習

28×28 ピクセルの画像（数字）を認識させる例題でのネットワーク構築の例



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

畳み込み層・プーリング層・正規化層
などの層を積み上げて定義

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

学習率や最大反復数などを定義して
学習の関数を呼び出す

GPU有無を自動で判定。あればGPU,なければCPUで学習。

<http://www.mathworks.com/help/releases/R2017b/nnet/ref/trainnetwork.html>

よくある質問

- GPUが合ったほうがいいと聞くけどどれくらい違うの

- 最適なLearningRateについて

確率的勾配降下法

$$\theta_{\ell+1} = \theta_{\ell} - \alpha \nabla E(\theta_{\ell}),$$

- 学習中の表示について

θ : パラメータ(WeightやBias)

α : Learning Rate

E : Loss Function

ネットワークをスクラッチで作るハードル

精度が高いネットワークはどのような学習をしているのか

・100万枚の画像セット

AlexNet

■ NVIDIA® GeForce® GTX 580 2機 による 5～6日間の学習

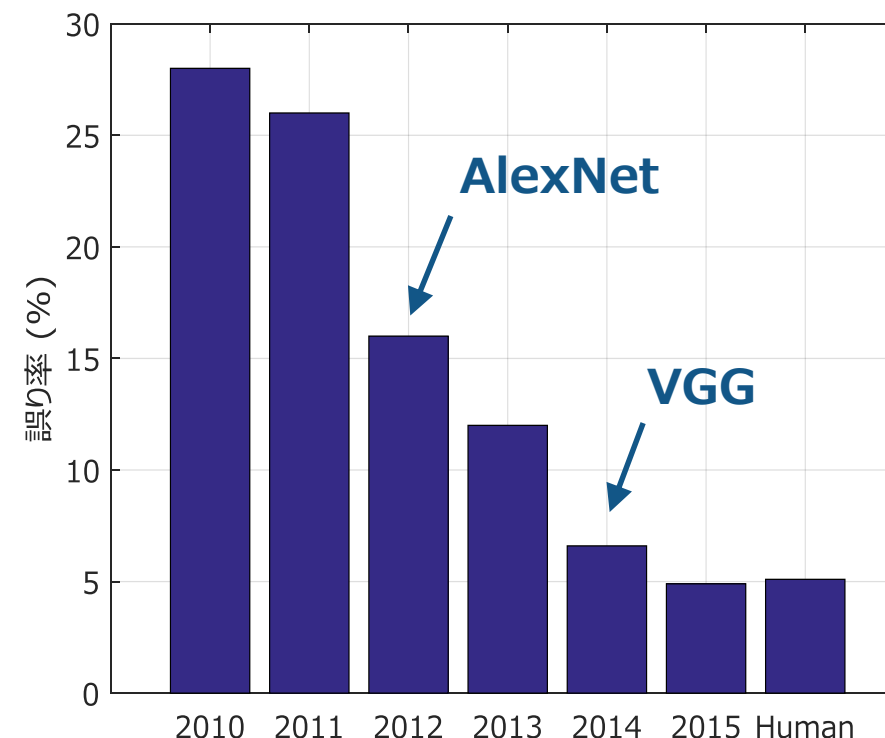
VGG Net

■ NVIDIA® GeForce® TITAN Black 4機 による 2～3週間の学習

ネットワークをスクラッチで作る際のハードル

- ・ネットワーク構築の知識
- ・大量の画像セット
- ・膨大な計算コスト

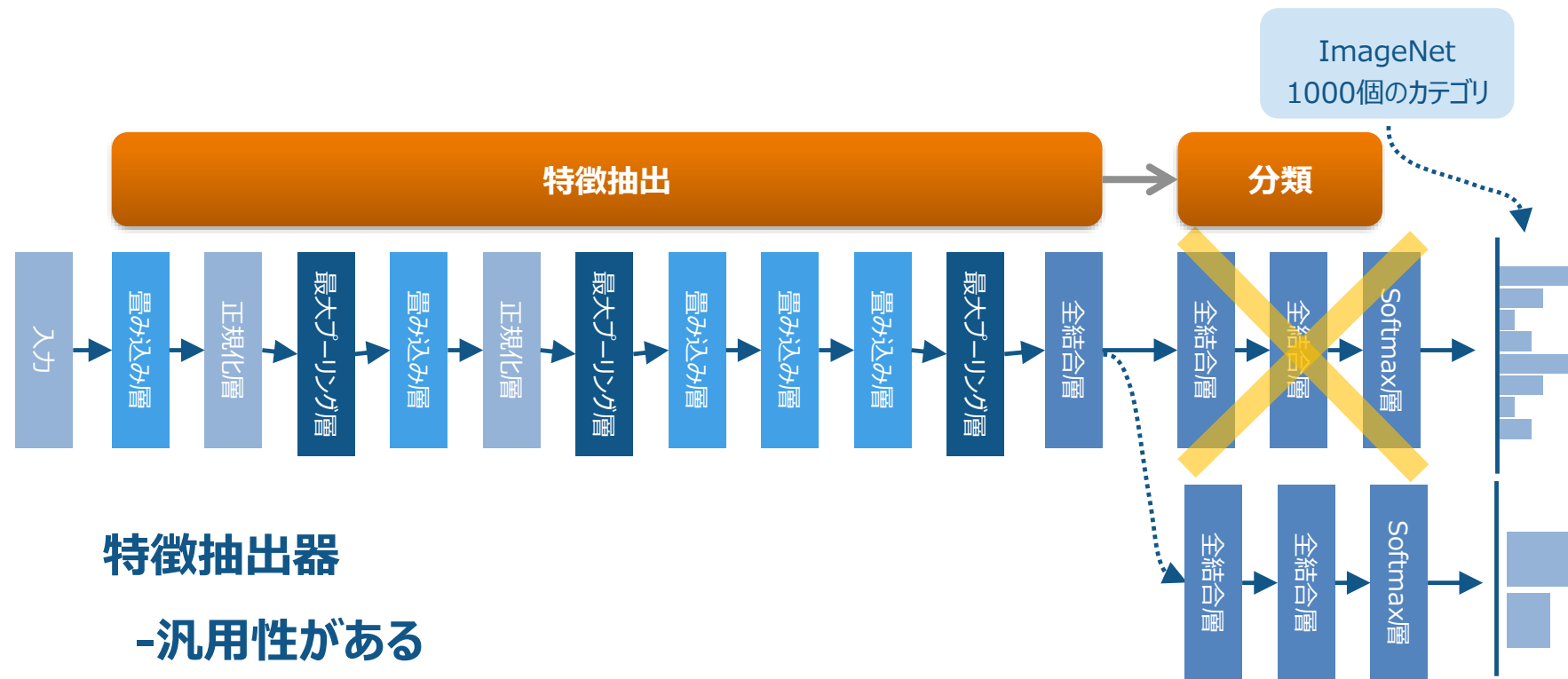
ILSVRC 2010 - 2015



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

K. Simonyan, A. Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition" arXiv technical report, 2014

CNN と転移学習

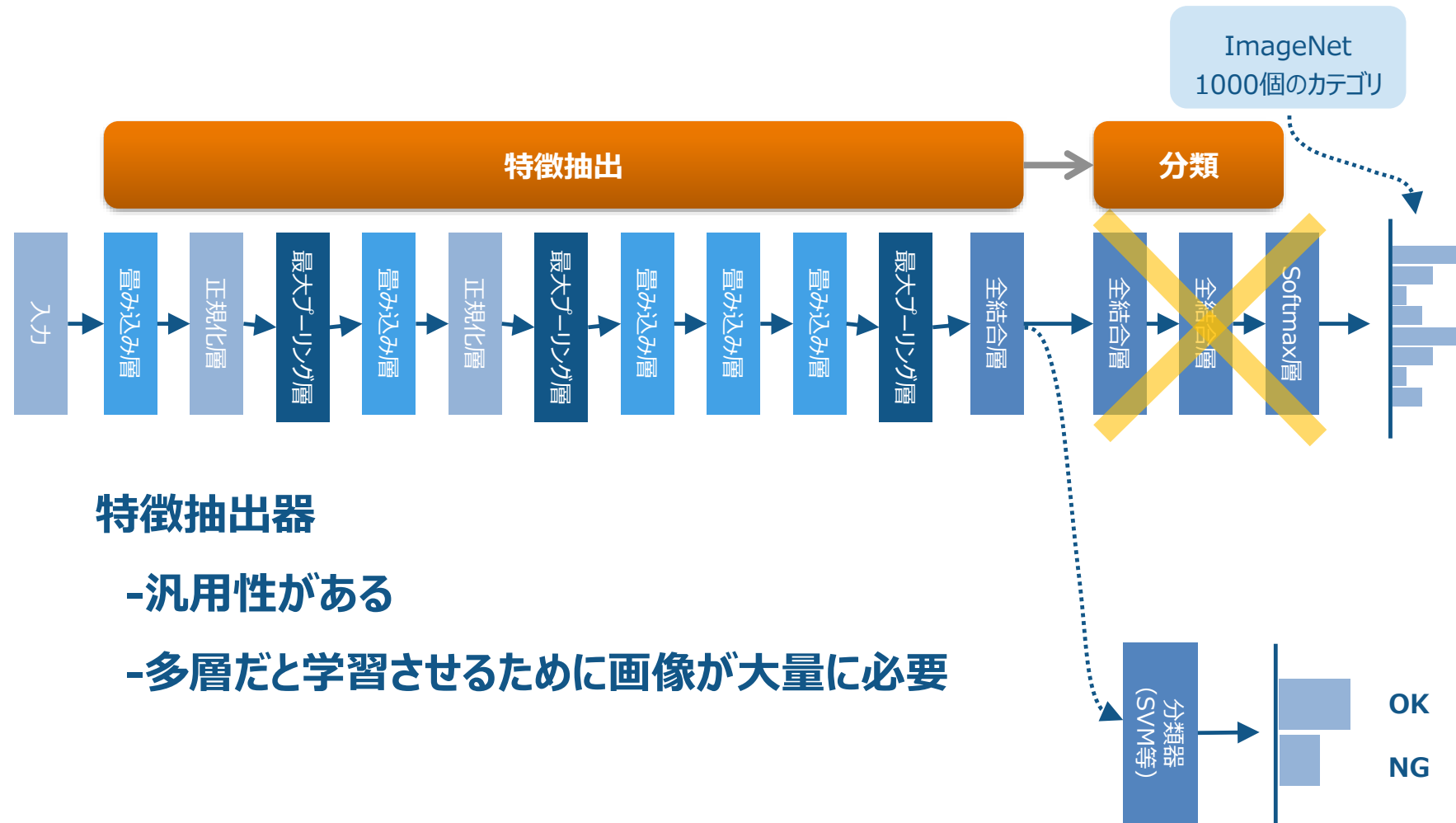


特徴抽出器

- 汎用性がある
- 多層だと学習させるために画像が大量に必要

優れた既存ネットワークの特徴抽出器 + 独自分類器

CNN と転移学習

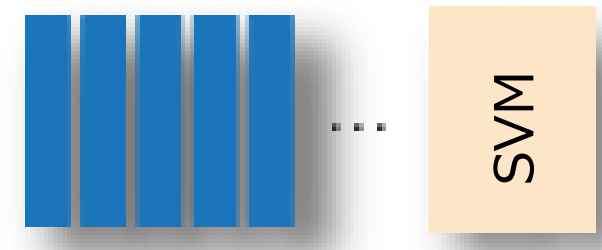
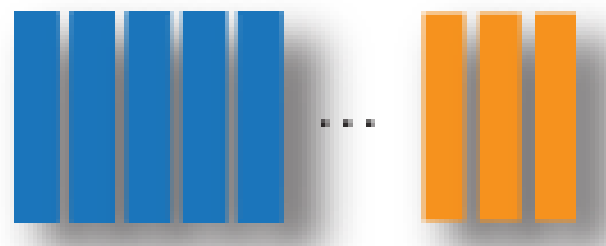


特徴抽出器

- 汎用性がある
- 多層だと学習させるために画像が大量に必要

優れた既存ネットワークの特徴抽出器 + 独自分類器

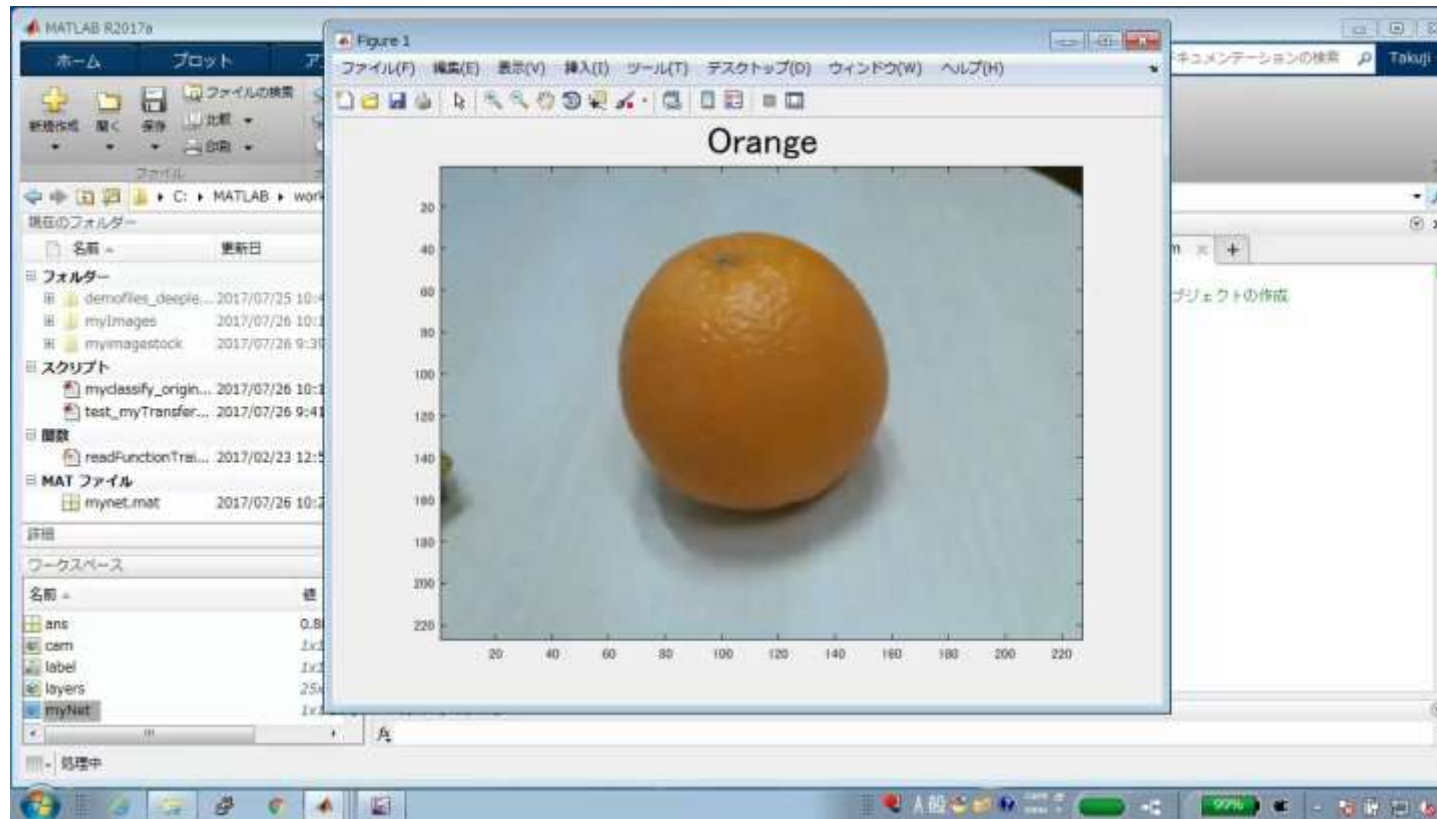
転移学習の種類



分類器をNNで置き換え		特徴抽出器として利用し、SVMなどを接続
再学習によるチューニング	特徴抽出	元のネットワークの特性
NNの全結合層・ソフトマックス層を取替え	分類器	SVM等 機械学習で利用する分類器を利用
独自分類への最適なネットワークを目指す。	メリット	分類器の変更など試すことがわかりやすい。 学習の計算コストが少ない。
学習のパラメータなどの知識が必要。	ハードル	特徴抽出の最適化はできない。

ディープラーニングによる物体認識

ディープラーニング：10行でできる転移学習 ～画像分類タスクに挑戦～



学習した種類：

- オレンジ
- みかん
- グレープフルーツ(ルビー)
- グレープフルーツ(ホワイト)
- レモン

学習画像数：各 20 枚

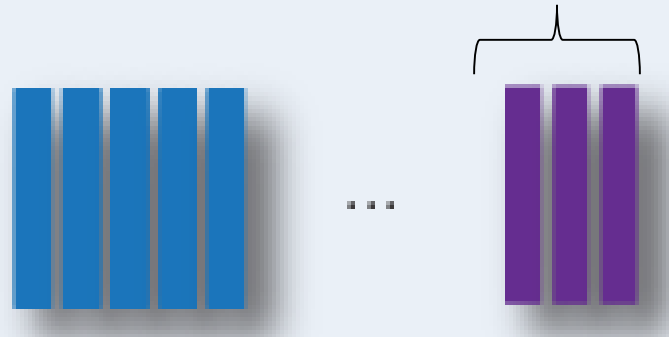
- 要件を満たすPC&MATLAB環境
 - 学習済みAlexnet
 - 画像セット
- で10行のコーディングで始められます

