

# AIを支えるディープラーニングとその判断根拠

---

山下 隆義

中部大学 機械知覚ロボティクスグループ



MPRG

MACHINE PERCEPTION AND ROBOTICS GROUP

**人工知能 = ディープラーニング？**

- 人間の脳が行っている知的な作業をコンピュータで模倣したソフトウェアやシステム
- 自然言語を理解したり，論理的な推論を行ったり，経験から学習したりするコンピュータプログラムなど
- **人工知能の応用例**
  - 専門家の問題解決技法を模倣するエキスパートシステム
  - 翻訳を自動的に行う機械翻訳システム
  - 画像や音声の意味を理解する画像理解システム
  - 音声理解システムなど

- **汎用人工知能 (強いAI)**

- 「強い」人工知能とも呼ばれ、自意識・創造性などあらゆる面で人間と同等以上の知性を示す、いわゆる「人工知能」

- **特化型人工知能 (弱いAI)**

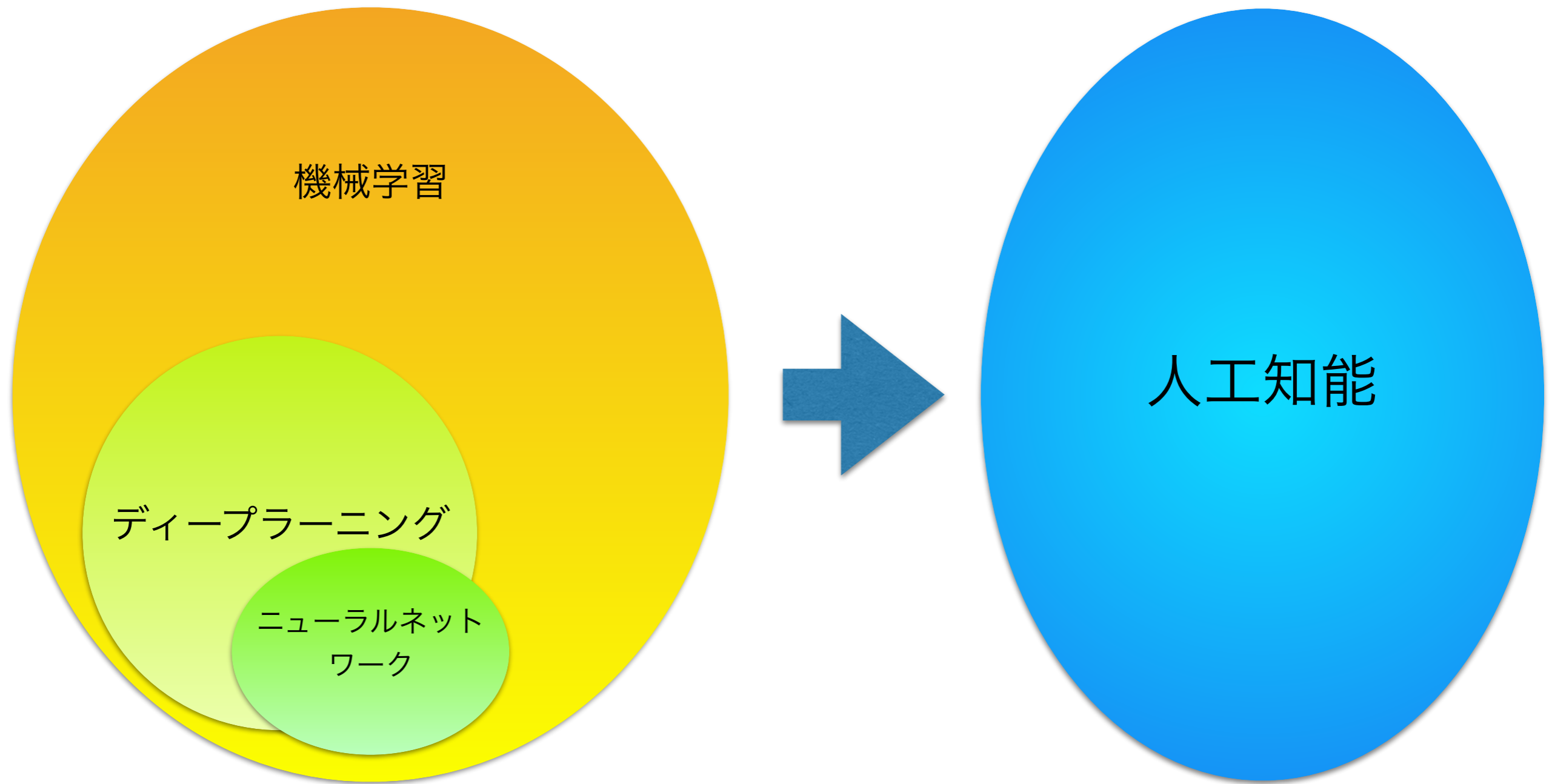
- ある特定のタスクで知性を示すソフトウェア、あるいはそれを組み込んだ自動機械
- 画像認識・翻訳システムは、特化型人工知能システム

→ディープラーニングは特化型人工知能を実現するための一手段である

人工知能技術の健全な発展のために Hiroshi Maruyama

<https://research.preferred.jp/2017/04/ai4future/>

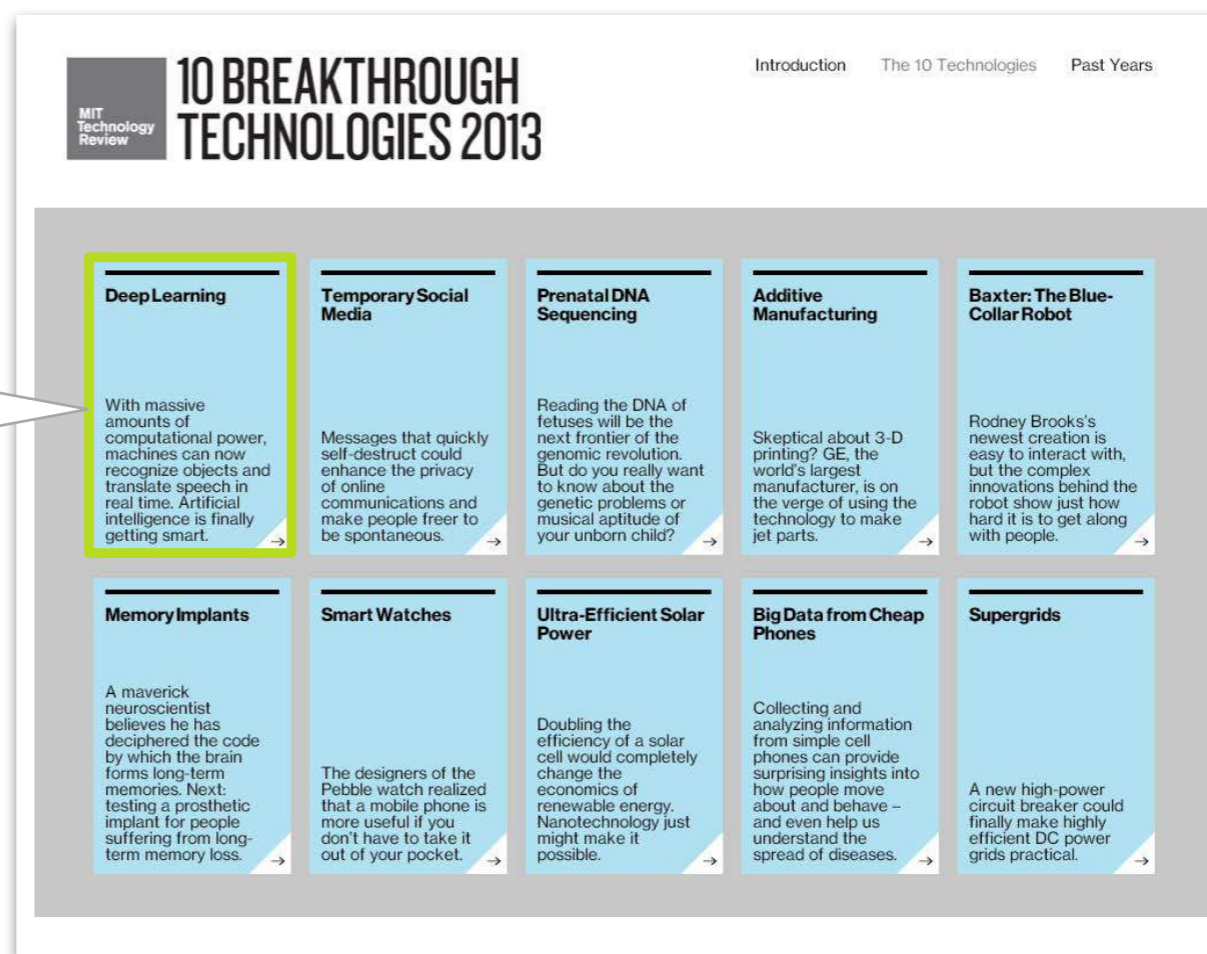
ディープラーニング ≠ 人工知能



機械学習 = 人工知能を実現する手段

- No.1 of top 10 breakthrough technologies in 2013 (MIT technology review)

With massive amounts of computational power, machine can now recognize objects and translate speech in real time.  
AI is finally getting smart.