<https://www.youtube.com/watch?v=XMcHiMIT8iE>

転移学習ワークフロー

学習済みネットワーク読み込み

特定のタスクのために
学習された分類器

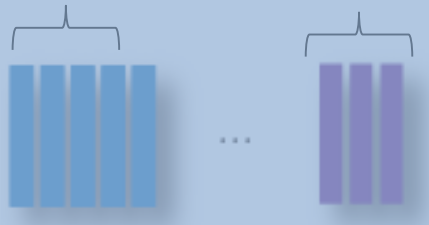


画像数 100万枚以上
クラス 1000以上

転移学習ワークフロー

学習済みネットワーク読み込み

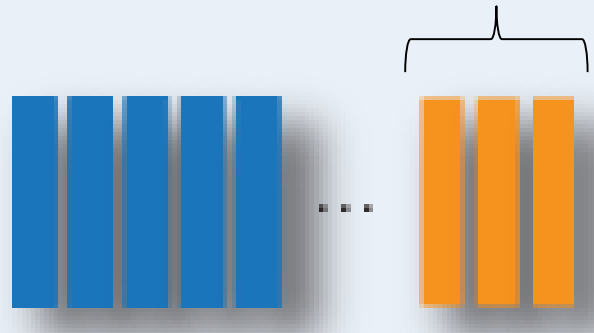
特定のタスクのために
学習された分類器



画像数 100万枚以上
クラス 1000以上

最後段の層を いくつか置き換え

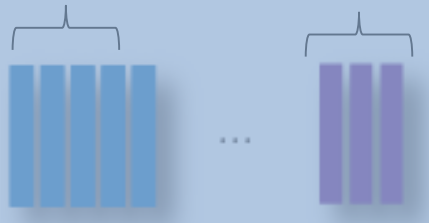
新しいデータを
学習するための新しい層



転移学習ワークフロー

学習済みネットワーク読み込み

特定のタスクのために
学習された分類器



画像数 100万枚以上
クラス 1000以上

最後段の層を いくつか置

新しいデー
タのための新

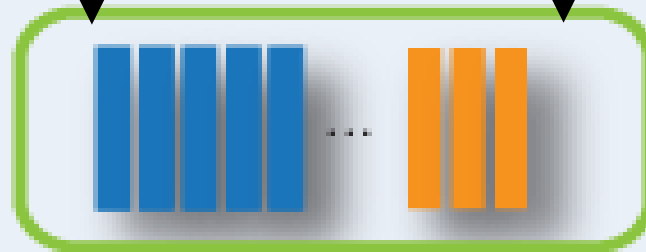


ネットワークの学習

学習用画像



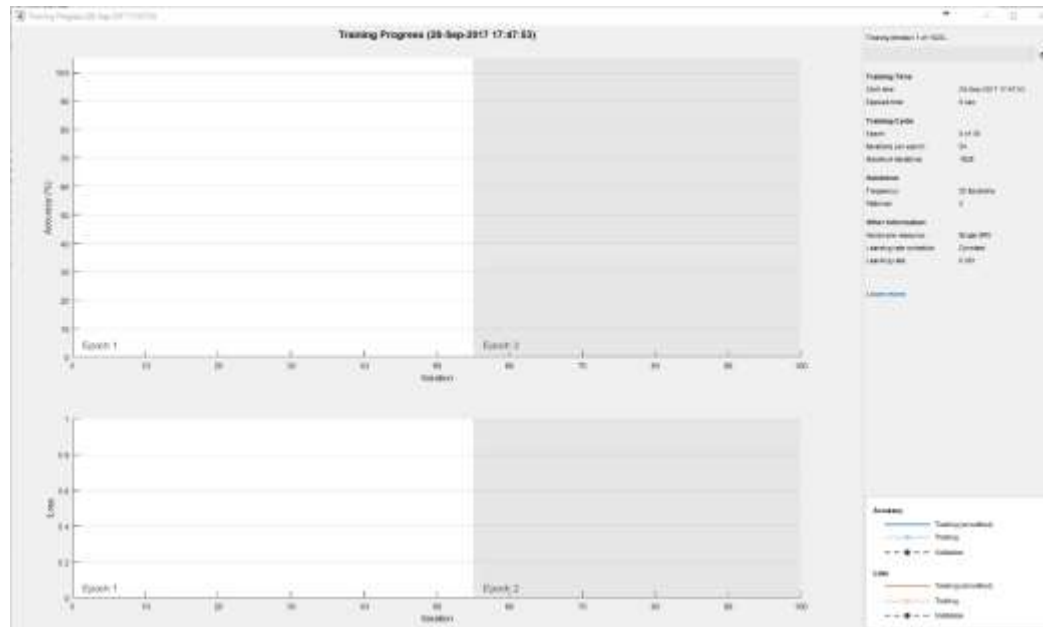
学習オプション



画像 数100枚
クラス 数10カテゴリ

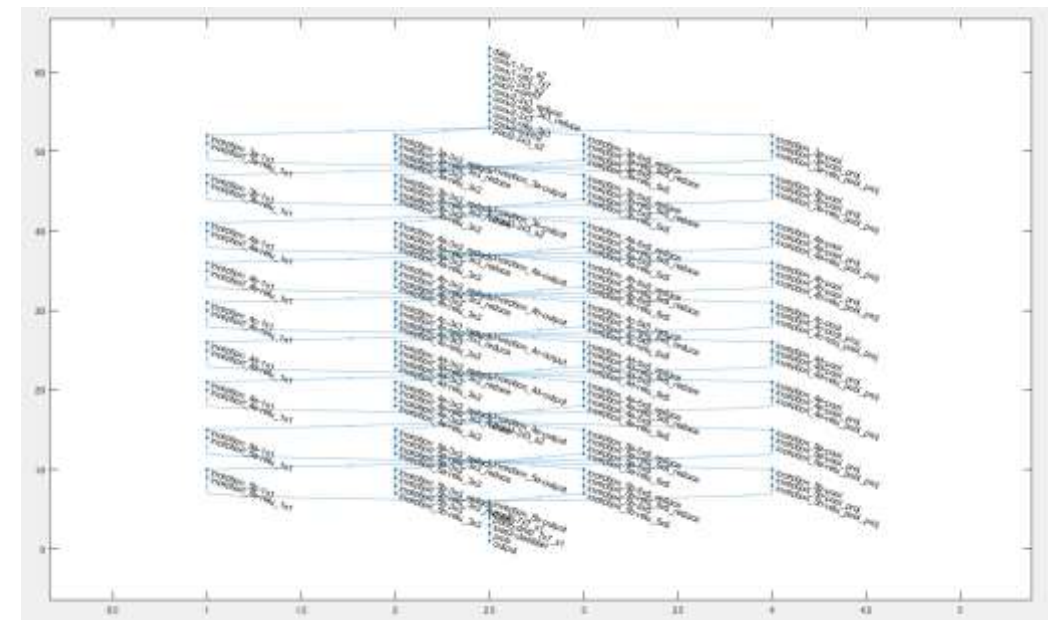
豊富な可視化機能

学習進捗の可視化と学習停止



```
opts = trainingOptions(...
    'Plots', 'training-progress', ...
    'ValidationPatience', 3);
```

ネットワークの可視化



```
lgraph = layerGraph(net);
plot(lgraph)
```

転移学習ワークフロー

学習済みネットワーク読み込み

特定のタスクのために
学習された分類器



画像数 100万枚以上
クラス 1000以上

最後段の層を いくつか置

新しいデー
タのための新



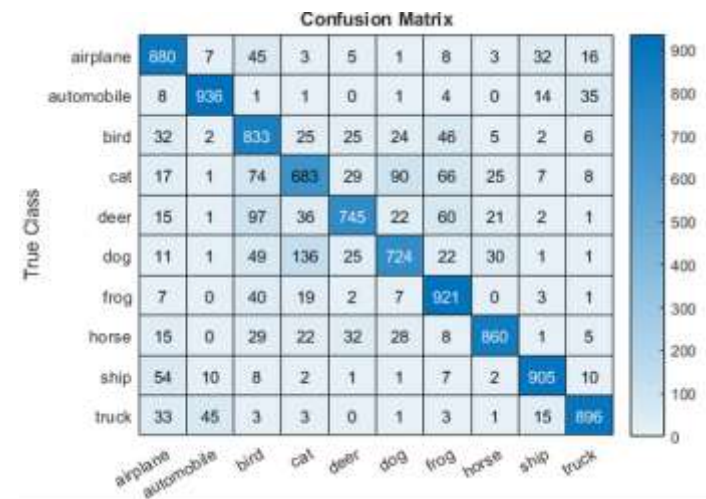
予測とネットワークの 性能評価



テスト画像

再学習したネットワーク

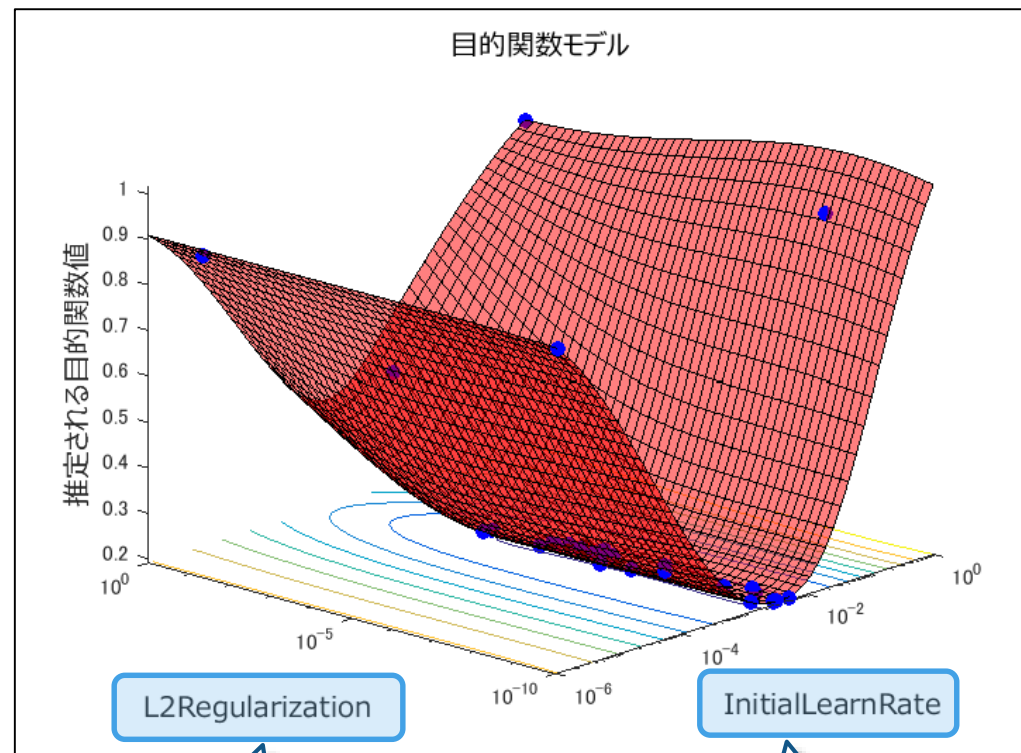
混合行列/ヒートマップ表示



```
[cmat,classNames] =  
confusionmat(testLabels,predictedLabels);  
h = heatmap(classNames,classNames,cmat);
```

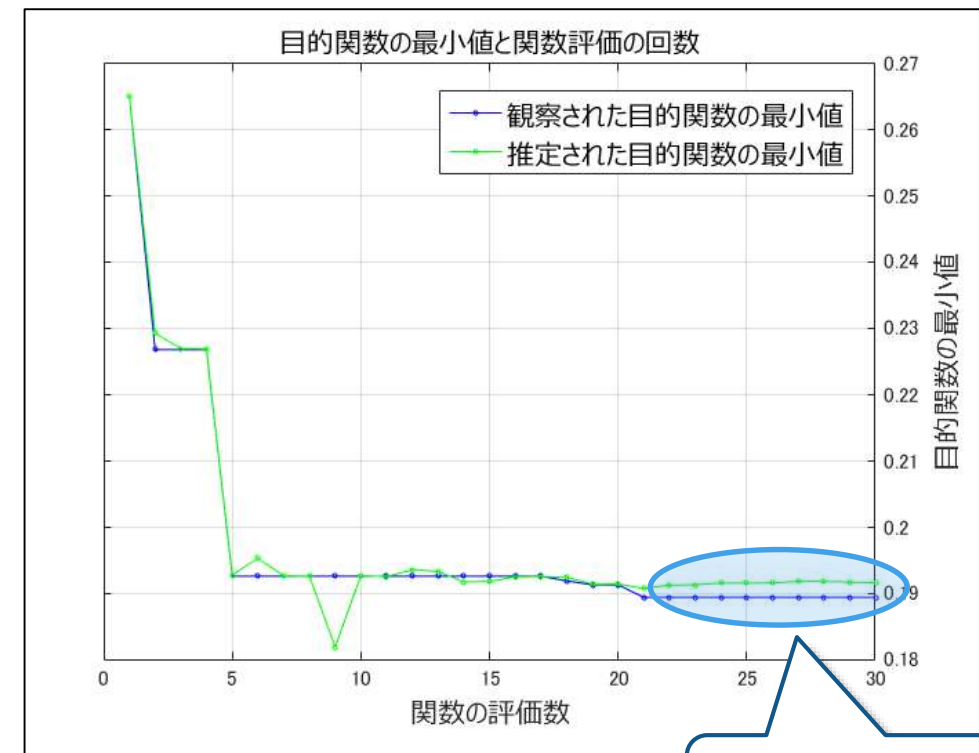
パラメータの自動最適化（ベイズ最適化）

以下は、CNN におけるパラメータ調整にベイズ最適化を用いた例



L2正則化の係数

学習率の初期値



最適なパラメータへ収束

転移学習のための学習済みモデル読み込み

学習済みネットワーク*

- AlexNet
- VGG-16
- VGG-19
- GoogLeNet **New**
- Resnet50 **New**
- InceptionV3 (coming soon)

* 一行でモデル読み込み

他のフレームワークのモデル読み込み

- Caffe Model Importer
- TensorFlow/Keras Model **New** Importer

AlexNet
PRETRAINED MODEL

VGG-16
PRETRAINED MODEL

ResNet
PRETRAINED MODEL

Caffe
MODELS

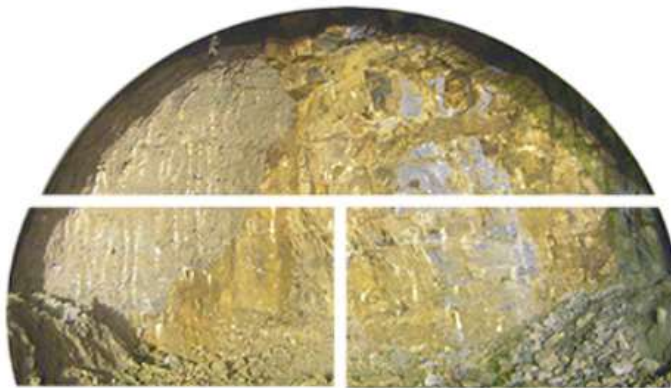
GoogLeNet
PRETRAINED MODEL

TensorFlow/Keras
MODELS

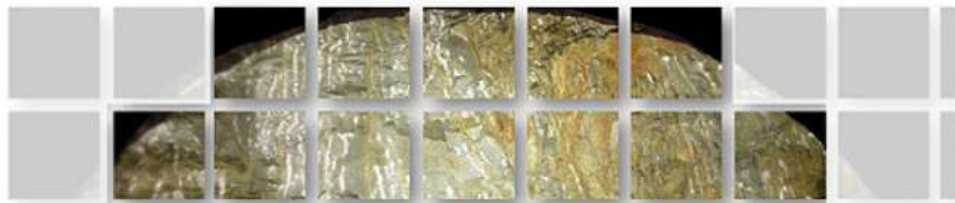
転移学習 + 最適化で多くのことを試すことができます。
どの程度タスクが難しいのか知ることが重要です。

MATLABを使った転移学習の活用例： 株式会社大林組 様

山岳トンネル工事の切羽（掘削面）評価にディープラーニングを適用



従来の切羽の評価領域（上方、左右の3分割）



AlexNet(※3)による切羽の評価領域（分割数は撮影時の画素数によって異なります）

3項目の評価にディープラーニングを適用

- ・風化変質（4分類）
- ・割目間隔（5分類）
- ・割目状態（5分類）

AlexNet + SVMの転移学習

割目状態では89%の的中率



土木学会

第**72**回年次学術講演会にて発表

※大林組様プレスリリースより参照

http://www.obayashi.co.jp/press/news20170912_01

転移学習の利用例が増えています

転移学習ワークフロー

学習済みネットワーク読み込み

特定のタスクのために
学習された分類器



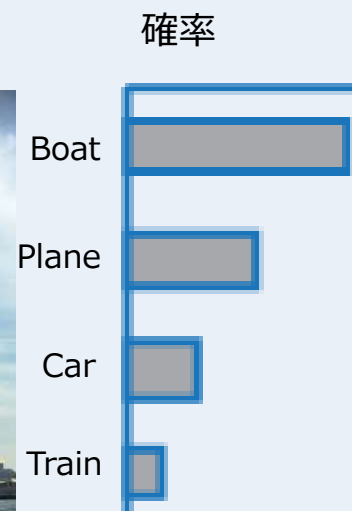
画像数 100万枚以上
クラス 1000以上

最後段の層を いくつか置

新しいデー
タのための新



システムへの統合



予測とネットワークの 性能評価

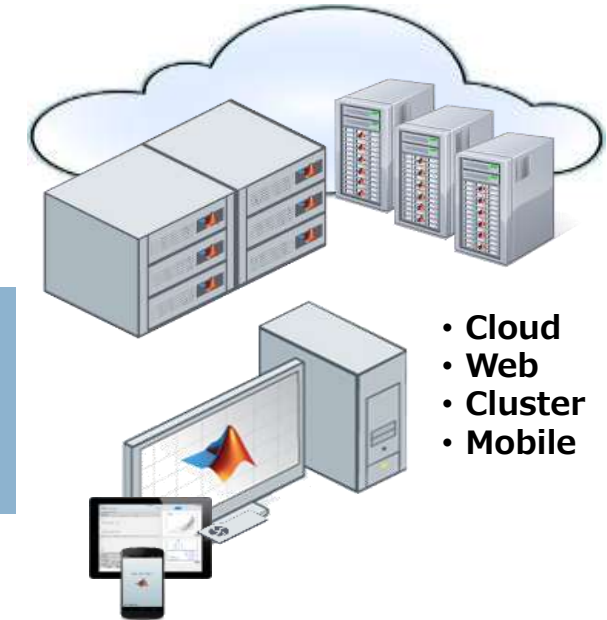


MATLABによるアプリケーション配布

遠隔地での判断、
スマホ、タブレットの利用



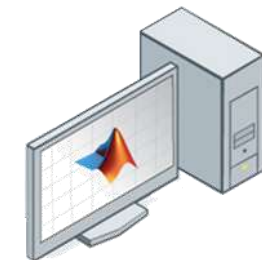
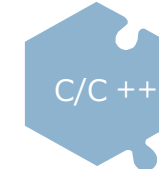
最適なアルゴリズムの探求



- Cloud
- Web
- Cluster
- Mobile



MATLABライセンスのない
PCへの配布



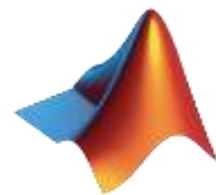
GPU Coder™

New in **R2017b**

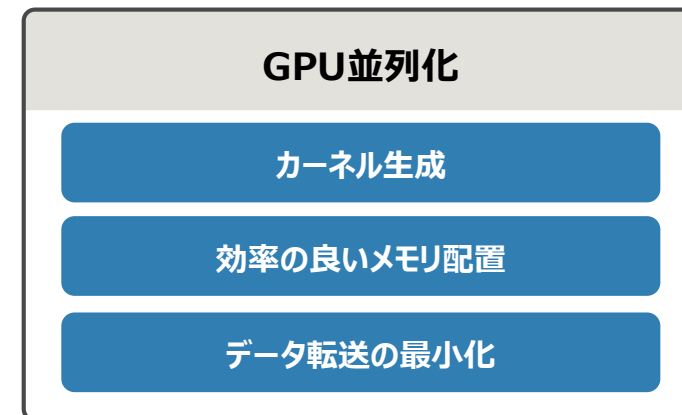
- プラグマによる関数解析とカーネル生成
 - CUDAの文法を知らなくても利用できる
- デザインパタンの利用も可能
 - より確実かつ効率の良いカーネル生成
- GPU Coder専用GUIを使ったコード生成
 - 初めてでも使いやすいGUI



製造ラインなど
リアルタイム性が
必要なケースに有効



MATLAB



NVIDIA GPU

MATLABコードからCUDA Cを生成します

転移学習ワークフロー

学習済みネットワーク読み込み

特定のタスクのために
学習された分類器



画像数 100万枚以上
クラス 1000以上

最後段の層を いくつか置き換え

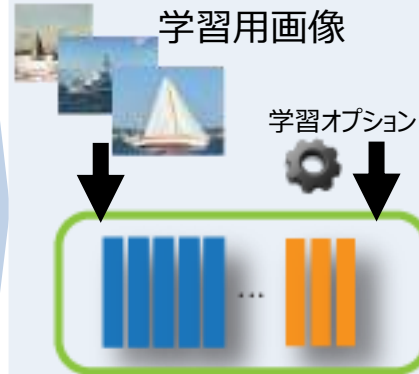
新しいデータを
学習するための新しい層



ネットワークの学習

学習用画像

学習オプション



画像数100枚
クラス 数10カテゴリ

予測とネットワークの 性能評価

テスト画像



再学習したネットワーク

システムへの統合



MATLABは学習済みネットワークからシステムへの統合までサポート
画像セットがあればすぐに始められます。

multi-GPU、クラウド利用による学習の高速化

```
opts = trainingOptions('sgdm', ...  
    'MaxEpochs', 100, ...  
    'MiniBatchSize', 250, ...  
    'InitialLearnRate', 0.00005, ...
```

```
'ExecutionEnvironment', 'auto' );
```

GPUの有無を自動で認識、
あればGPU、なければCPUで学習



CPUのみ



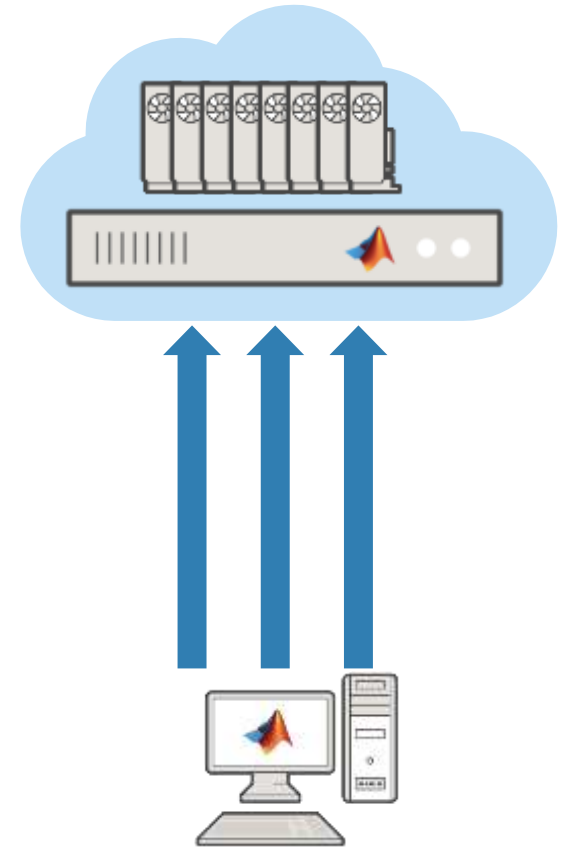
GPU搭載



複数GPU



サーバー/クラウド



More GPUs

multi-GPU、クラウド利用による学習の高速化

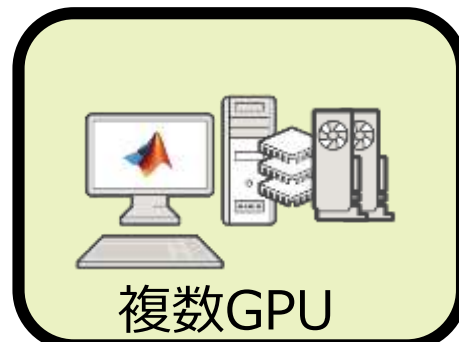
```
opts = trainingOptions('sgdm', ...  
    'MaxEpochs', 100, ...  
    'MiniBatchSize', 250, ...  
    'InitialLearnRate', 0.00005, ...  
    'ExecutionEnvironment', 'multi-gpu' );
```



CPUのみ



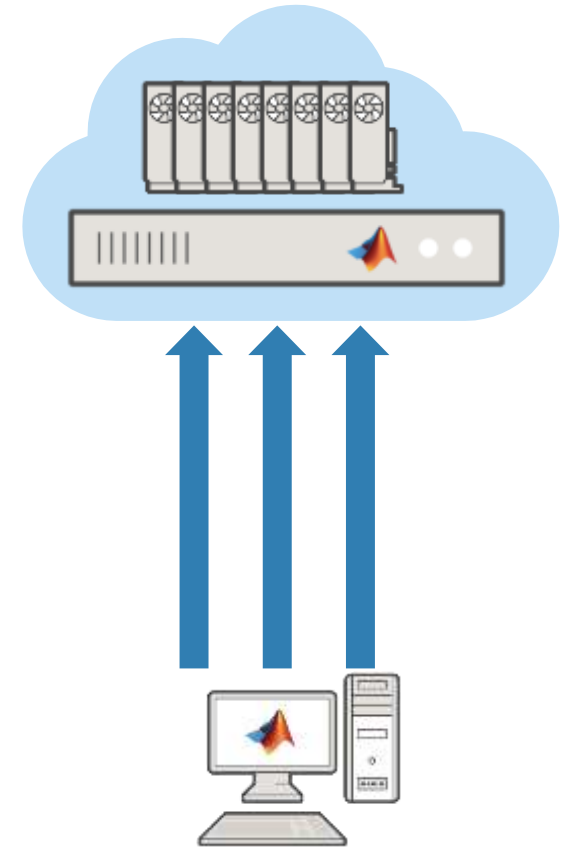
GPU搭載



複数GPU



サーバー/クラウド



More GPUs

multi-GPU、クラウド利用による学習の高速化

```
opts = trainingOptions('sgdm', ...  
    'MaxEpochs', 100, ...  
    'MiniBatchSize', 250, ...  
    'InitialLearnRate', 0.00005, ...
```

```
'ExecutionEnvironment', 'parallel' );
```