<https://www.technologyreview.com/lists/technologies/2013/>

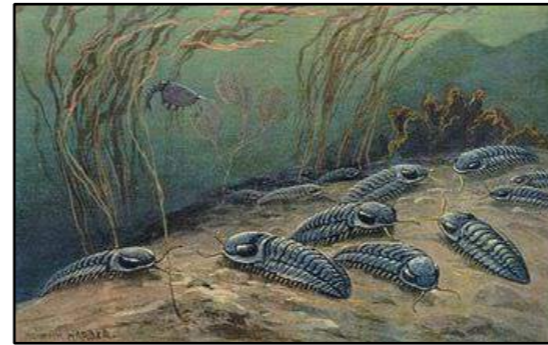
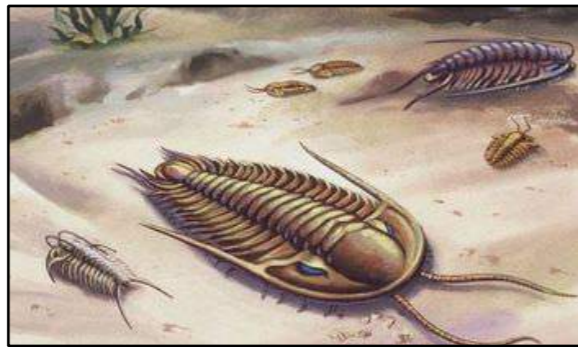
これまでの“*intelligence*”は人が設計したインテリジェンス



これからの“*intelligence*”はデータドリブンで獲得したインテリジェンス

- ・ カンブリア爆発
  - 5億4,200万年前から5億3,000年前の間に突如として今日見られる動物の「門」が出揃った現象
  - 古生物学者アンドリュー・パーカーは、「眼の誕生」がその原因だったという光スイッチ説を提唱

三葉虫：



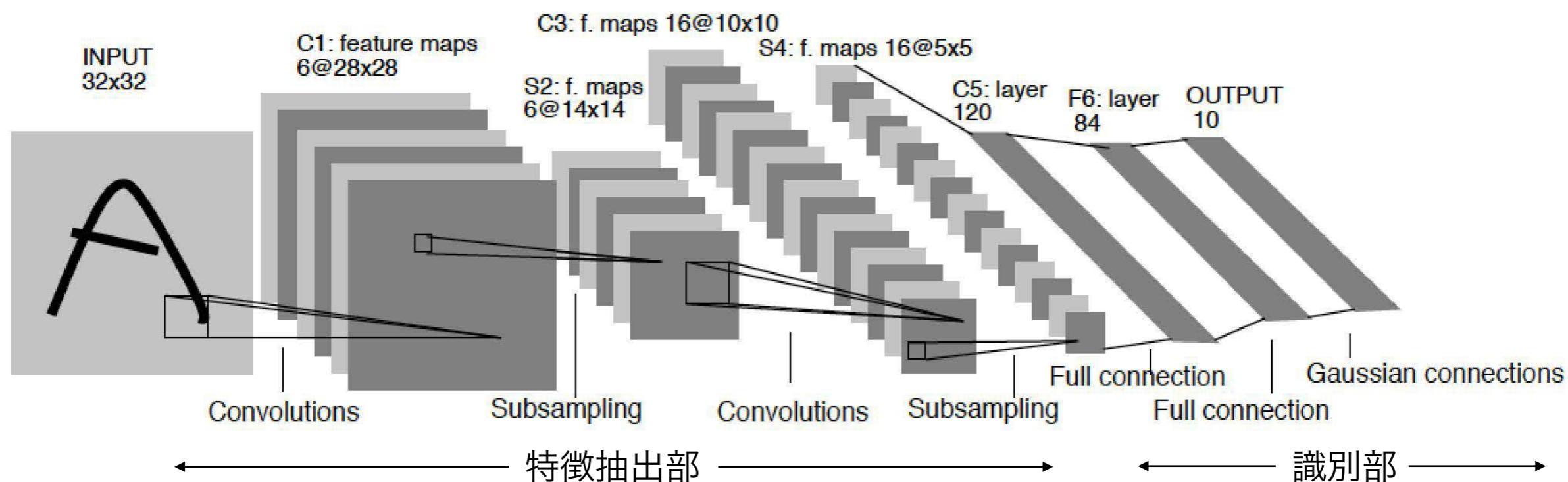
史上初めて眼を持った生物



<https://www.amazon.co.jp//dp/4794214782/>

- ・ 「眼を持った機械」が誕生する
  - 機械・ロボットの世界でカンブリア爆発が起こる
  - 日本の企業が産業活用できるか？

- ・ 畳み込みニューラルネットワーク [LeCun1998]
  - 畳み込み・プーリング・全結合の各層から構成
  - 手書き文字認識に応用
  - 平行移動に対する不変性がある

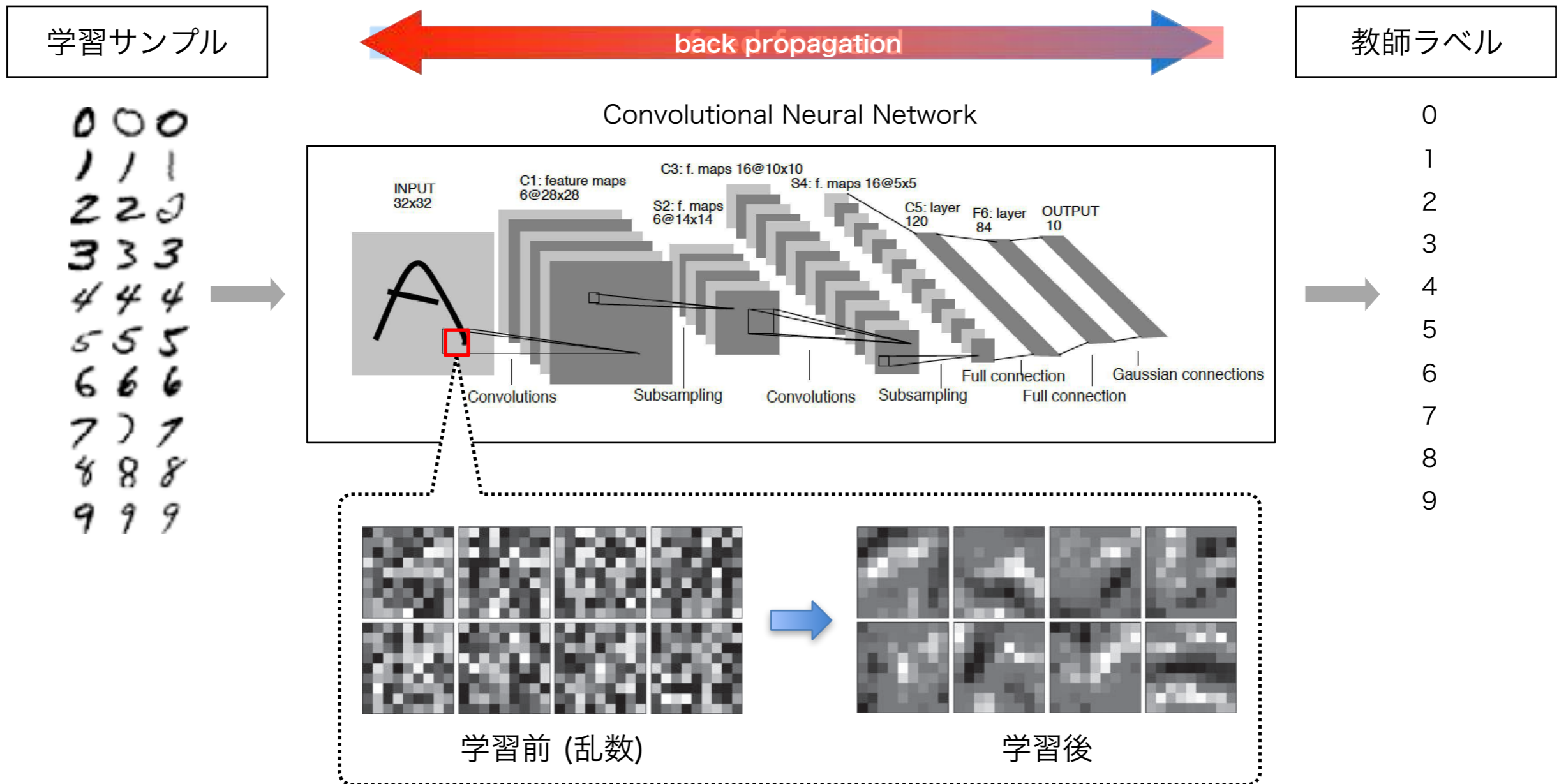


→学習により特徴表現を自動獲得



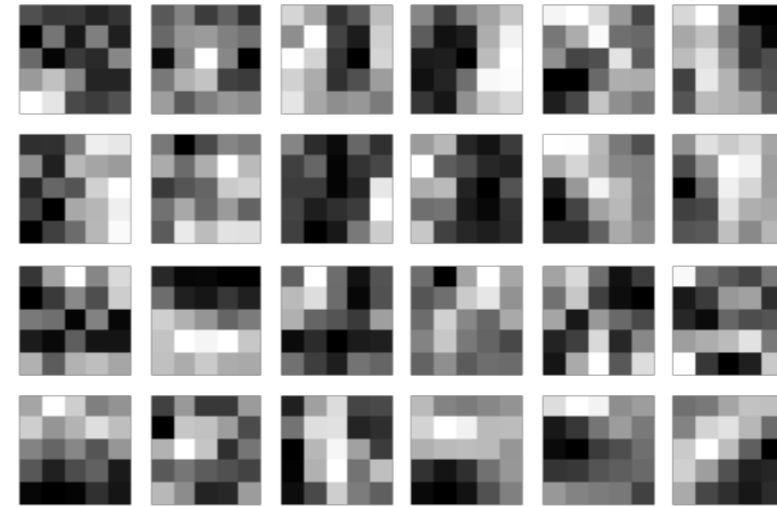
# 学習によるカーネル(フィルタ)の獲得

- 特徴表現の自動獲得：学習によりカーネル(フィルタ)を更新



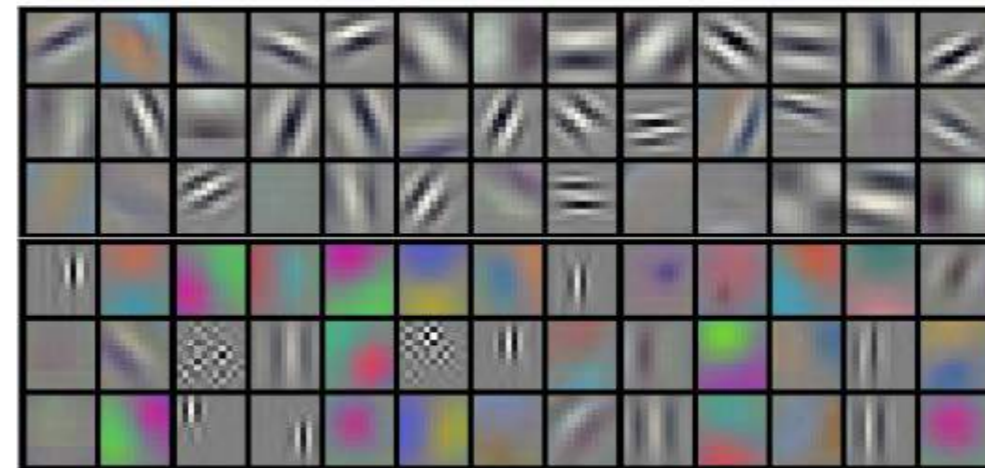
# 学習によって獲得したカーネルの例

文字認識：MNIST



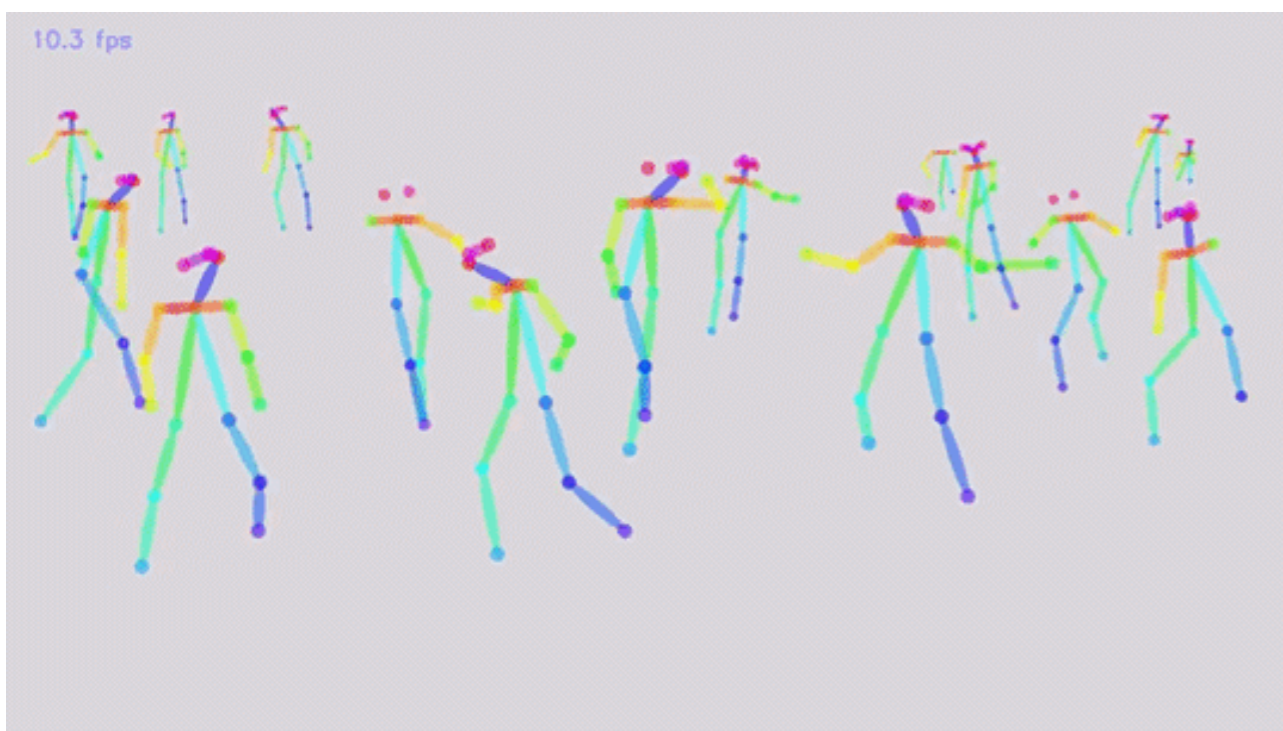
Kernel (5x5)

1000クラス画像分類：ImageNet



Kernel (11x11)