• わずかな書き換えでスケールアップ&高速化が可能



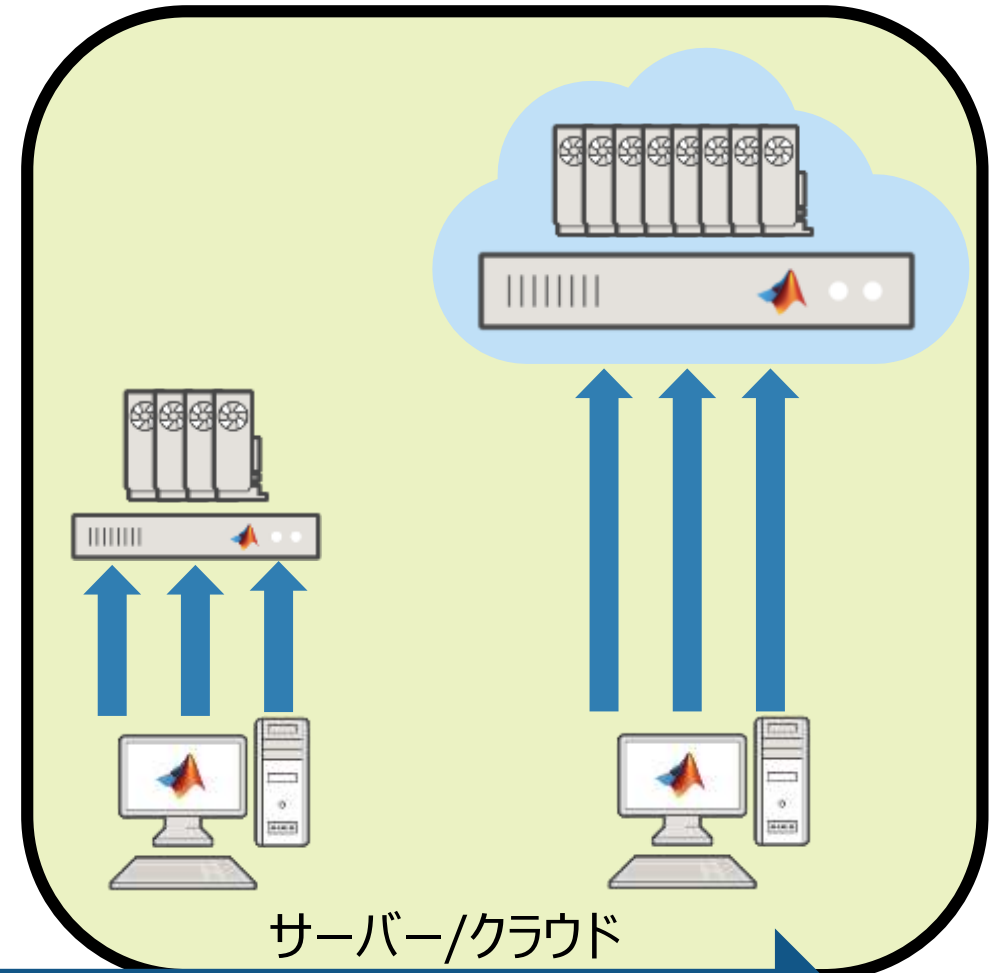
CPUのみ



GPU搭載



複数GPU

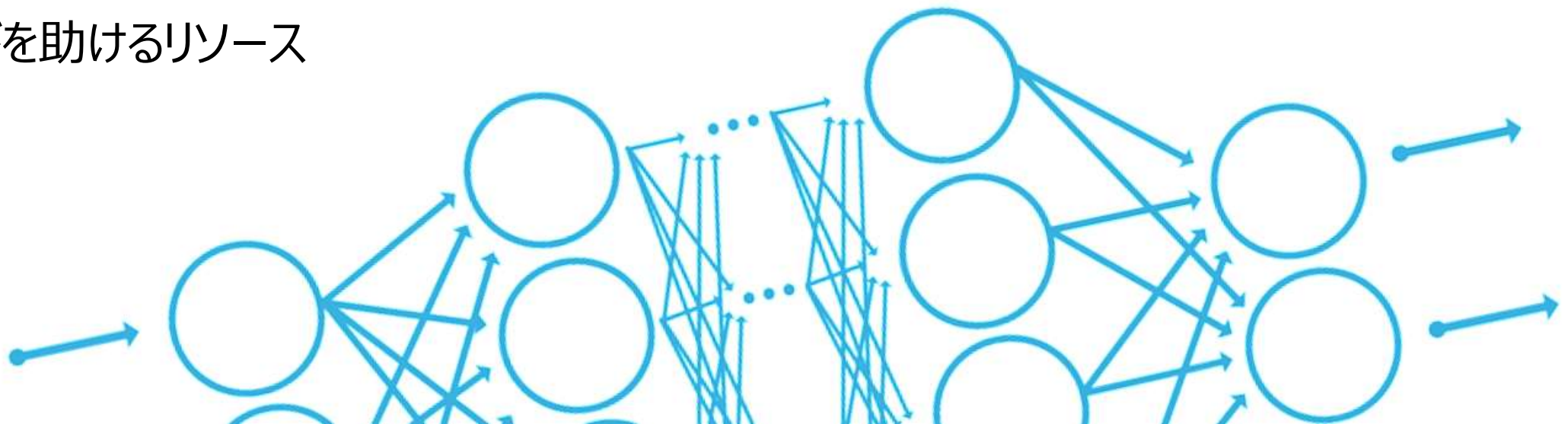


サーバー/クラウド

More GPUs

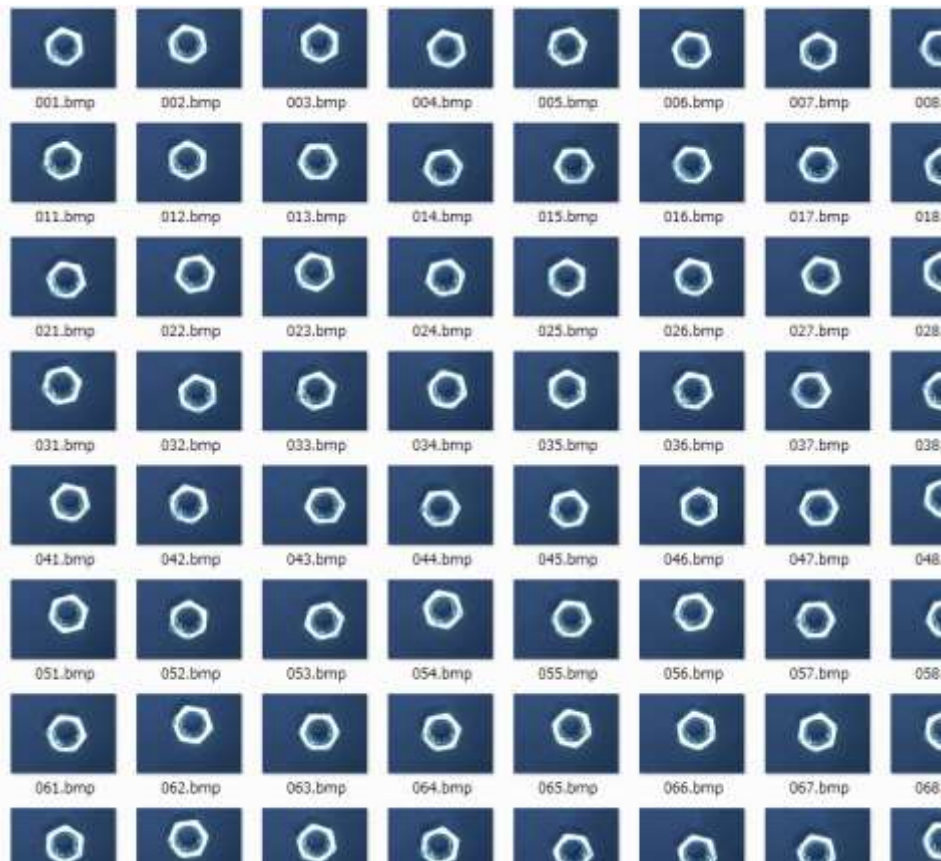
Agenda

- これからはじめるディープラーニング活用ワークフロー
- ディープラーニングの異常検知への適用
- 物体検出と領域の切り出し
- 学びを助けるリソース



【例題】特徴量を使った異常検出

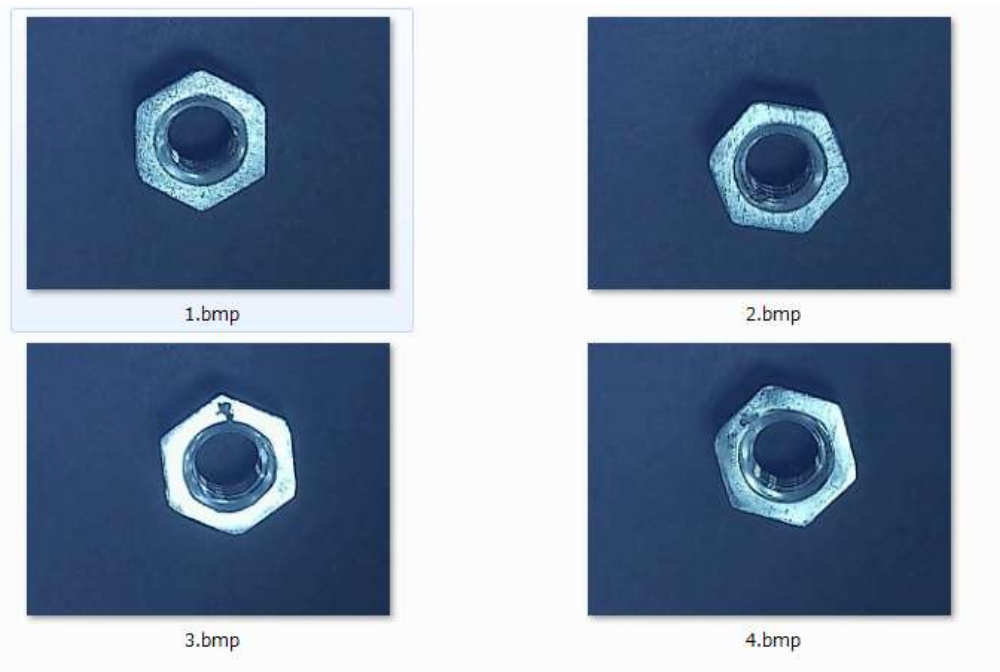
100個のナット



一般的に
正常なモノに対して異常なモノが少ない

デモ

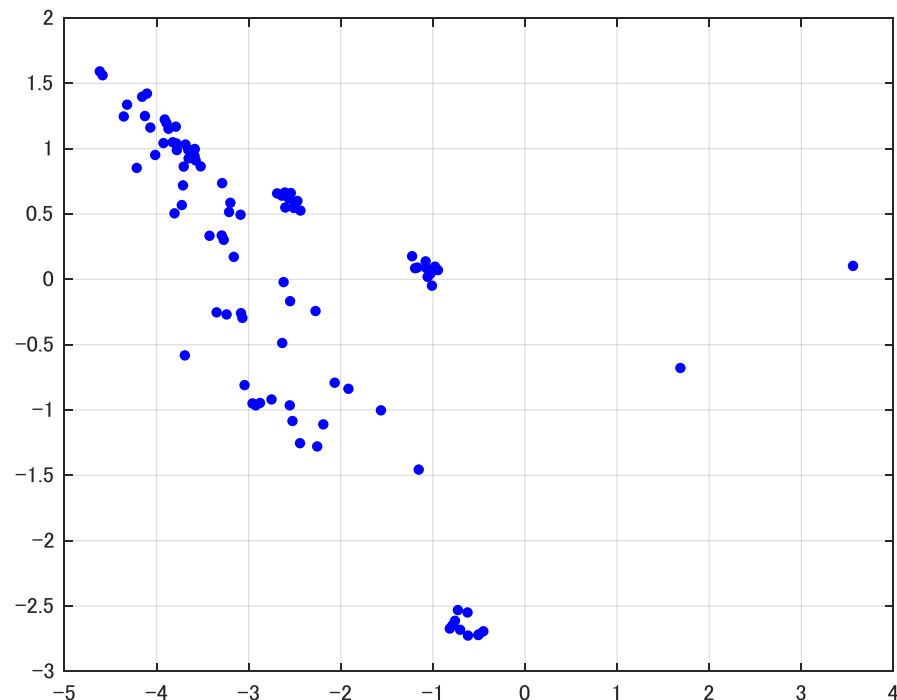
異常



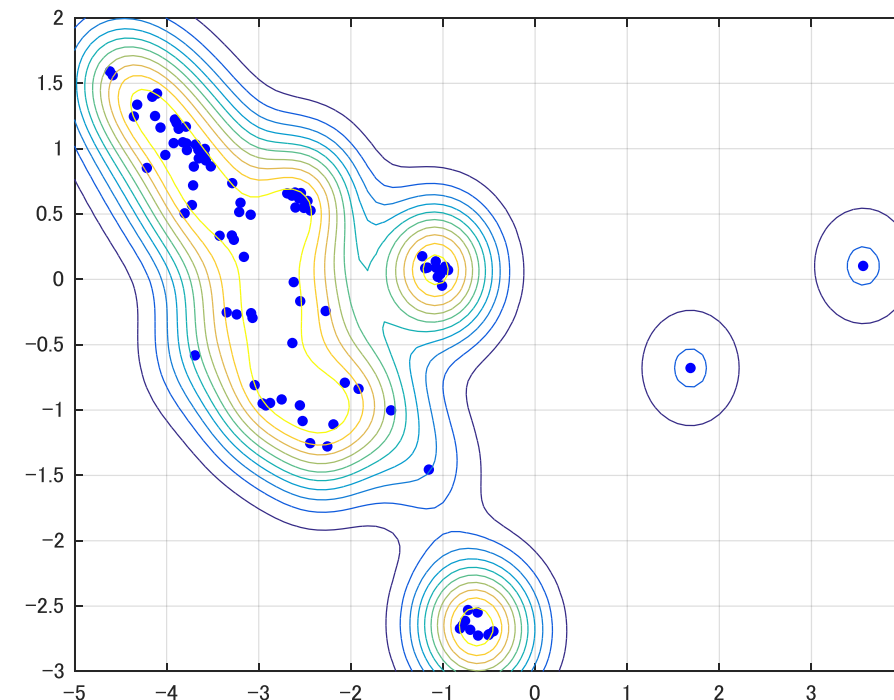
100個の中に紛れ込んでいる4個の異常を見つけられるか

1-Class SVM とは？

カーネル法を使った正常度スコアの推定アルゴリズム



異常データを含むデータの例

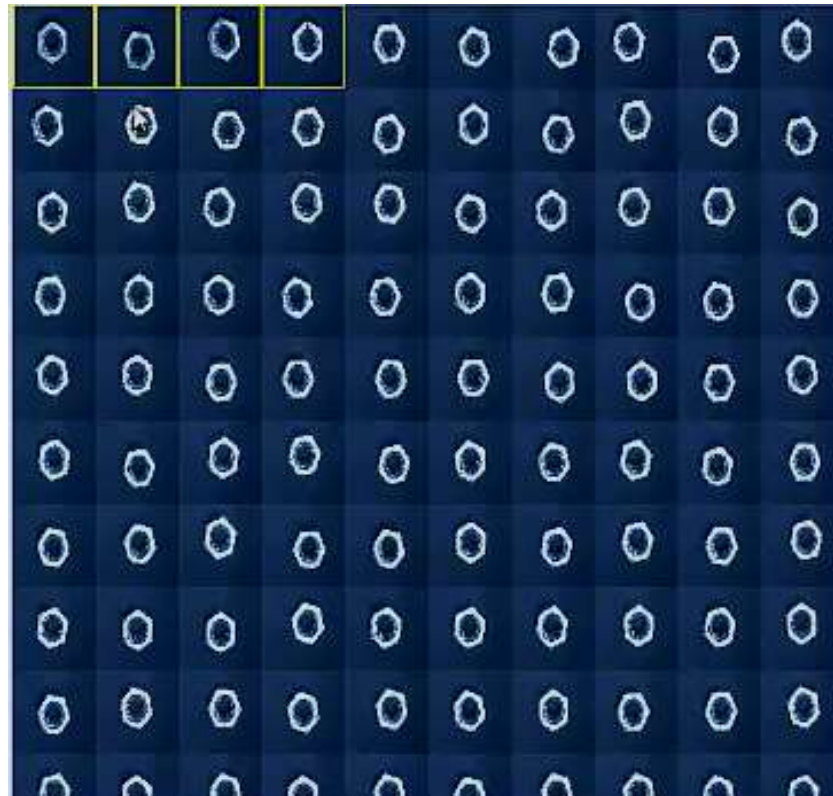


スコアの値が高い方が正常度が高い

1-Class SVM

【例題】特徴量を使った異常検出

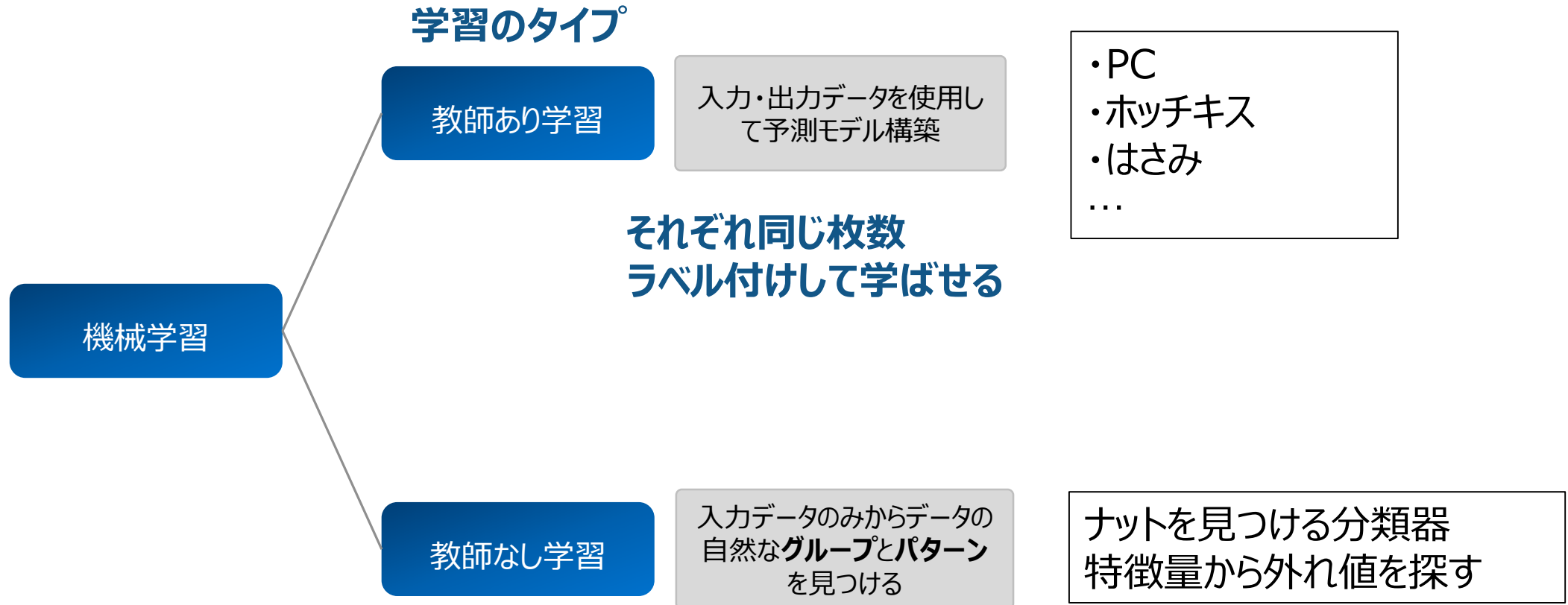
異常度：高



人が下した判定と
ディープラーニングの判定
が一致

異常度：低

取り扱った二つの学習タイプ



正常(高スコア) 画像の加工と評価

行列を反転させる関数： $B = \text{fliplr}(A)$

正常元画像



9.30 (1位)

3.53 (75位)

左右反転 \Rightarrow スコア悪化

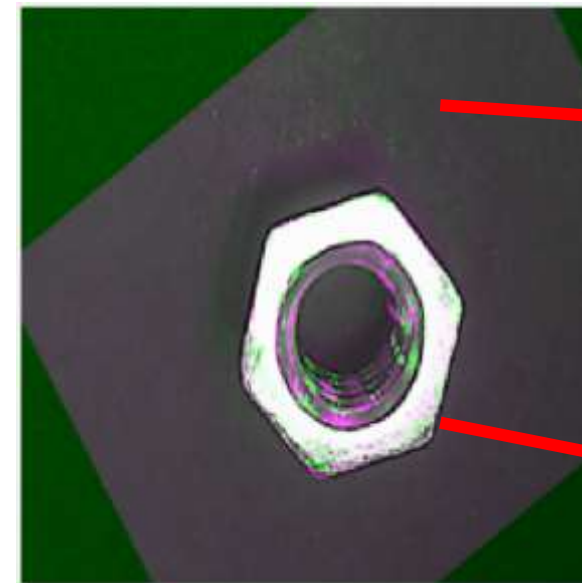
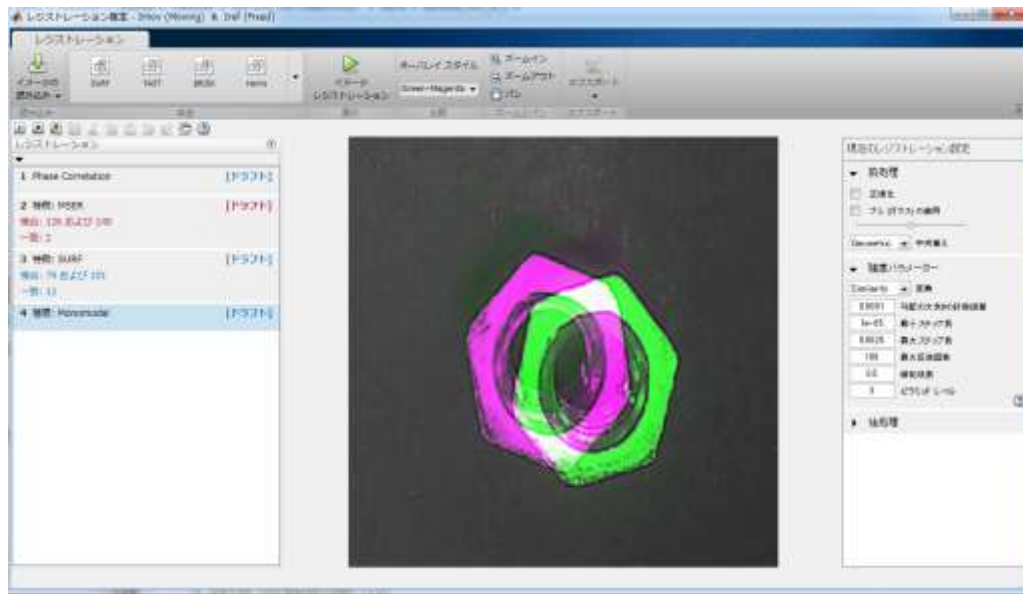
位置関係や輝度グラデーションなどもスコアに大きく影響がありそう

解析結果を元に前処理をおこなう

表面のキズに注目したい。

解析からナットの位置や背景の輝度分布などもスコアに影響している。

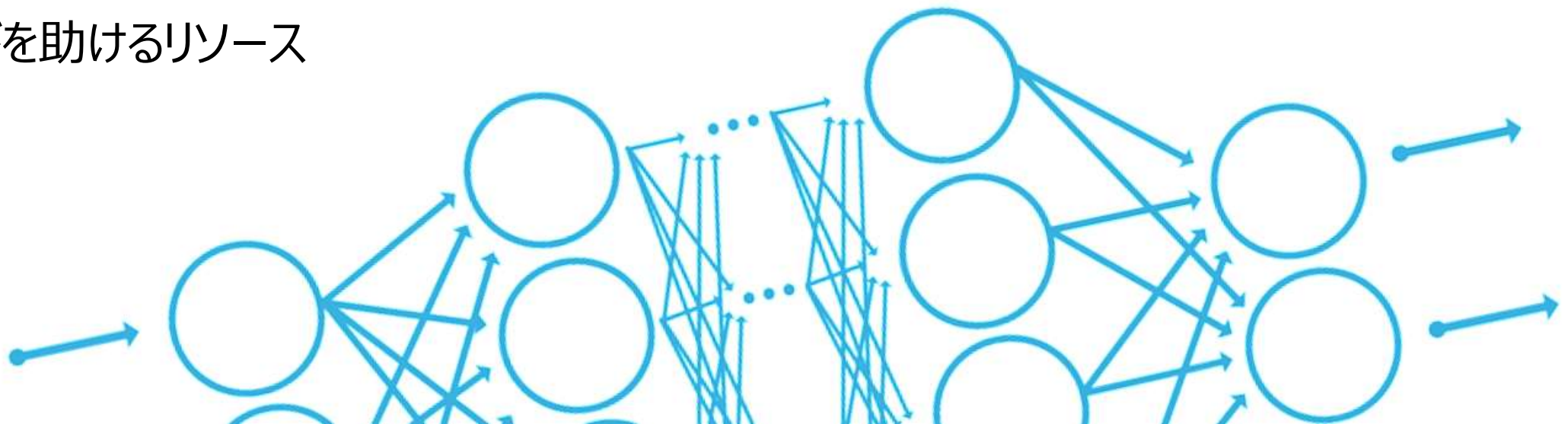
⇒輝度のレジストレーションと幾何学変換により、位置合わせ。
ナット以外のところはマスク。



- ・画像の前処理により精度が向上
- ・

Agenda

- これからはじめるディープラーニング活用ワークフロー
- ディープラーニングの異常検知への適用
- 物体検出と領域の切り出し
- 学びを助けるリソース



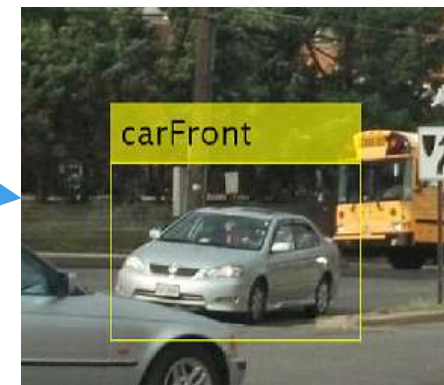
物体の検出

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

R2017a



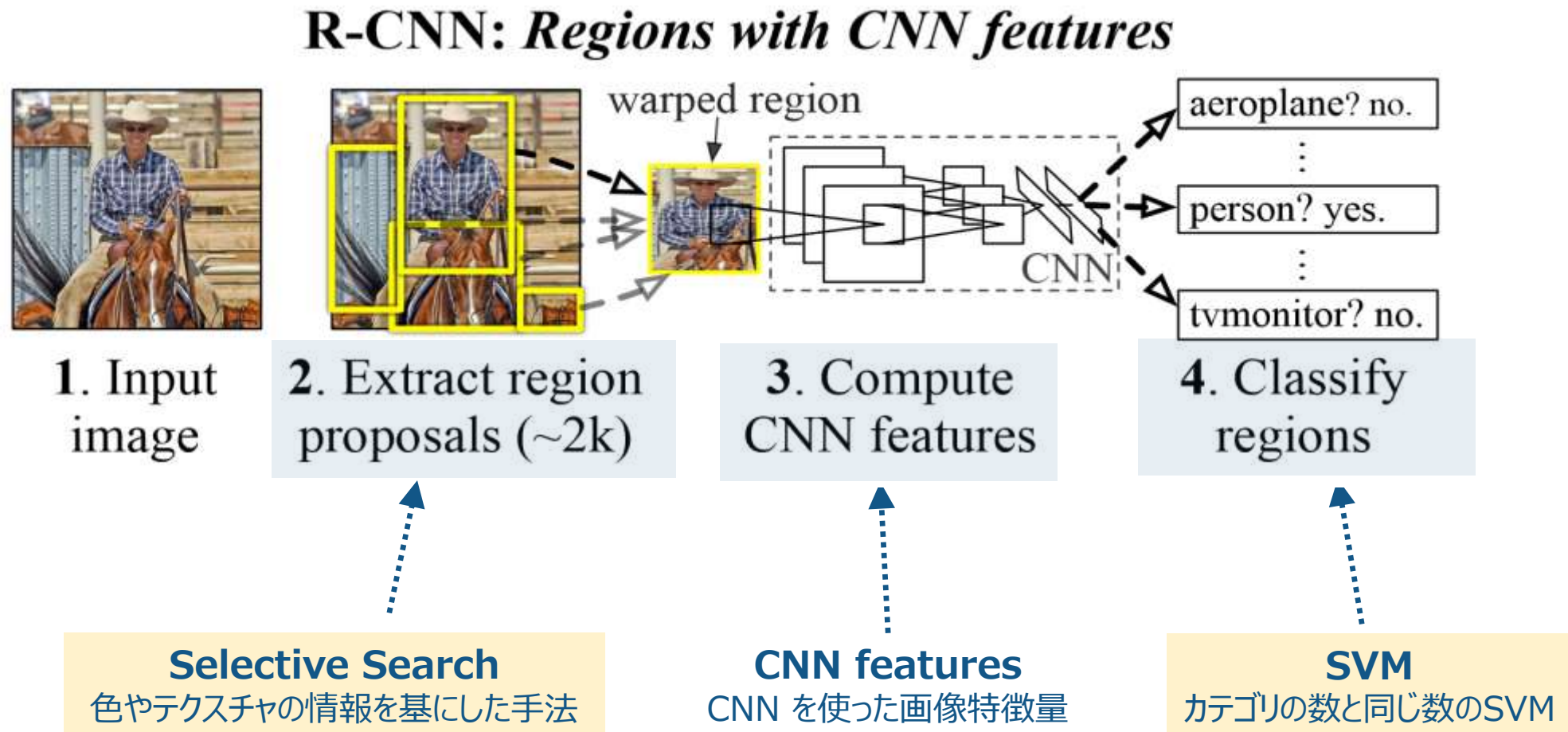
停止標識 (Stop Sign)