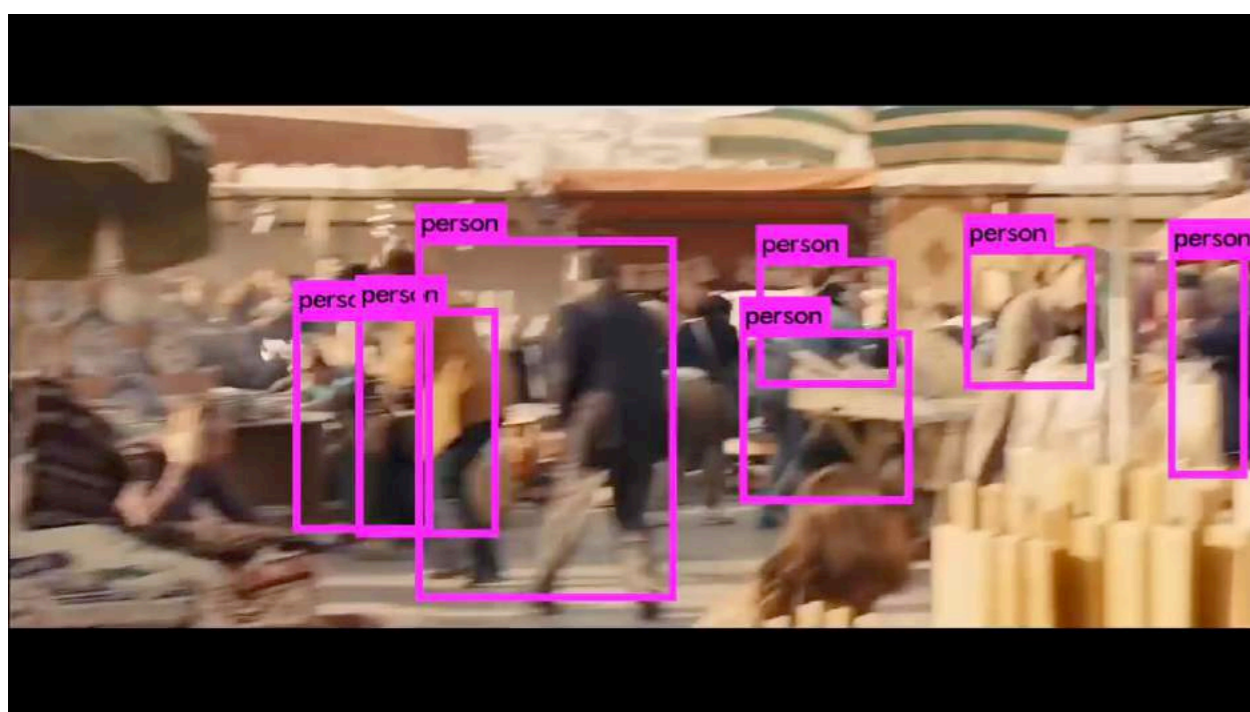
入出力の組み合わせ次第で色々な問題に応用できる



姿勢推定



セグメンテーション



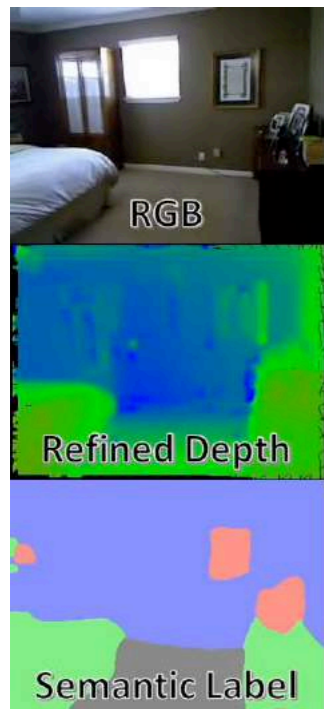
物体検出



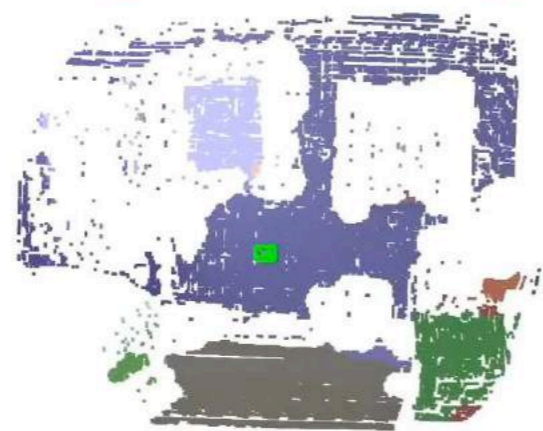
行動認識



## フロー推定



FPS: 31.158907  
■:Floor ■:Vertical structure/Wall  
■:Large structure/furniture ■:Small structure

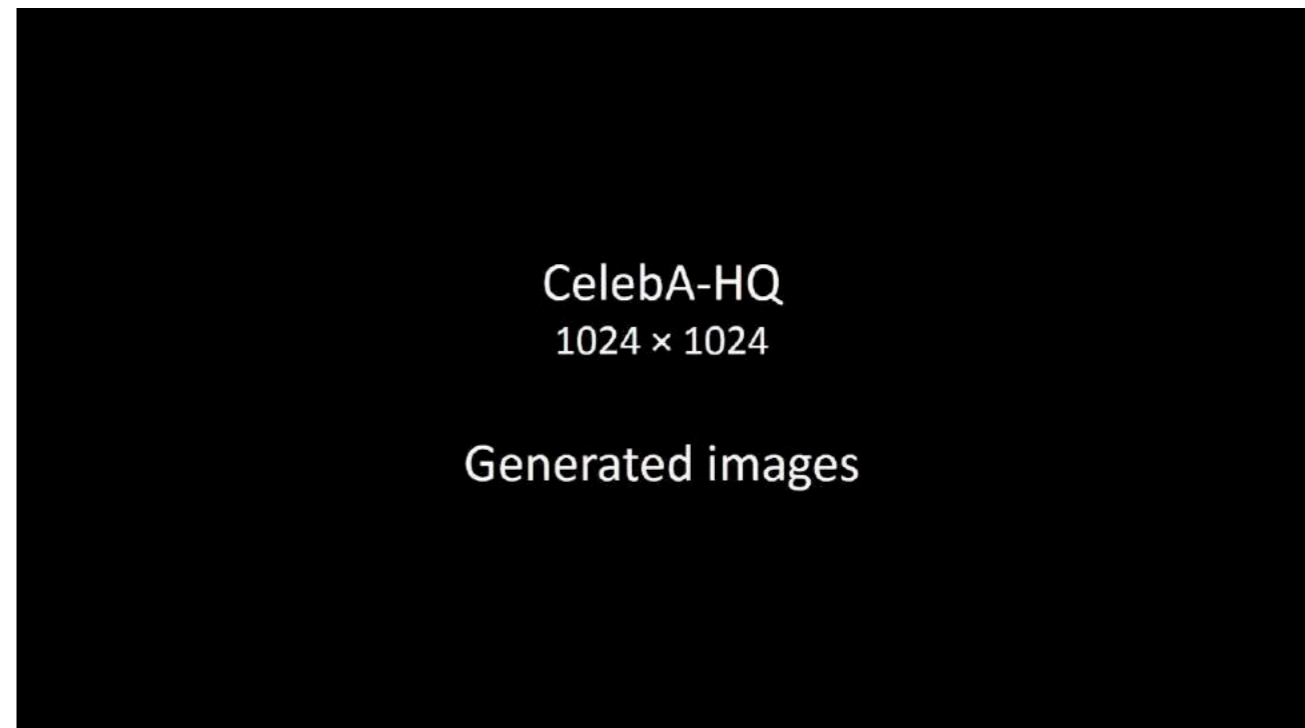


Result of dense 3D reconstruction and semantic label fusion

## SLAM



## 強化学習



## 画像生成(GAN)

## ものづくり

- ・自動車
- ・産業ロボット
- ・家電
- ・農業機械
- ・建設機械
- ・医療機器
- ...

大企業中心  
ベテラン技術者が戦力

×

ディープラーニング  
「眼の技術」

=

## 作業の自動化

- ・自動運転
- ・組立加工の自動化
- ・外観検査の自動化
- ・建設作業の自動化
- ・農作業の自動化
- ・画像診断の自動化
- ・物流の自動化
- ・介護の自動化
- ・陳列廃棄の自動化
- ...

地方での活用  
人手不足の解消

## ものづくり

- ・自動車
- ・産業ロボット
- ・家電
- ・農業機械
- ・建設機械
- ・医療機器
- ...

×

ディープラーニング  
「眼の技術」

=

## 作業の自動化

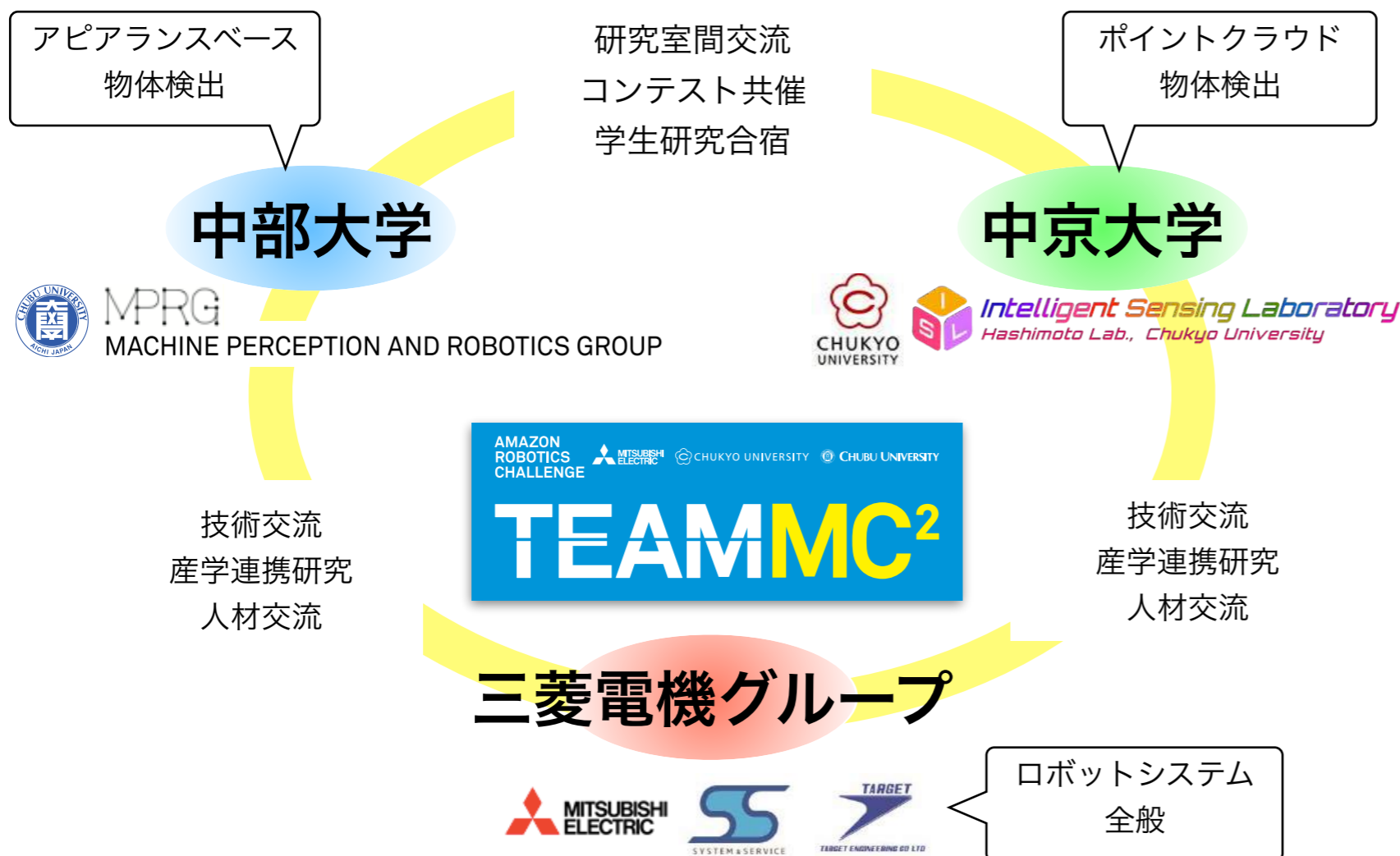
- ・自動運転
- ・組立加工の自動化
- ・外観検査の自動化
- ・建設作業の自動化
- ・農作業の自動化
- ・画像診断の自動化
- ・物流の自動化
- ・介護の自動化
- ・陳列廃棄の自動化
- ...

MPRGの取り組み：

Amazon Robotics Challenge (アマゾンが主催する国際ロボット大会)

# Amazon Robotics Challenge 2017

- 日時：2017年7月26日～7月30日
- 場所：名古屋市, 日本





# Amazon fulfillment( 注文処理 ・ 出荷)



- 2台の産業ロボットによる物体形状に合わせた三種類のハンドリングを実現

把持対象物体  
(柔軟物や透明なアイテムを含む)

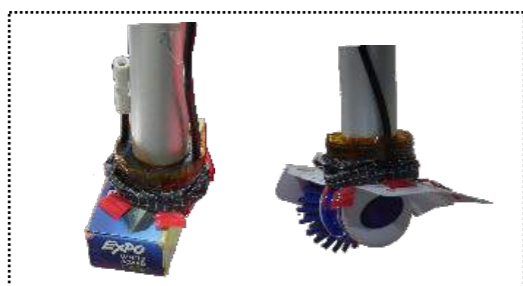


## Hand 1: Powerful vacuum

変形パッドと強力な吸引による把持



変形パッド  
(シリコン)



→ラフな形状を把持可能

Robot 1

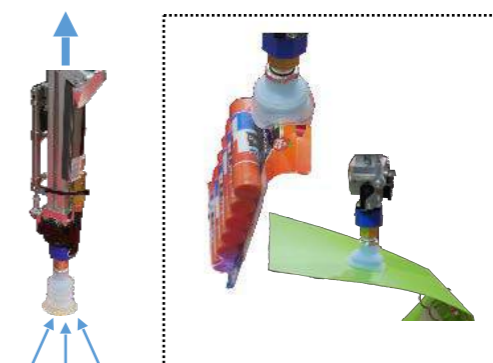
Robot 2



Team MC^2 ロボットシステム

## Hand 2: Slim vacuum

小さなパッドによる吸引



→ 狭い空間にある物体の把持

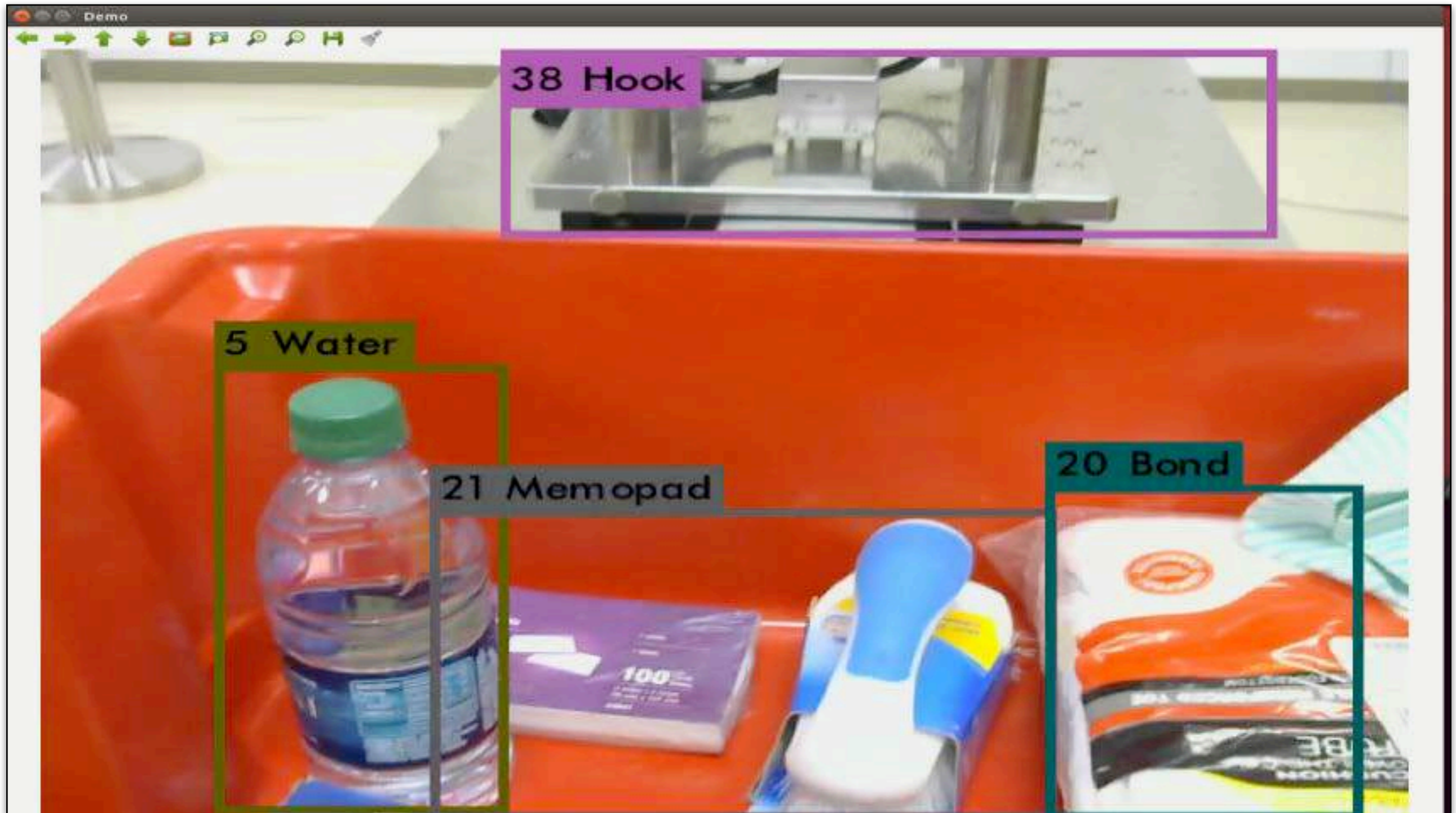
## Hand 3: Gripper

二指ハンドによる挟持



**Grip!**

→ 吸引できない物体を挟持して把持



15 FPS@GeForce GTX1080

# Team MC<sup>2</sup>: Stow task in ARC2017

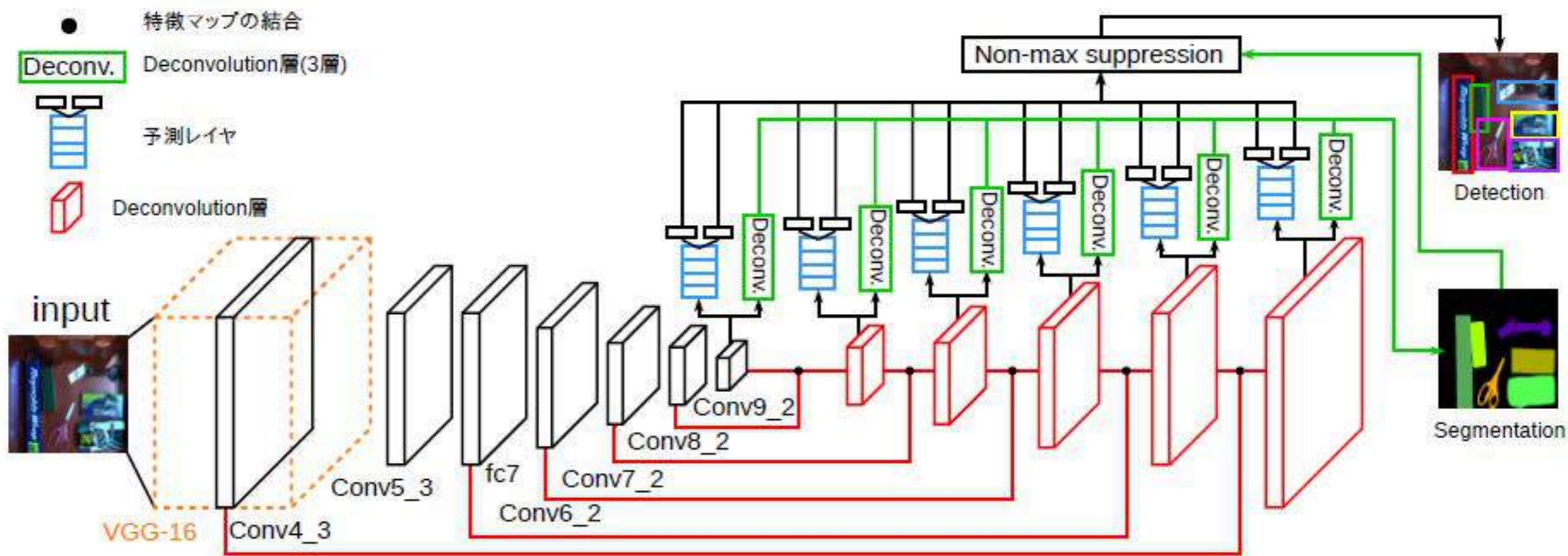
<https://vimeo.com/227401334>



10:00	ACRV	55	12:30	CMU PLAID	20	15:00	MC <sup>2</sup>	120	17:00	NimbRo Picking	20
10:30	Team Duke	10	13:00	Team T2	10	15:30	TKU M-Bot	5	17:30	Nanyang	125
11:00	NAIST-Panasonic	110	13:30	IITK-TCS	105	16:00	UJI RobInLab	20	18:00	MIT-Princeton	160
12:00	IFL PiRo	25	14:30	Team K	5	16:30	Applied Robotics	5	18:30	GMU-Negev	0