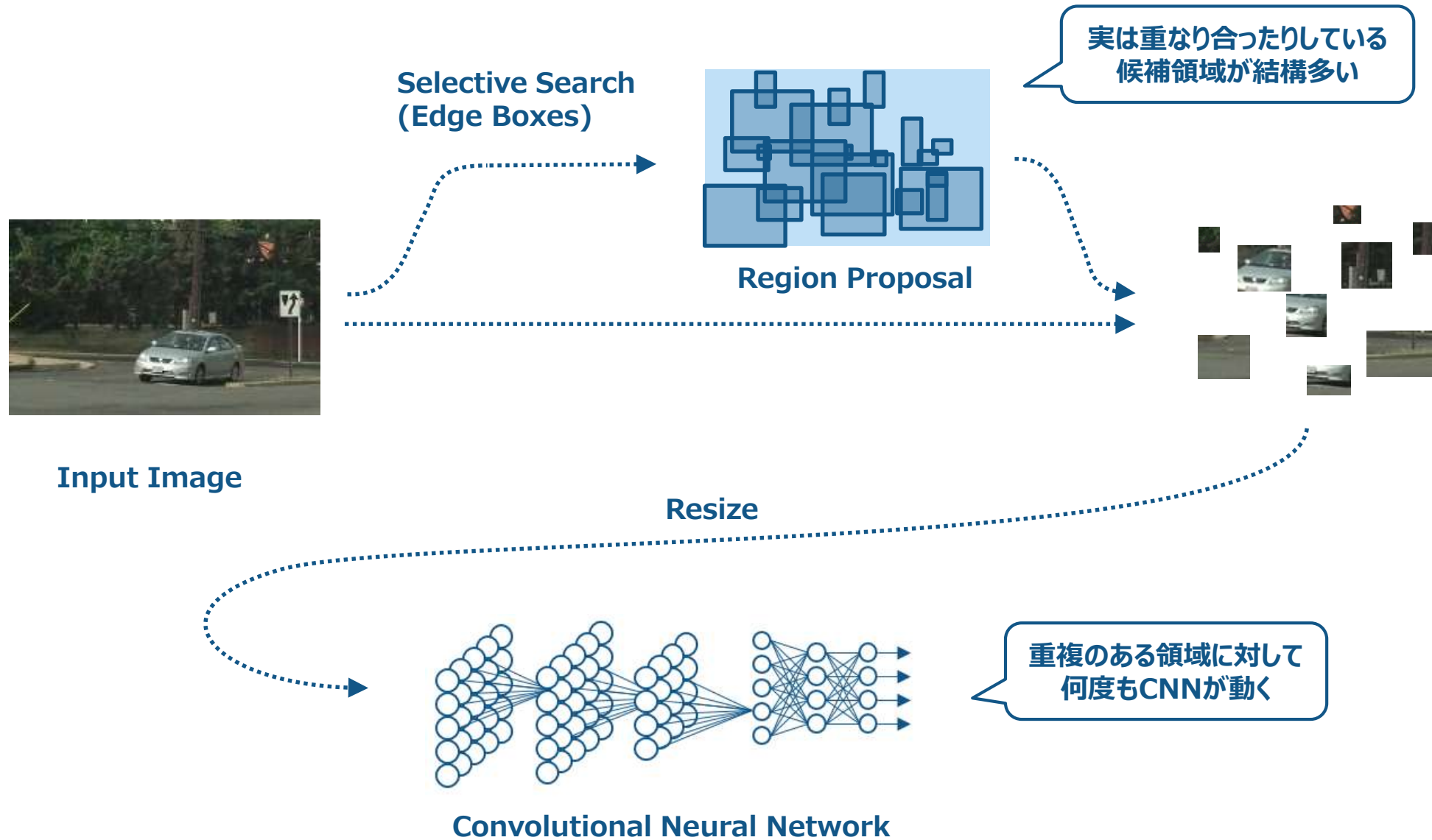
自動車の前面 (Car Front)

R-CNNに自動車の前面と停止標識を学習させた場合の検出例

R-CNN (Regions with CNN features) とは？

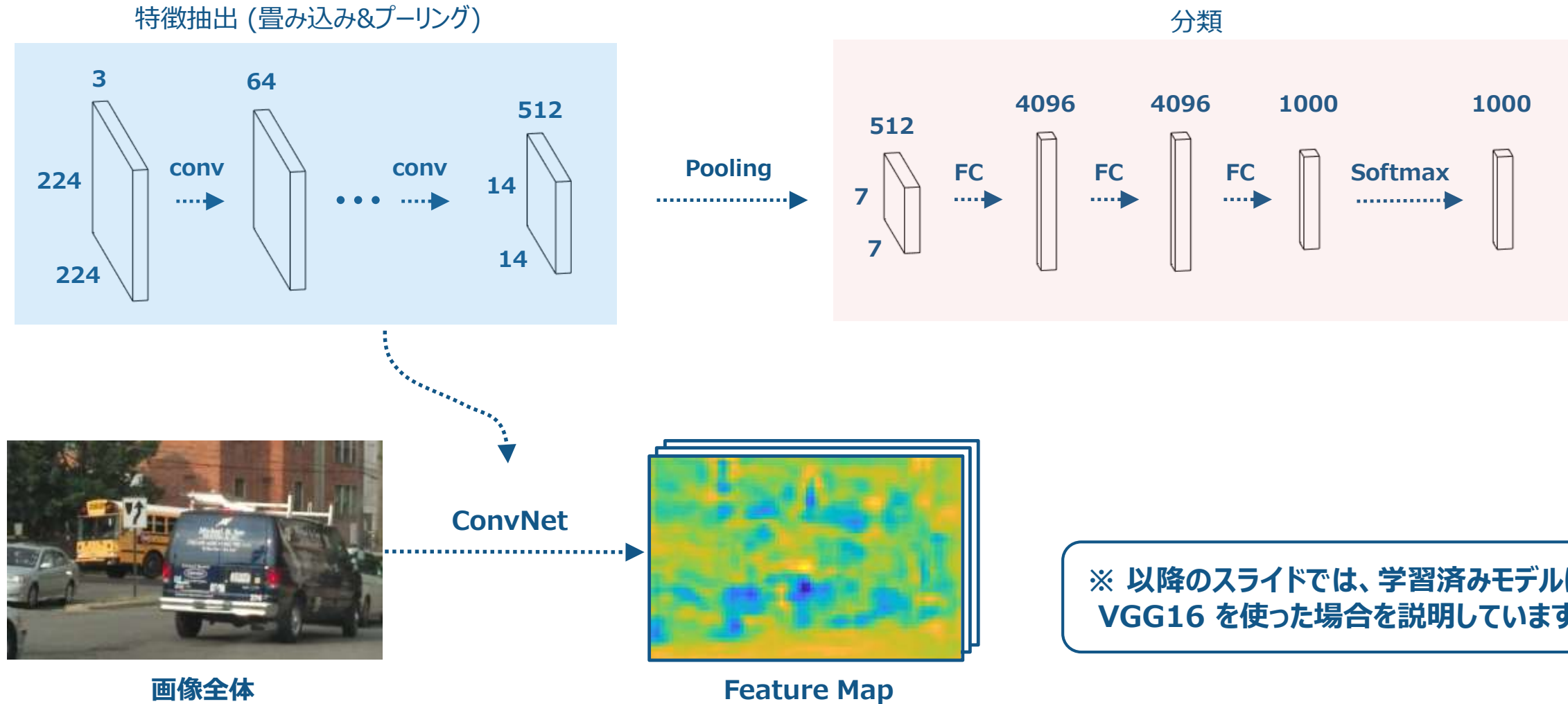


R-CNN はなぜ遅くなってしまうのか？



Fast R-CNN とは？

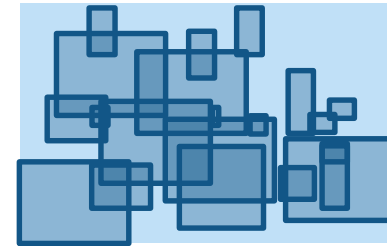
Step 1) 画像全体に CNN の前半部分（畳み込み&プーリング）を実行して、Feature Map を生成する



Fast R-CNN とは？

実はRegion Proposal の生成に
時間がかかる

Selective Search
(Edge Boxes)

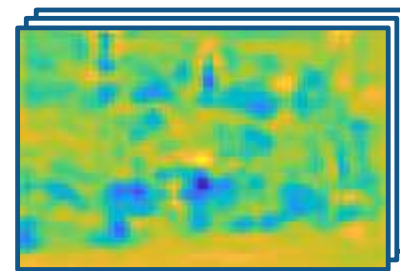


Region Proposal



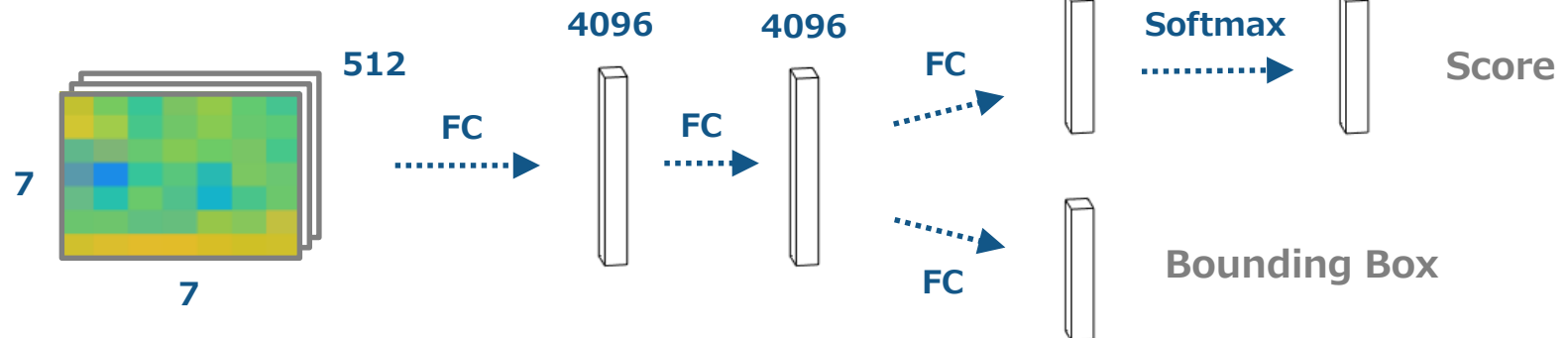
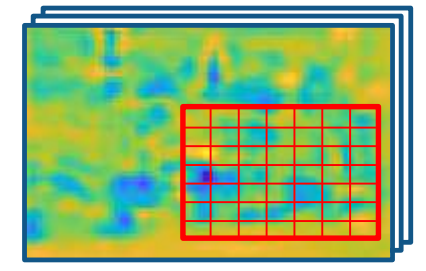
Input Image

ConvNet

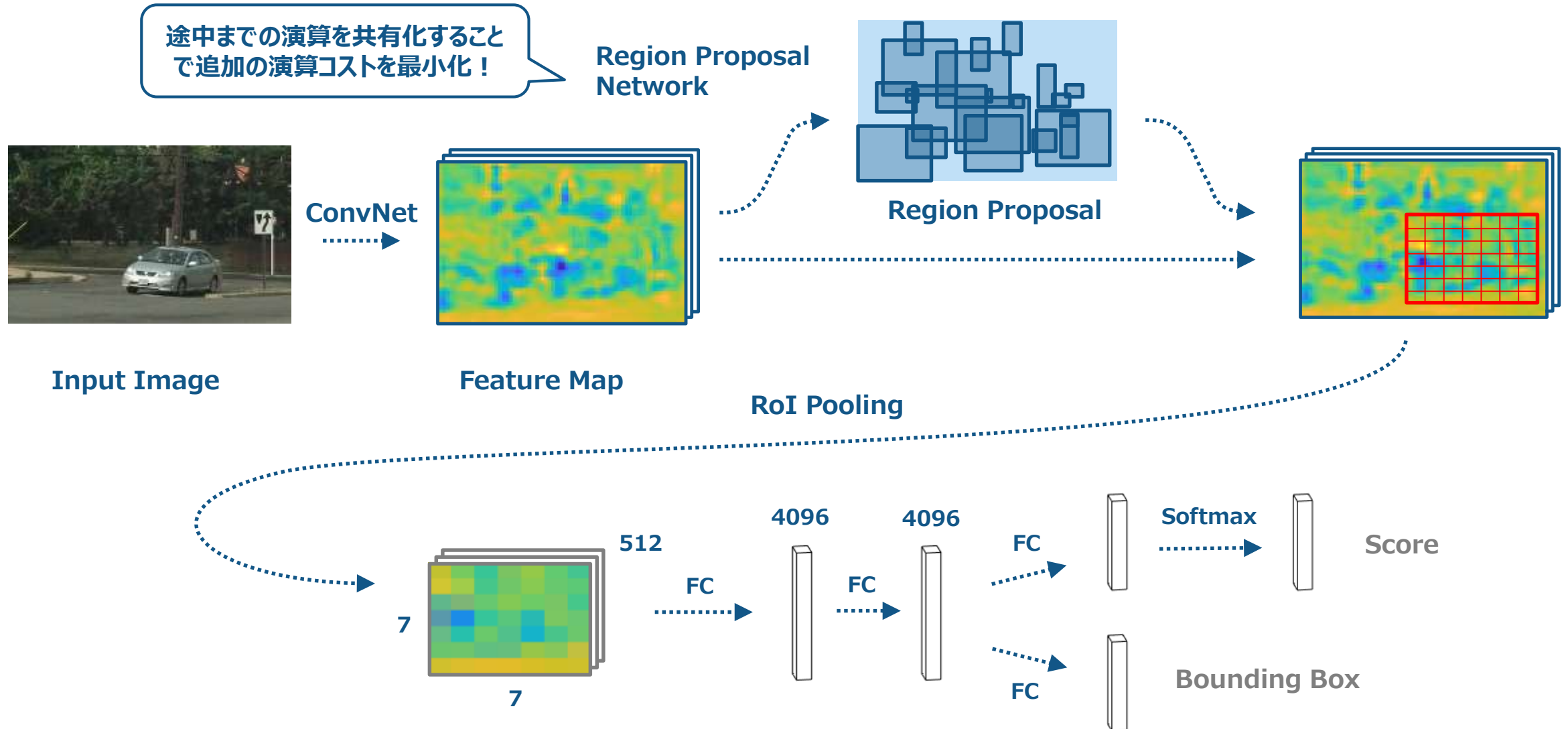


Feature Map

RoI Pooling



Faster R-CNN とは？



R-CNN / Fast R-CNN / Faster R-CNN の選び方

	R-CNN	Fast R-CNN	Faster R-CNN
認識・検出の速度	× (遅い)	△ (割と速い)	○ (速い)
必要なGPUメモリ	○ (少なめ)	× (多い)	× (多い)
小さな物体の認識	○ (得意)	× (不得意)	× (不得意)
カスタムの領域候補	○ (可)	○ (可)	× (不可)
学習のさせ易さ	○ (簡単)	○ (簡単)	× (難しい)
学習に必要な時間	○ (短め)	○ (短め)	× (長い)

画像全体の Feature Map を生成するため大量の GPU メモリを消費しやすい

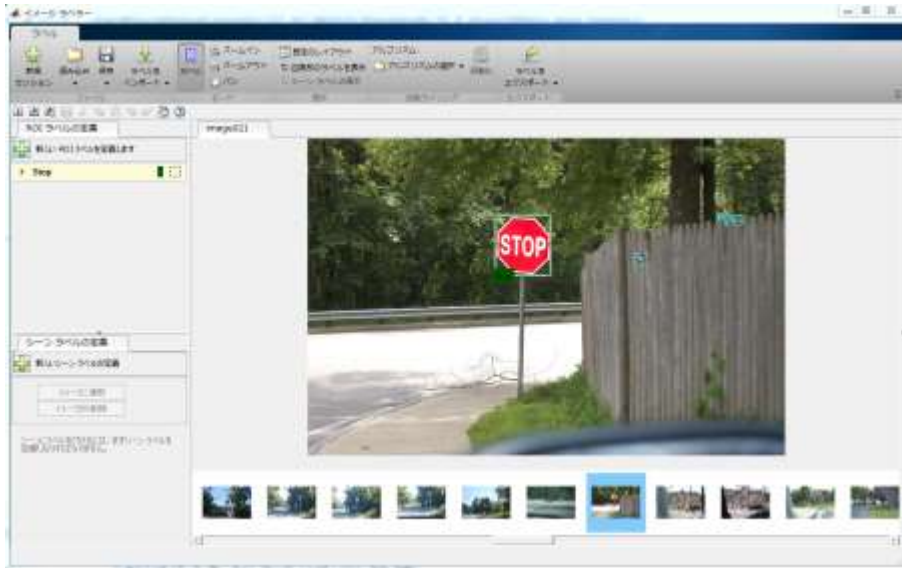
4 回の学習が必要であり、時間がかかる
うまく収束させるにはコツがいる

切り出した領域をリサイズする操作が入るため
小さな領域では拡大が行われる

画像のラベル付けをアプリでサポート

Image Labeler

ROIのラベリング



Ground Truth Labeler

動画のROIラベリング



Automated Driving System Toolbox™

- シーン、オブジェクト、ピクセルに対してのラベリング
- 独自のアルゴリズムによる自動化

3種類の学習方法の使い分け

R-CNN

```
detector = trainRCNNObjectDetector(groundTruth, network, options)
```

Fast R-CNN

```
detector = trainFastRCNNObjectDetector(data, layers , options)
```

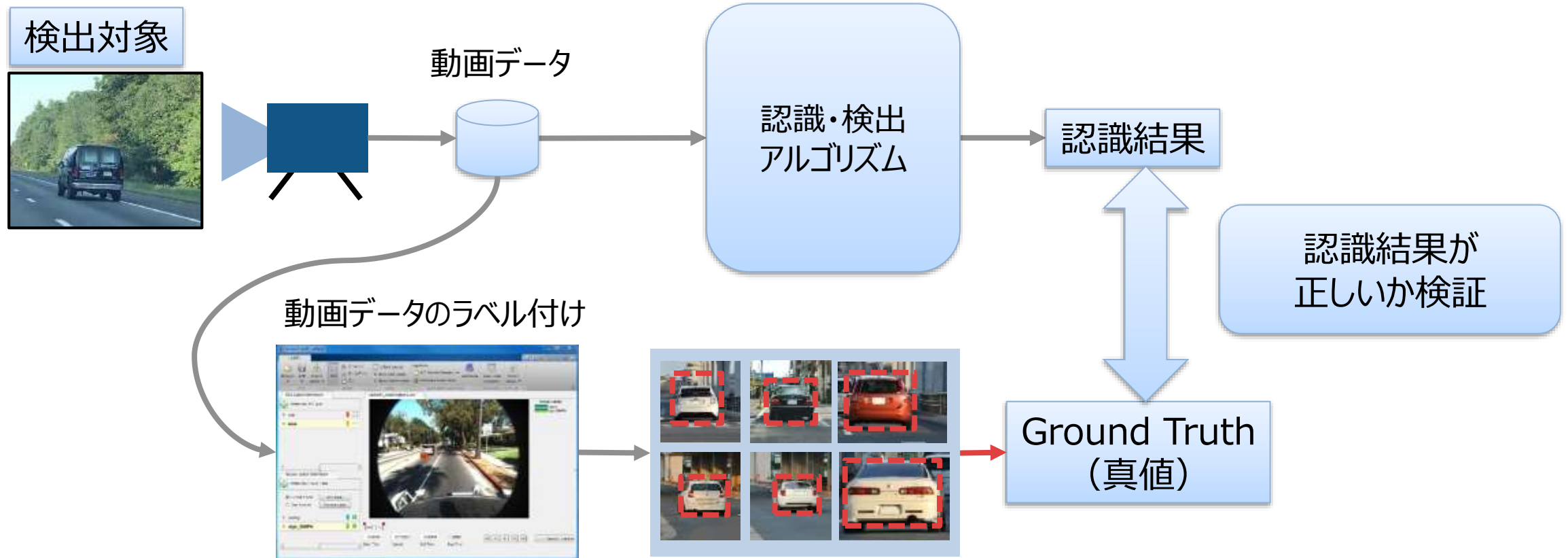
Faster R-CNN

```
detector = trainFasterRCNNObjectDetector(data, layers, options)
```

わずかな書き換えで3種類を試すことができます。

機械学習ワークフロー⑧：認識器の性能検証

Computer Vision System Toolbox™
Automated Driving System Toolbox™



`evaluateDetectionPrecision`

検出精度の算出

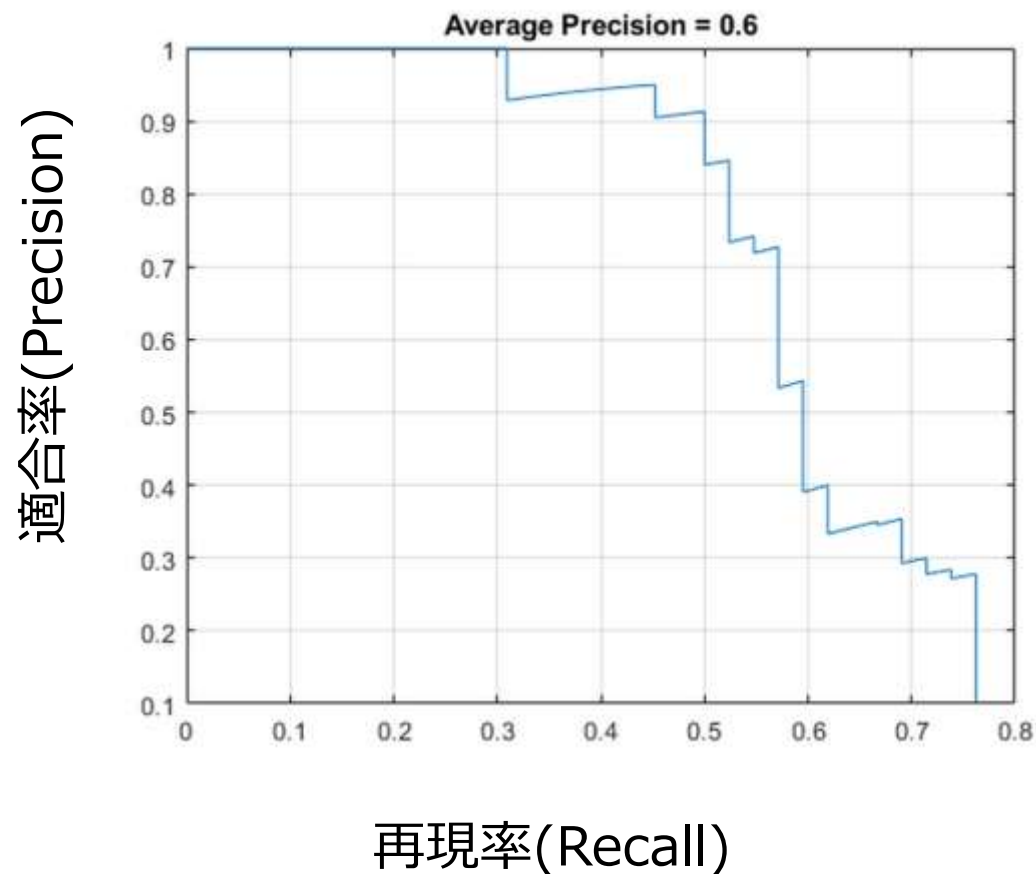
R2017a

`evaluateDetectionMissRate`

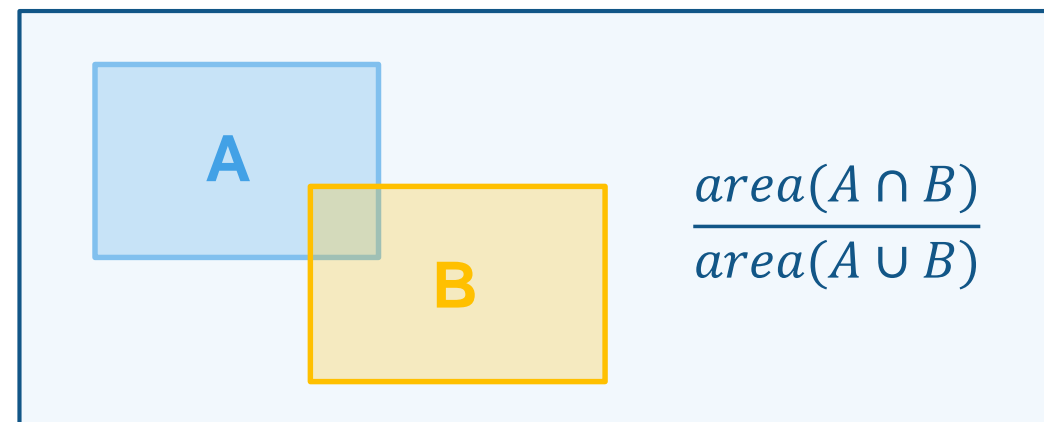
誤検出率の算出

R2017a

機械学習ワークフロー⑧：認識器の性能検証



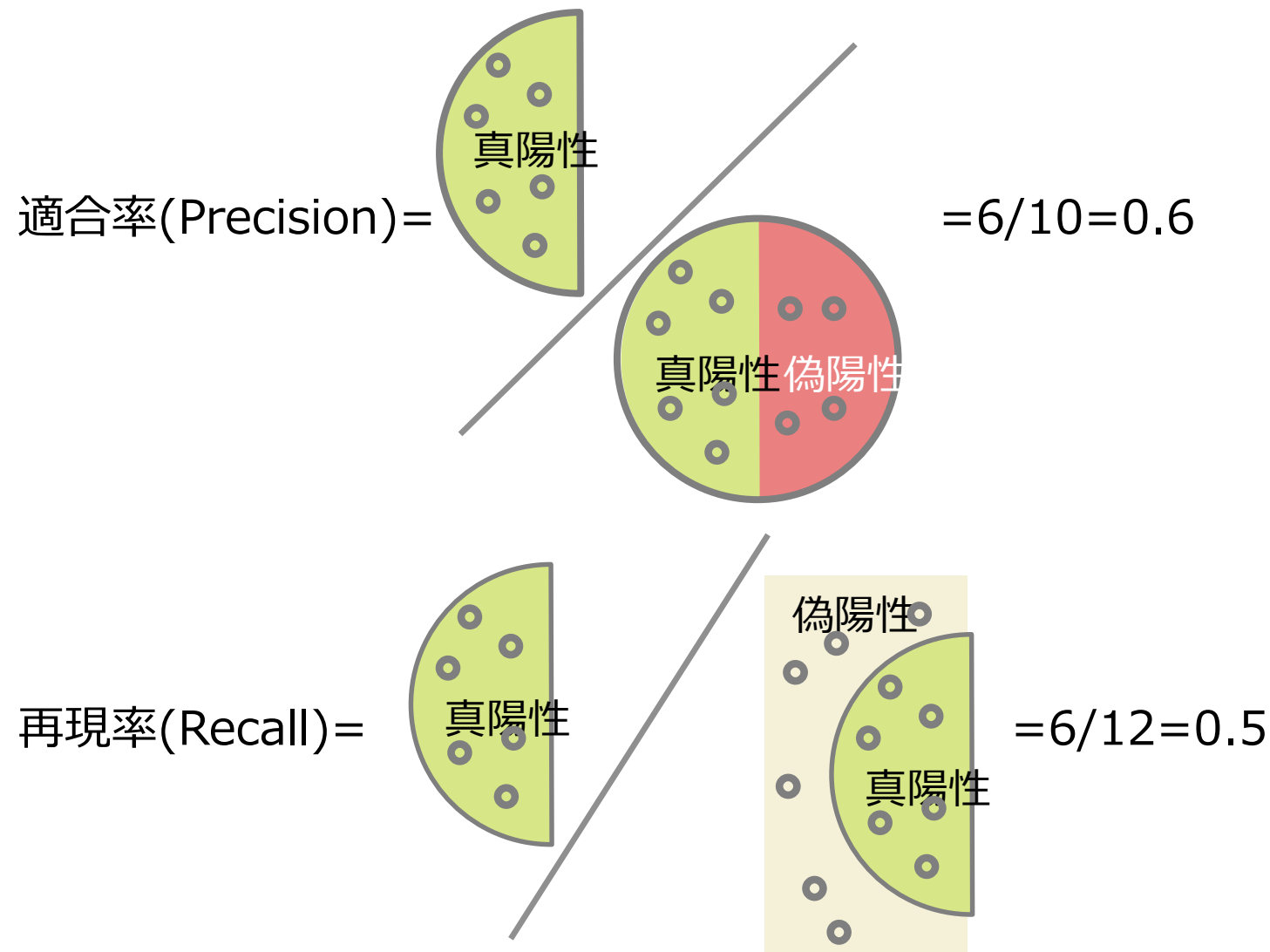
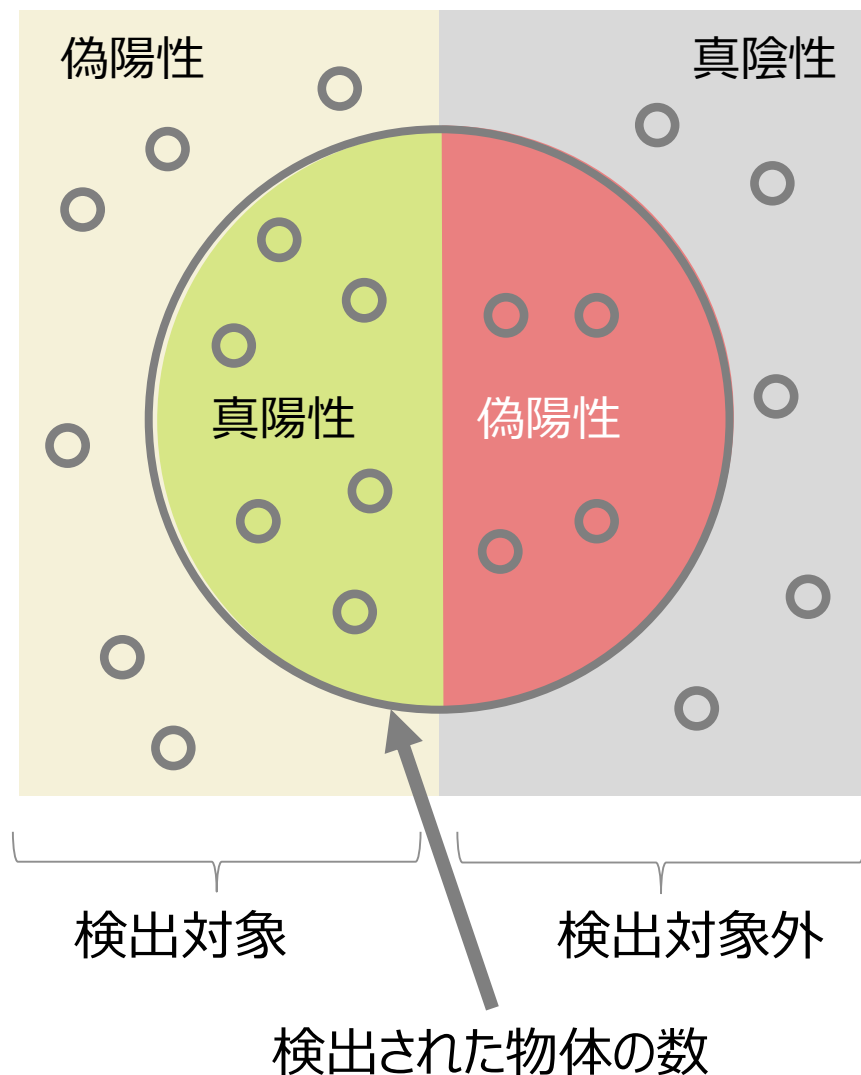
IoU (Intersection over Union)



理想は再現率が大きくなっても、
適合率が落ちないこと

機械学習ワークフロー⑧：適合率・再現率による性能評価

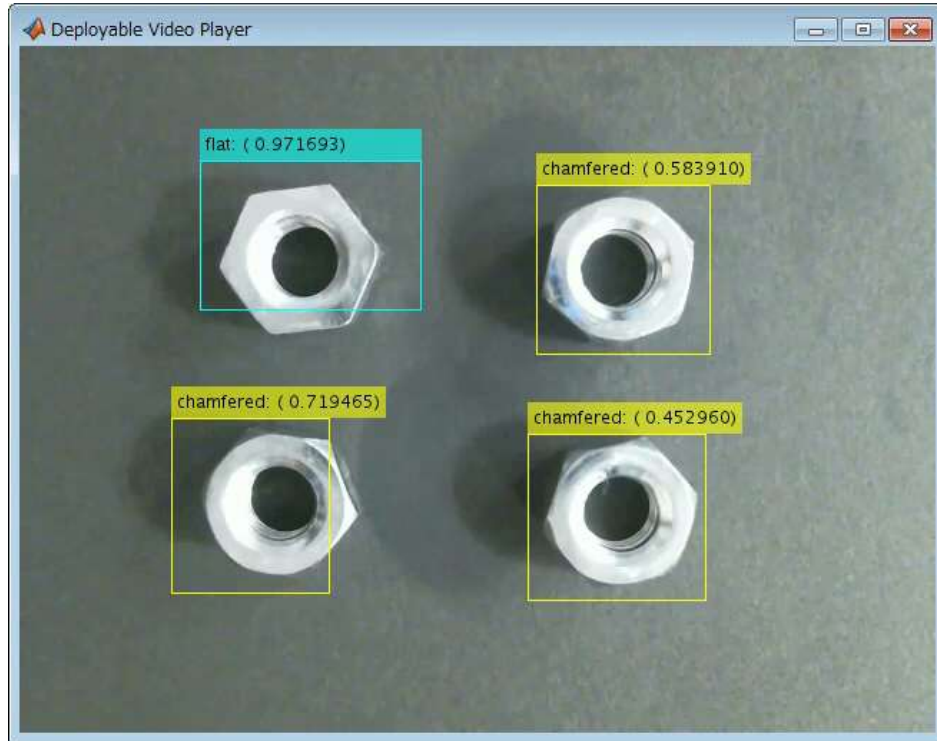
Computer Vision System Toolbox™



どちらも1に近いほうが優秀な検出器

Faster R-CNNによる物体検出

Faster R-CNNにより高速に物体
の位置と状態を判断



物体検出を用いたロボット制御



物体識別（ピクセル毎）

Semantic Segmentation (SegNet)