# STOW ROUND SCHEDULE

- シングルショット系の物体検出の出力において、セグメンテーションを同時に行うようにする
  - DSSD : エンコード・デコード構造の物体検出
  - セグメンテーション結果をNMSに反映

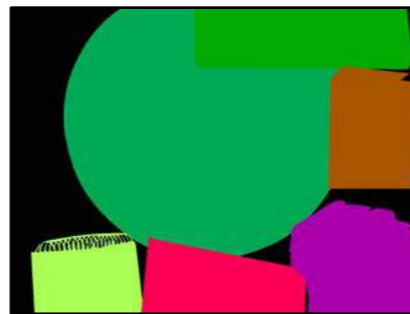


# 評価結果(セグメンテーション)

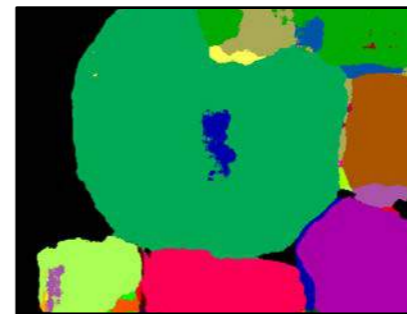
手法	Global Accuracy[%]	Class Accuracy[%]	Mean IoU[%]
SegNet	78.19	72.16	53.97
	↓ +14%	↓ +15%	↓ +24%
提案手法	91.80	87.31	78.30



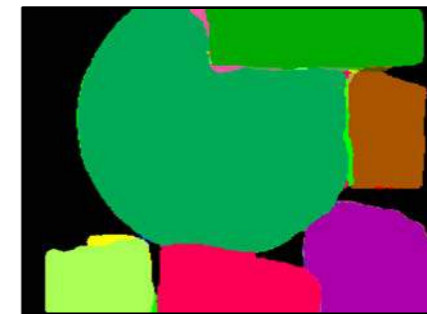
入力画像



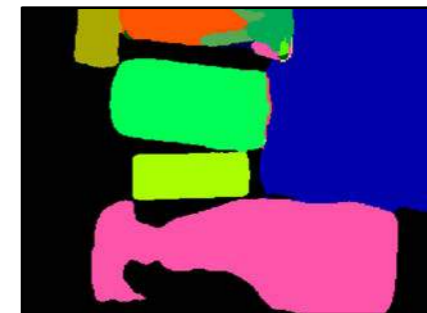
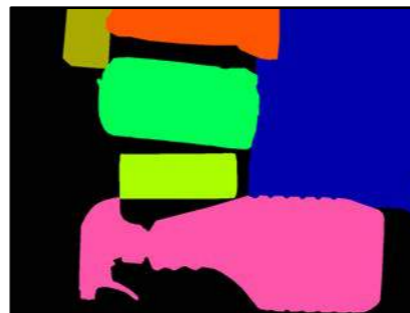
教師画像



SeaNet

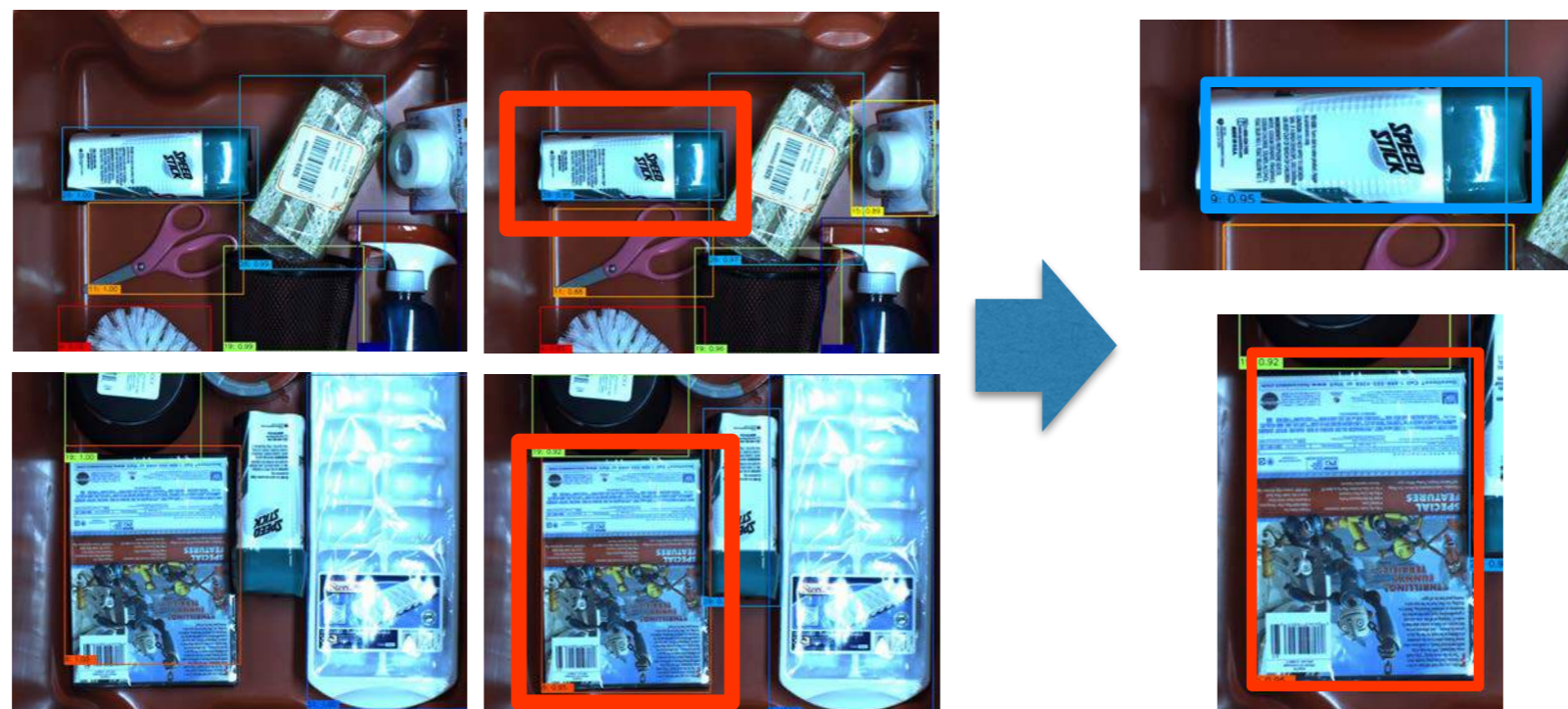


提案手法



# 評価結果(物体検出)

手法	識別率[%]	未検出率[%]	boxの重なり率[%]
DSSD	88.43	30.19	81.37
提案手法	91.50	29.68	83.68



検出位置精度が向上

DSSD

提案手法



## ものづくり

- ・自動車
- ・産業ロボット
- ・家電
- ・農業機械
- ・建設機械
- ・医療機器
- ...

×

ディープラーニング  
「眼の技術」

=

## 作業の自動化

- ・自動運転
- ・組立加工の自動化
- ・外観検査の自動化
- ・建設作業の自動化
- ・農作業の自動化
- ・画像診断の自動化
- ・物流の自動化
- ・介護の自動化
- ・陳列廃棄の自動化
- ...

MPRGの取り組み：

WRSフューチャーコンビニエンスストアチャレンジ

- ・ 日本（経済産業省、NEDO）が主催するロボットの国際大会
  - ものづくり、サービス、インフラ・災害対応、ジュニア
- ・ フューチャーコンビニエンスストアチャレンジ (FCSC)
  - おにぎり、お弁当などの自動補充および消費期限切れ商品の廃棄を競う



おにぎり  
商品名: 熟成紅しゃけ  
外寸: 約H75 x W80 x D35 mm  
重さ: 約110 g



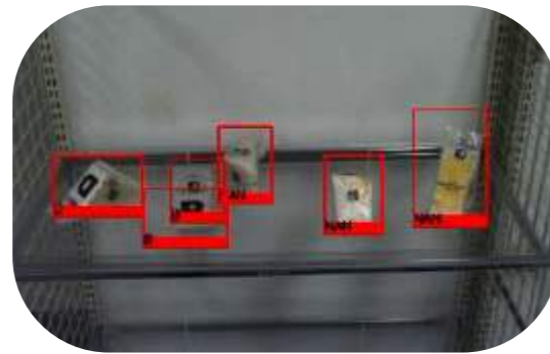
ドリンク  
商品名: カフェラテ (240 ml)  
外寸: 約H108 x W76 x D76 mm  
重さ: 約260g