Semantic Segmentation

畳み込みニューラルネットによるセグメンテーション

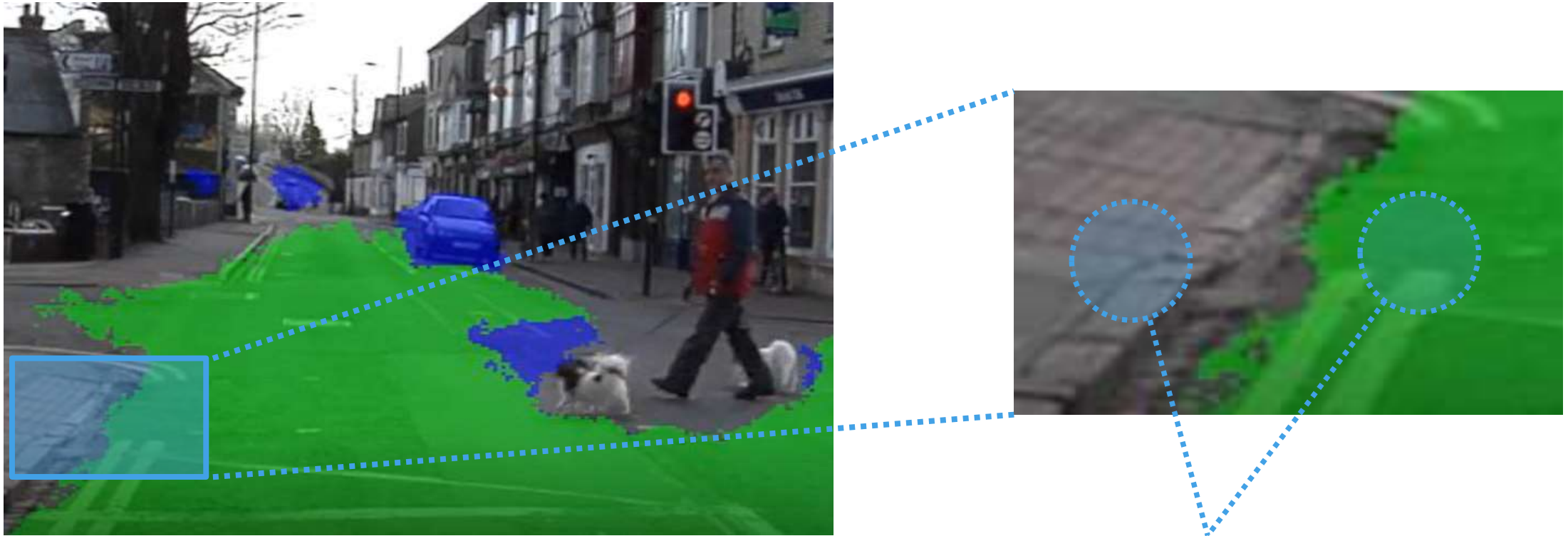
R2017b



Semantic Segmentation とは？

R2017b

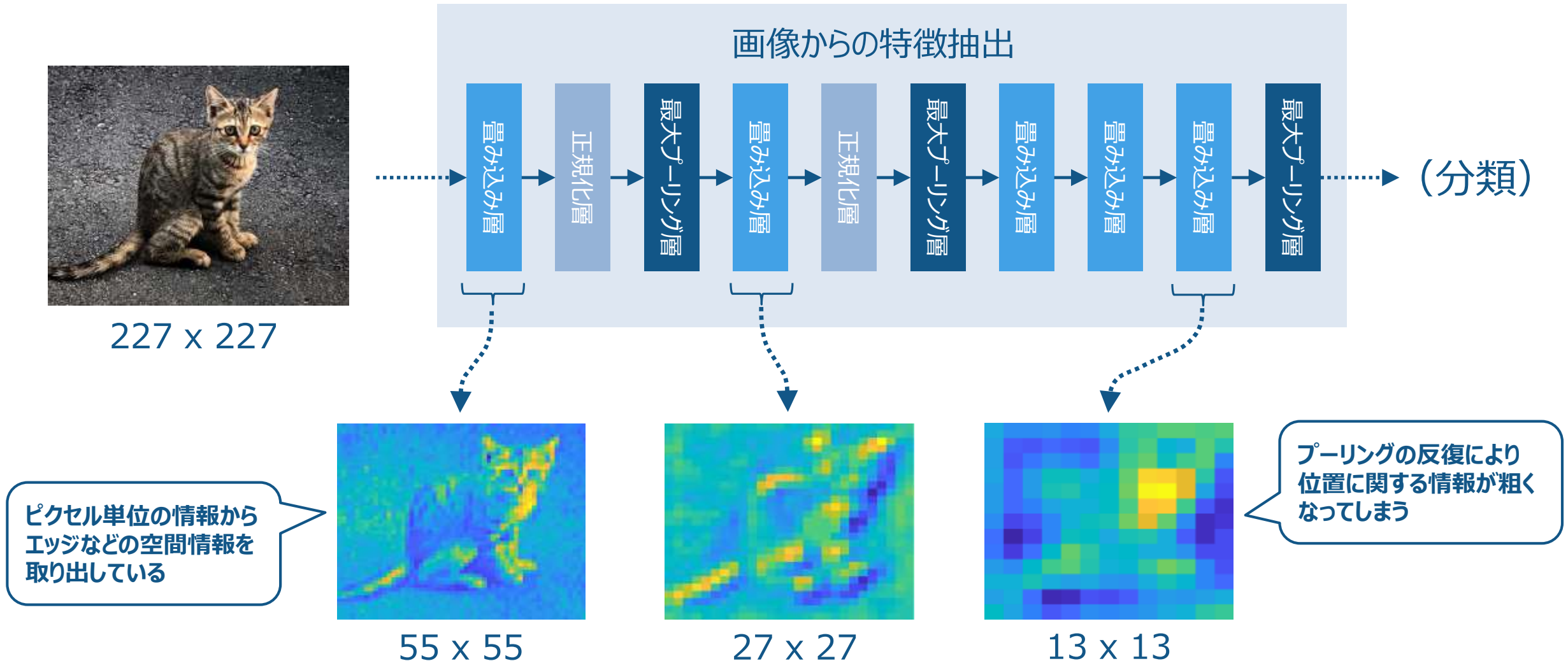
各ピクセルをその意味（周辺のピクセルの情報）に基づいて、カテゴリ分類する手法



ちゃんと歩道と車道を区別できている！
色だけを見ているわけではない

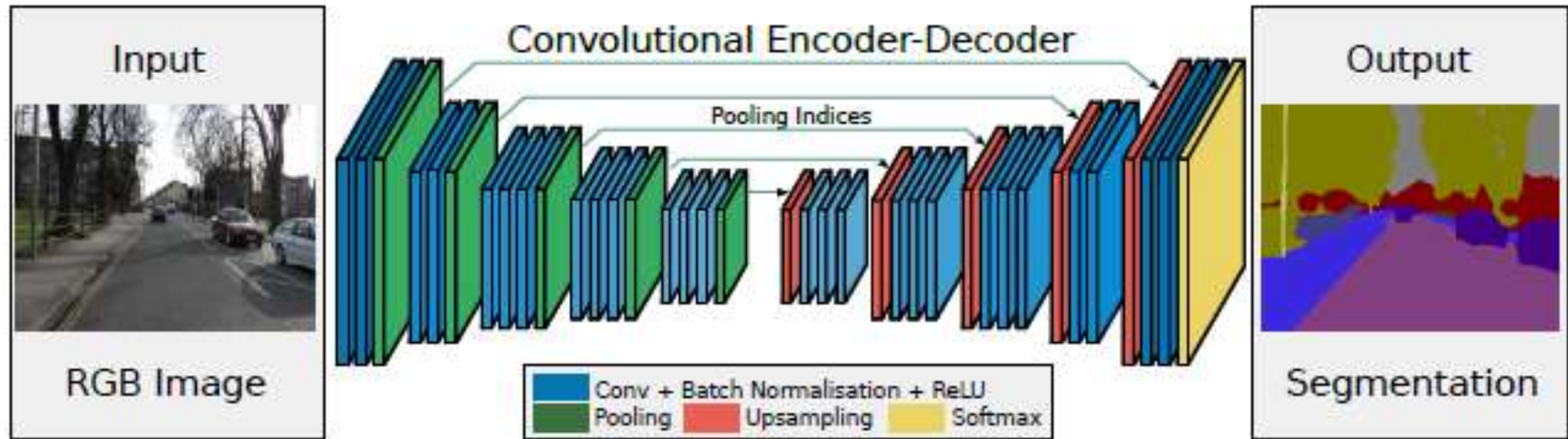
Feature Map は位置情報が粗くなる

畳み込みニューラルネットワークの計算過程で出てくる畳み込みの出力



SegNet (Semantic Segmentation)

Max Pooling時のIndexを転送して
位置に関する情報を補充している



Badrinarayanan, V., A. Kendall, and R. Cipolla. "Segnet: A deep convolutional encoder-decoder architecture for image segmentation." arXiv. Preprint arXiv: 1511.0051, 2015.

逆プーリング (Unpooling)

2	5	3	9
4	8	4	8
3	7	5	4
5	6	3	6

Max Pooling
Indices



0	0	0	9
0	8	0	0
0	7	0	0
0	0	0	6

最大プーリング
(Max Pooling)



8	9
7	6

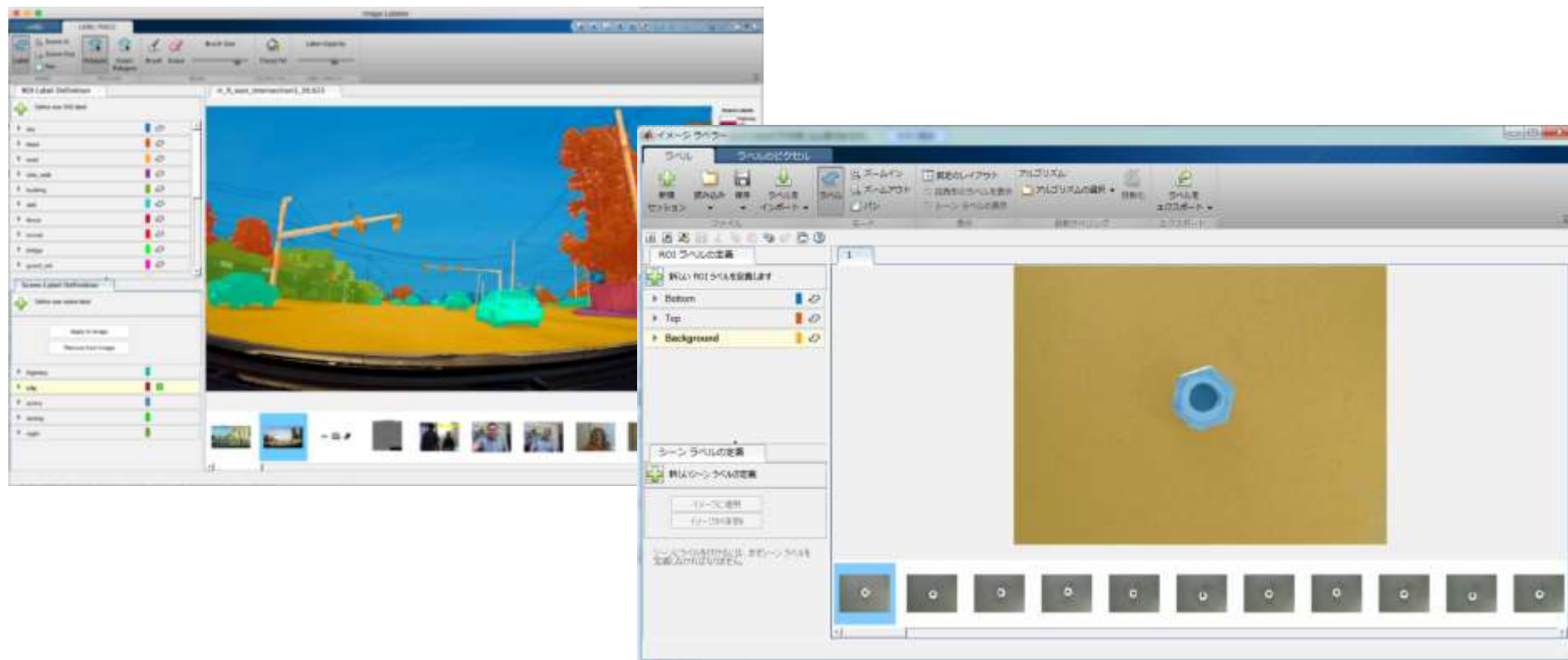


8	9
7	6

逆プーリング
(Unpooling)



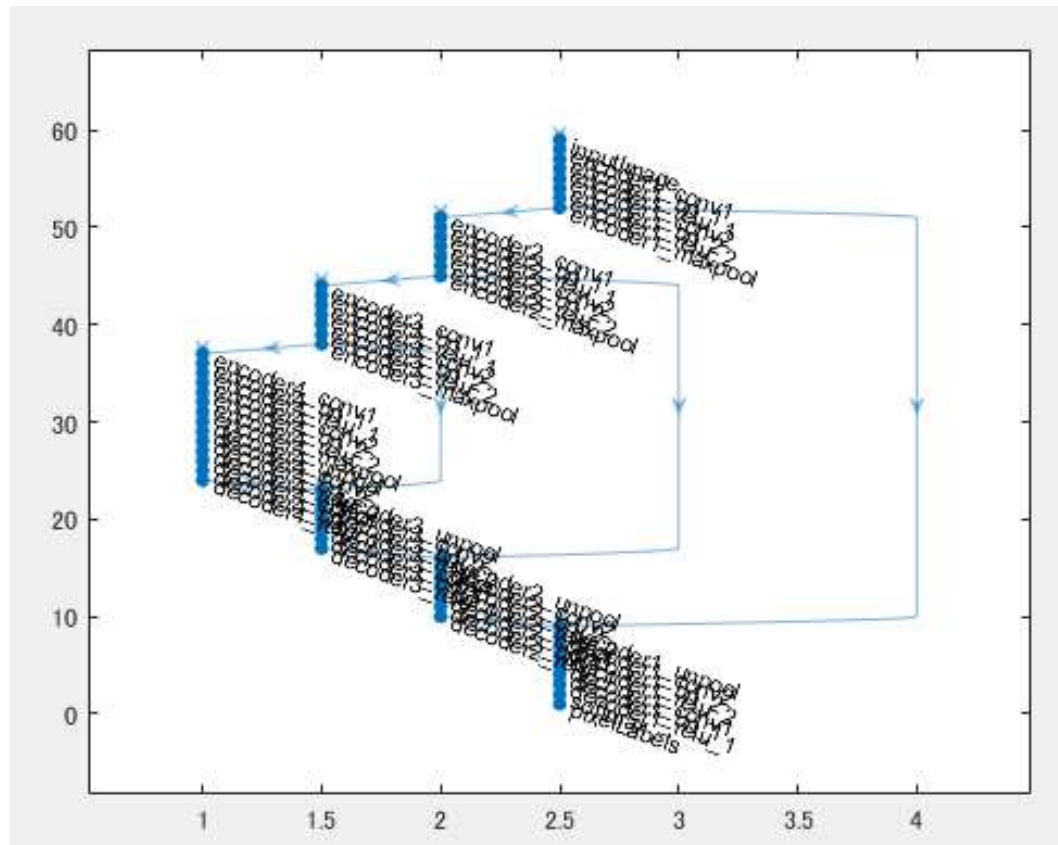
ピクセルに対するラベリング



Segnetの層の定義

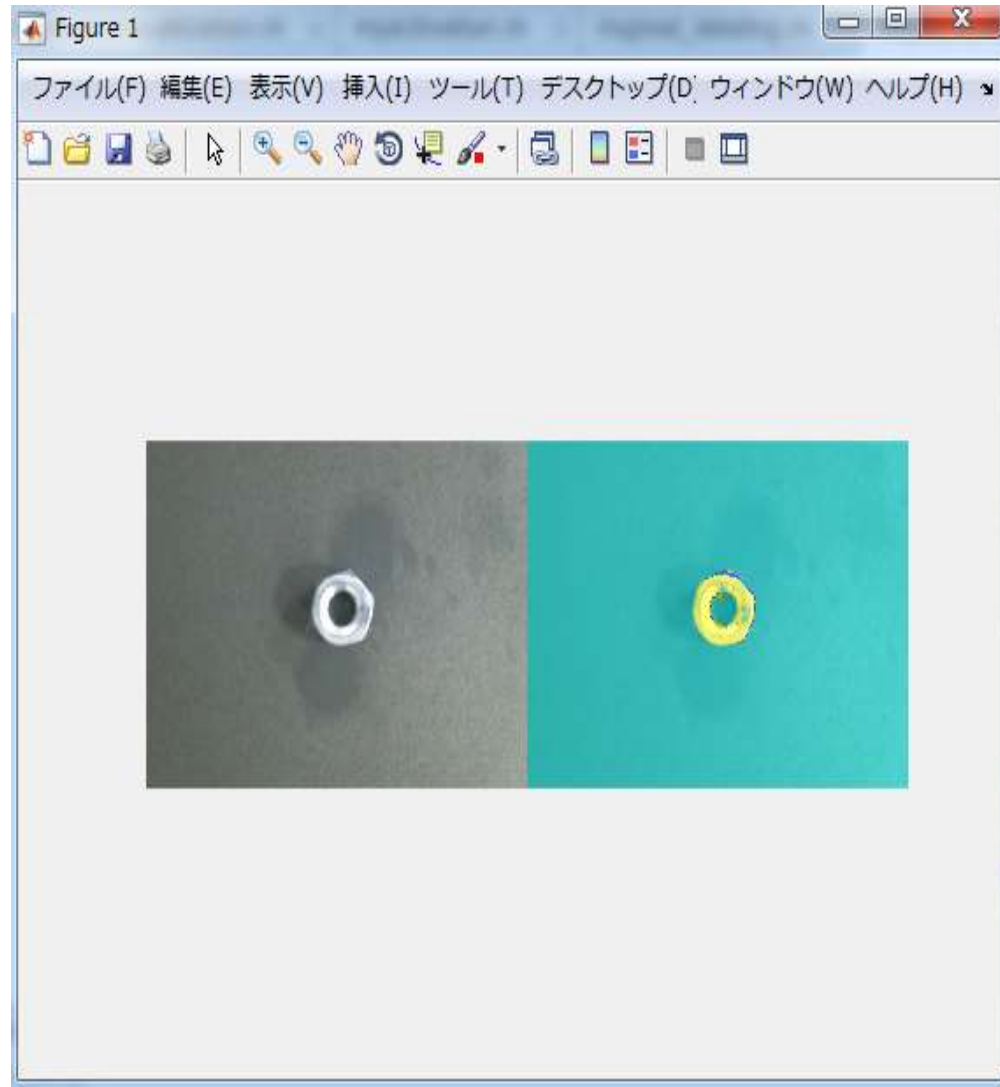
```
lgraph = segnetLayers(imageSize, numClasses, encoderDepth)
```

```
figure  
plot(lgraph)
```