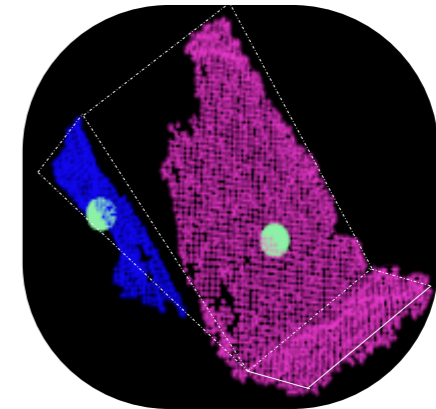
弁当  
商品名: チキンカツ弁当  
外寸: 約H50 x W250 x D175 mm  
重さ: 約535g



サンドイッチ  
商品名: たっぷりたまごサンド  
外寸: 約H140 x W90 x D70 mm  
重さ: 約105g



Deep learningを用いた  
対象物の場所推定技術



点群セグメンテーションを  
用いた吸着点の推定技術



## ROC2/Japan



フューチャーコンビニエンスストアチャレンジ/Future Convenience Store Challenge

Task-3

陳列・廃棄タスク  
Stocking and Disposing Tasks

00:15:25



結果発表 陳列・廃棄タスク

TEAM	SCORE	TEAM	SCORE
Mei Blue	50	homer@UniKoblenz	14
TAK	16	ST	40
Happy Mini @ Girls Art Project	—	HSRL-CoR	0
Happy Robot	9	TCR	0
<b>ROC2</b>	<b>74</b>	Hi-KCCT	59
OCU-KDEL	56	<b>U.T.T.</b>	<b>74</b>
TIMDA	44	NAIST-RITS-Panasonic	42



2位入賞 (1位と同点)  
日本ロボット学会特別賞

2018/11/07 愛知県 近郊版 (犬山、小牧、春日井など)

2018年(平成30年)11月7日(水曜)

中部大(春日井市)と中京大(名古屋市昭和区)は京都市の電気機器メーカー・オムロンと合同でつくるロボット研究チーム「ROC2(ロックツ)」で、国際ロボット競技会「World Robot Summit 2018(WRS)」に出場し、準優勝した。競技会は、高齢化や担い手の減少など課題の多い働く場におけるロボットの技術開発を目指している。  
(高岡涼子)

競技会場でロボットを囲んで記念撮影するROC2のメンバーたち



## 国際ロボット競技会



## 中部大チーム準V

競技会は十月十七〜二十一日、東京都江東区のビッグサイトで開催。両大学のチームは、コンビニエンスストアでの商品陳列・廃棄を想定した部門に出場。ロボットが弁当やおにぎり、飲み物を棚に並べたり、消費期限を確認して廃棄したりする正確さと速さを競った。同部門には十四チームが出場した。

ROC2が開発したロボット=いずれも東京都の東京ビッグサイトで(藤吉教授提供)

ROC2では、オムロンがロボット全体のシステム構築を担当し、中部大はカメラを通してどこにどんな商品があるかを認識する技術、中京大は商品のどの部分を的確につかむかを判断する技術を担当した。

ROC2は、オムロンがロボット全体のシステム構築を担当し、中部大はカメラを通してどこにどんな商品があるかを認識する技術、中京大は商品のどの部分を的確につかむかを判断する技術を担当した。

ROC2は、オムロンがロボット全体のシステム構築を担当し、中部大はカメラを通してどこにどんな商品があるかを認識する技術、中京大は商品のどの部分を的確につかむかを判断する技術を担当した。

ROC2は、オムロンがロボット全体のシステム構築を担当し、中部大はカメラを通してどこにどんな商品があるかを認識する技術、中京大は商品のどの部分を的確につかむかを判断する技術を担当した。

## 高度な商品認識技術 搭載

(c)中日新聞社 無断転載、複製、頒布は著作権法により禁止されています

2018年11月7日 中日新聞(近郊版)

## ものづくり

- ・自動車
- ・産業ロボット
- ・家電
- ・農業機械
- ・建設機械
- ・医療機器
- ...

×

ディープラーニング  
「眼の技術」

=

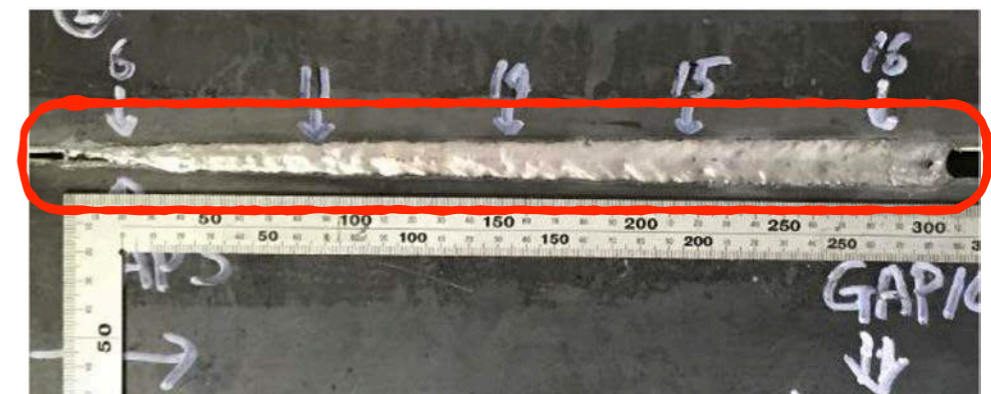
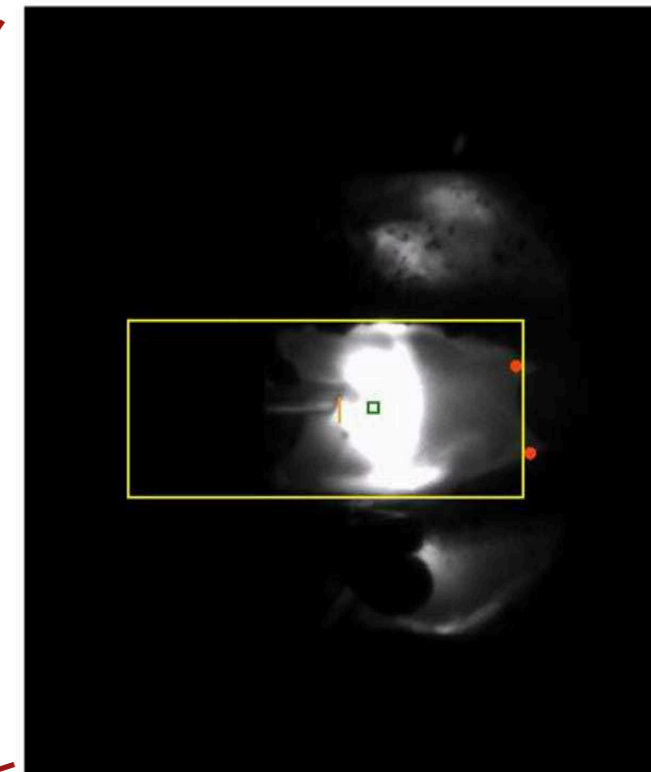
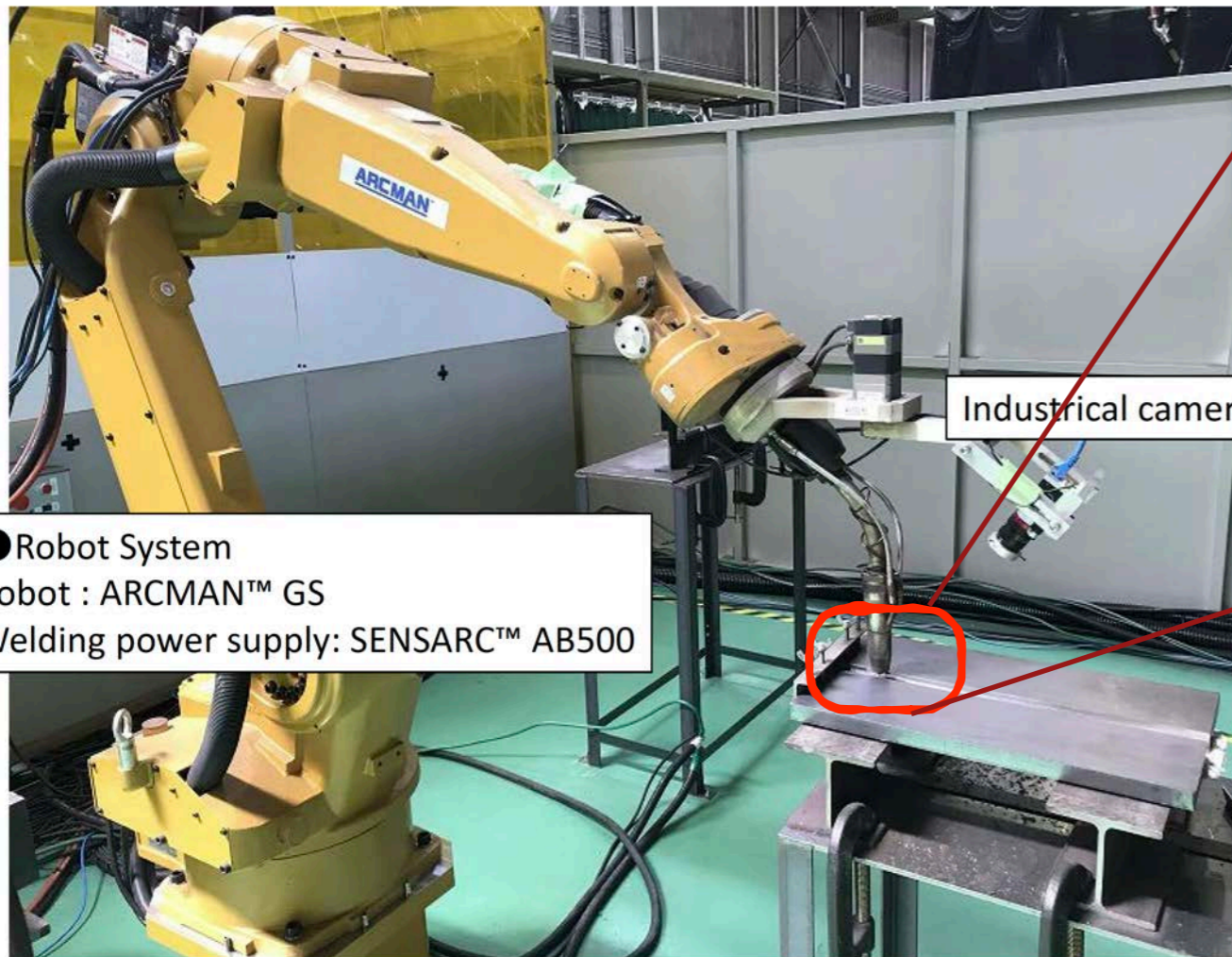
## 作業の自動化