- プーリングによるダウンサンプルの 픽셀位置情報を保持し、アップサンプリングに利用

セマンティックセグメンテーションによる領域の抽出と評価



Ground
Truth

Result



評価用の関数 `evaluateSemanticSegmentation`

- 精度
- IoU
- 平均BFスコア

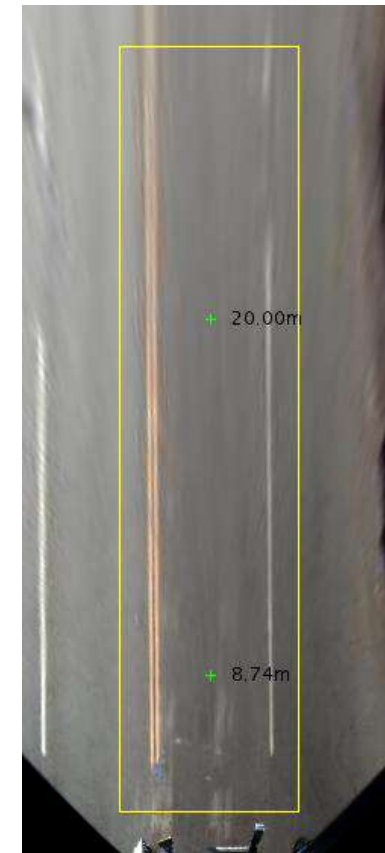
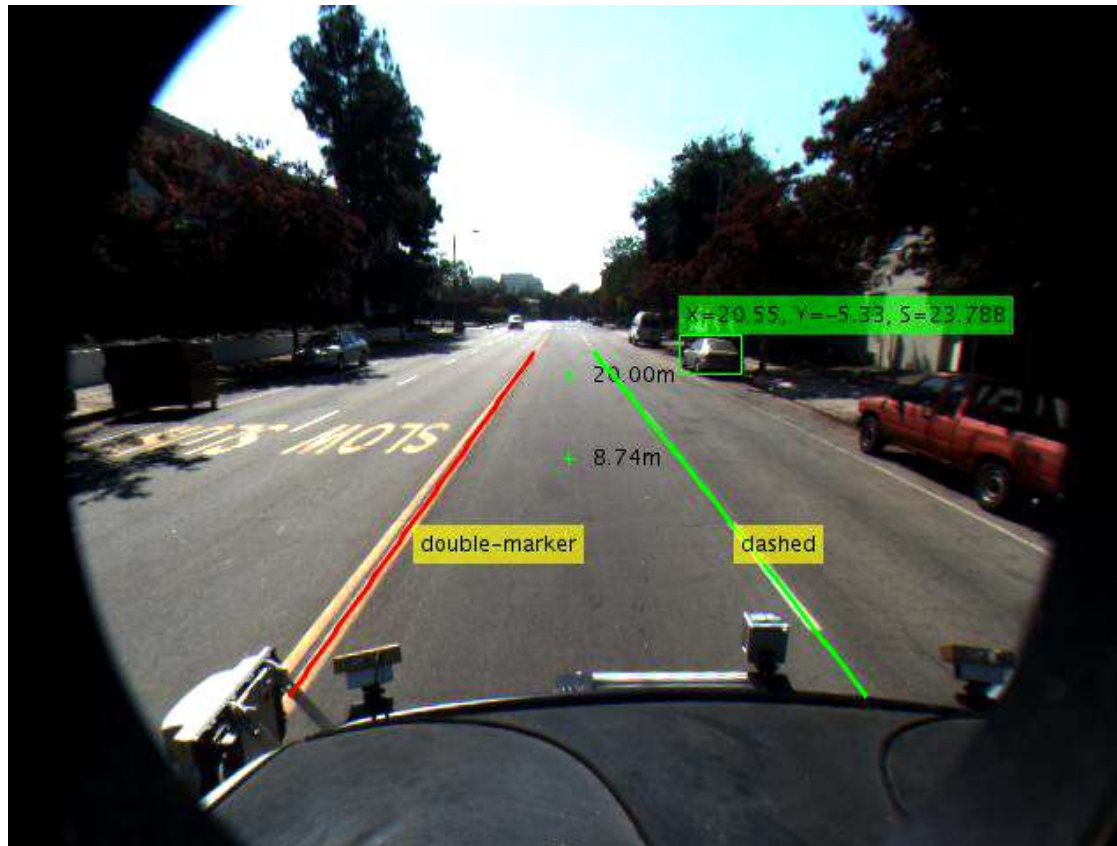
その他のディープラーニング適用

CNNの回帰

R2017a

Automated Driving System Toolbox™

- 車載カメラの画像処理・鳥瞰図(バードビュー)変換
- 前方車両の認識や白線認識のアルゴリズム提供



ディープニューラルネットワークによるノイズ除去

R2017b

```
net = denoisingNetwork('DnCNN');  
denoisedI = denoiseImage(noisyI, net);
```

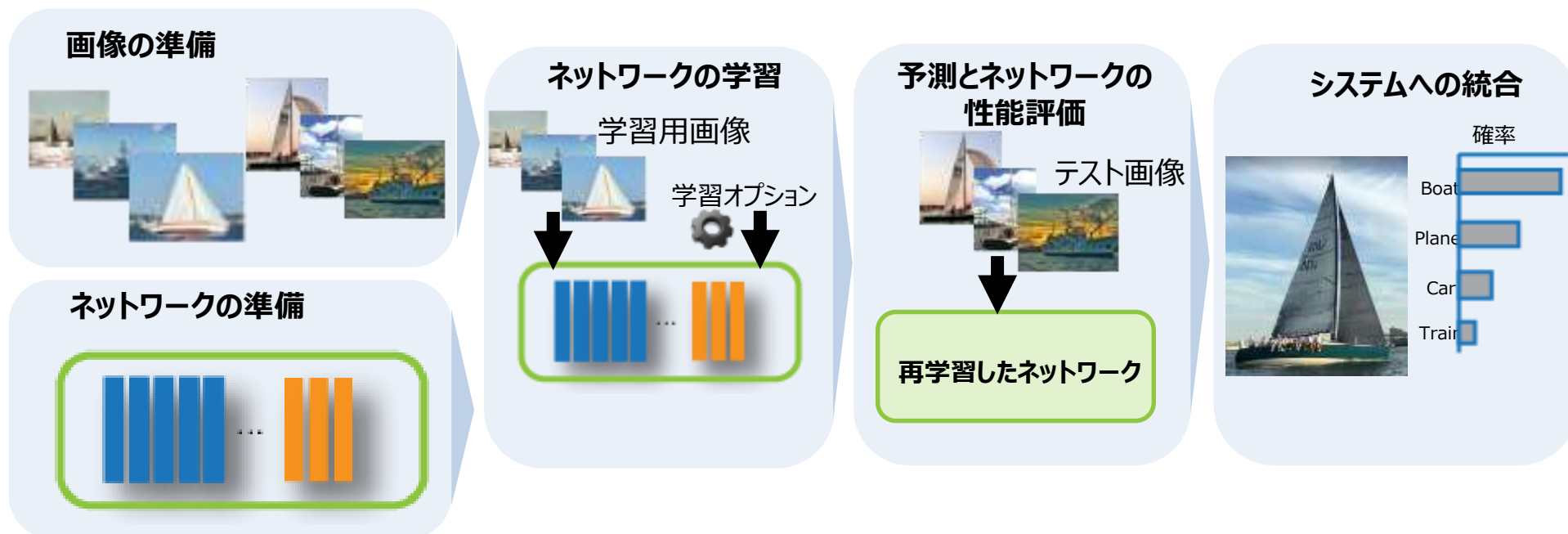


Image Processing toolbox
Neural Network toolbox

まとめ

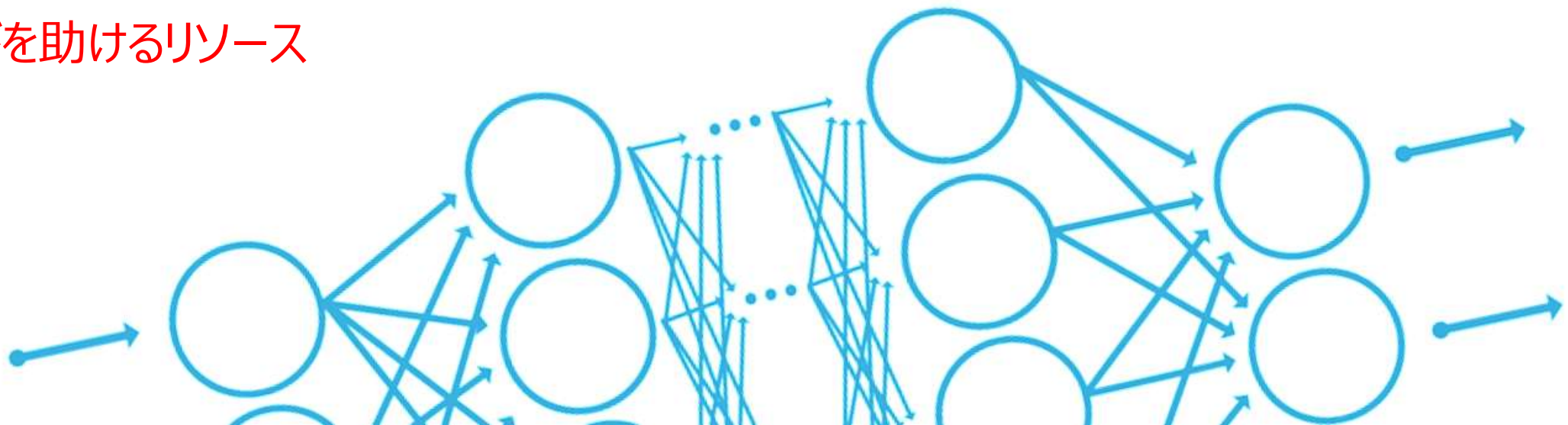
ディープラーニング分野でMATLABを利用するメリット

- ・画像があれば簡単にはじめられるフレームワーク
 - 転移学習を強力にサポート
 - 画像拡張、精度向上、可視化
 - CNN、R-CNN系、Semantic Segmentation
- ・作ったネットワークをすぐにシステムに統合できる



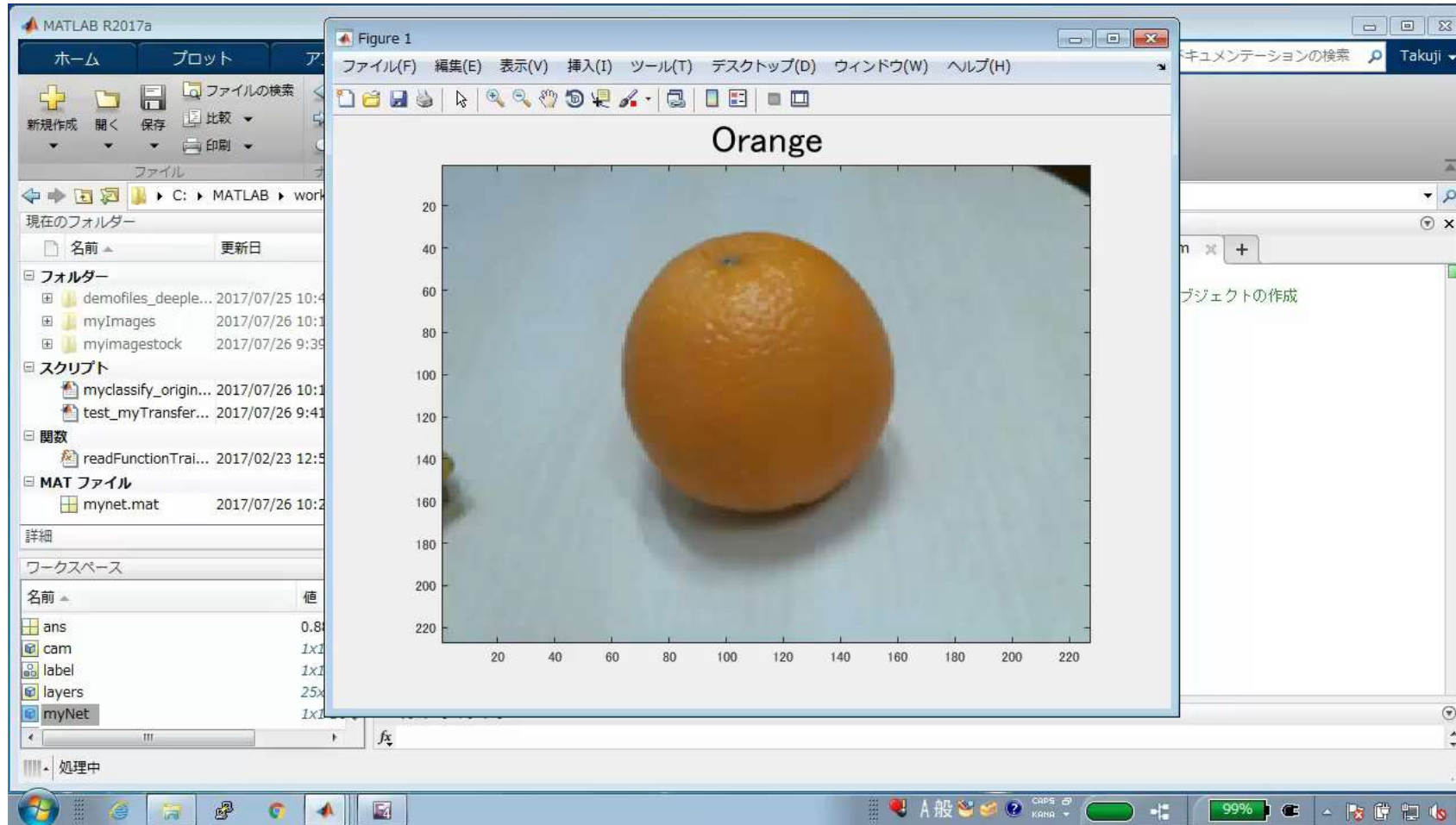
Agenda

- これからはじめるディープラーニング活用ワークフロー
- ディープラーニングの異常検知への適用
- 物体検出と領域の切り出し
- 学びを助けるリソース



ディープラーニングによる物体認識

ディープラーニング：10行でできる転移学習 ～画像分類タスクに挑戦～



学習した種類：

- オレンジ
- みかん
- グレープフルーツ(ルビー)
- グレープフルーツ(ホワイト)
- レモン

学習画像数：各 20 枚

- 要件を満たすPC&MATLAB環境
 - 学習済みAlexnet
 - 画像セット
- で10行のコーディングで始められます

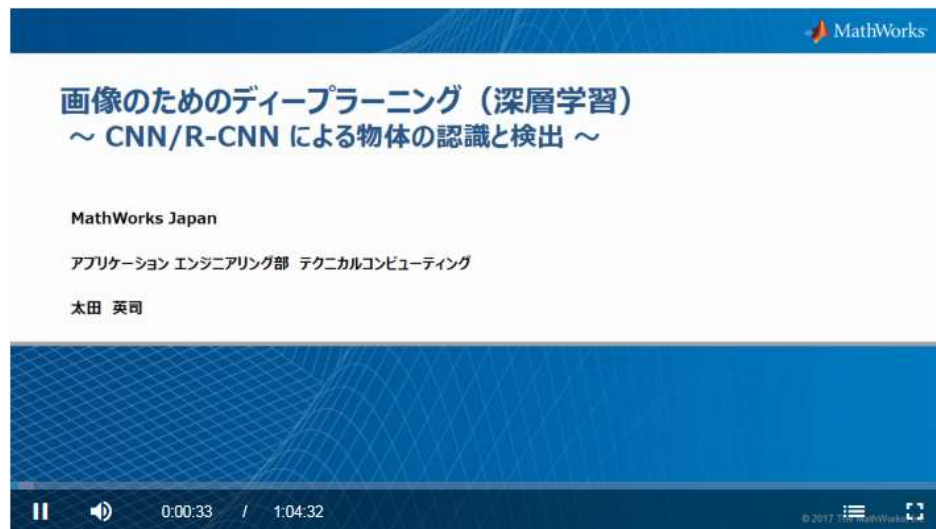
<https://www.youtube.com/watch?v=XMCHiMIT8iE>



実践的なディープラーニング物体認識を自分のペースで

画像のためのディープラーニング（深層学習） ～ CNN/R-CNN による物体の認識と検出 ～

<https://jp.mathworks.com/videos/object-recognition-and-detection-using-deep-learning-1490903520762.html>



MATLAB Academy

<https://matlabacademy.mathworks.com/jp>



WEBブラウザ上で
MATLABディープラーニングの操作体験



© 2017 The MathWorks, Inc. MATLAB and Simulink are registered trademarks of The MathWorks, Inc. See www.mathworks.com/trademarks for a list of additional trademarks. Other product or brand names may be trademarks or registered trademarks of their respective holders.

CNNをMATLAB®で利用するための必要要件

● MATLABライセンス

- MATLAB®
- Image Processing Toolbox™ (前処理、Computer Visionの必須要件)
- Computer Vision System Toolbox™ (物体認識の関数群)
- Statistics and Machine Learning Toolbox™ (SVM等の分類器)
- Neural Network Toolbox™ (CNN関数)
- Parallel Computing Toolbox™ (GPU使用(CNNで必須))

● ハードウェア

- NVIDIA CUDA対応GPU(Computing Capability 3.0以上) 搭載PC

最先端のディープラーニングをMATLABで手軽にお試しください