- ・自動運転
- ・組立加工の自動化
- ・外観検査の自動化
- ・建設作業の自動化
- ・農作業の自動化
- ・画像診断の自動化
- ・物流の自動化
- ・介護の自動化
- ・陳列廃棄の自動化
- ...

MPRGの取り組み：

溶接作業の自動化（共同研究）

- ・ アーク溶接の状態を認識, ロボットを制御



**高い性能を達成するために必要なものは？**



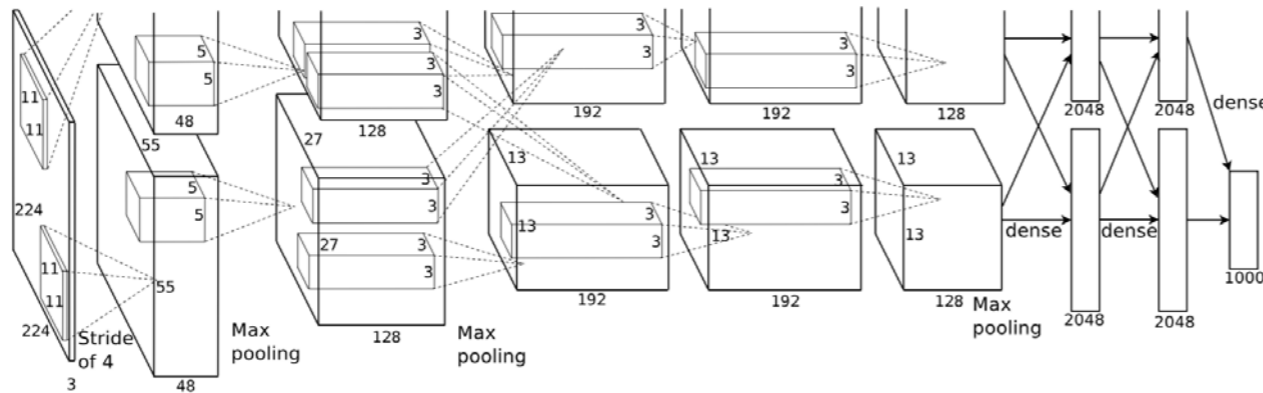
アルゴリズム

データセット

計算機リソース

- 物体認識のコンペティションを通じて飛躍的に進歩

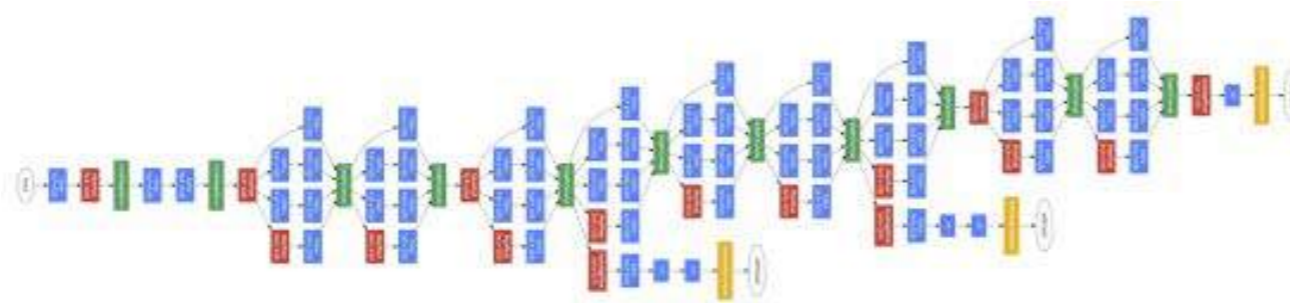
AlexNet



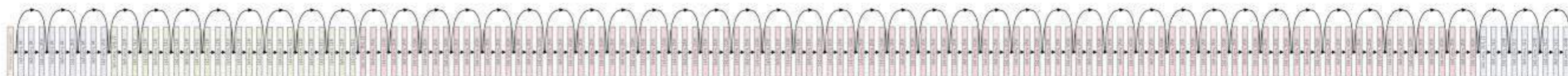
VGG16



GoogLeNet



ResNet



Team	year	Error
AlexNet	2012	15.3%
Clarifai	2013	11.2%
VGG	2014	7.32%
GoogLeNet	2014	6.67%
ResNet	2015	3.57%
ResNet+	2016	2.99%
SENet	2017	2.25%
human expert		5.1%

# アルゴリズム (物体検出)

**R-CNN** → **OverFeat** → MultiBox → SPP-Net → MR-CNN → DeepBox → AttentionNet →  
2013.11 ICLR' 14 CVPR' 14 ECCV' 14 ICCV' 15 ICCV' 15 ICCV' 15

**Fast R-CNN** → DeepProposal → **Faster R-CNN** → **OHEM** → **YOLO v1** → G-CNN → AZNet →  
ICCV' 15 ICCV' 15 NIPS' 15 CVPR' 16 CVPR' 16 CVPR' 16 CVPR' 16

Inside-OutsideNet(ION) → HyperNet → CRAFT → MultiPathNet(MPN) → **SSD** → GBDNet →  
CVPR' 16 CVPR' 16 CVPR' 16 BMVC' 16 ECCV' 16 ECCV' 16

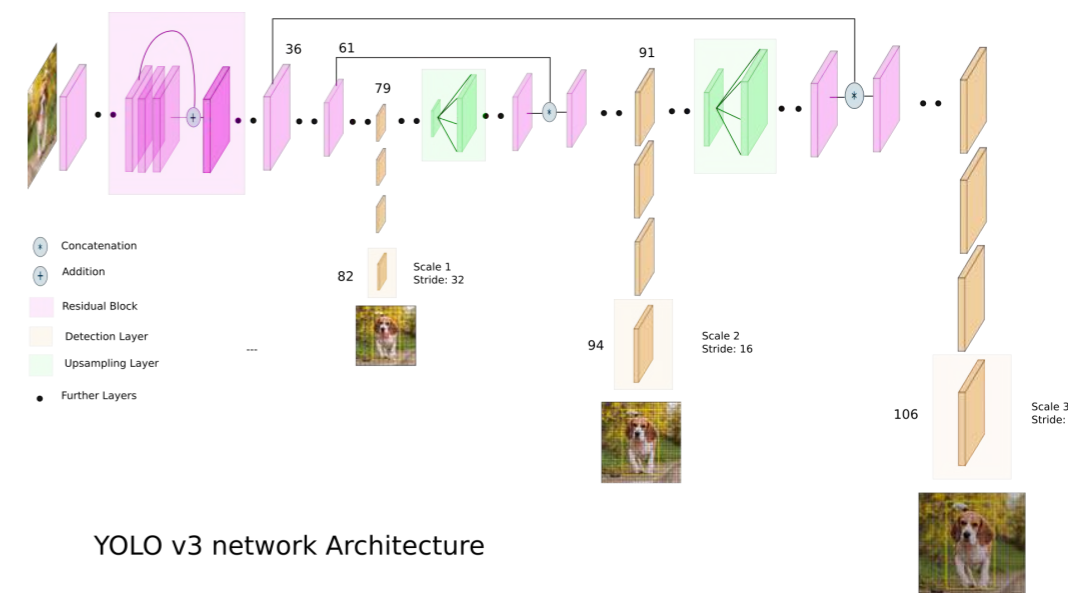
CPF → MS-CNN → **R-FCN** → PVANET → DeepID-Net → NoC → DSSD → TDM → **YOLO v2** →  
ECCV' 16 ECCV' 16 NIPS' 16 NIPSW' 16 PAMI' 16 TPAMI' 16 arXiv' 17 CVPR' 17 CVPR' 17

Feature Pyramid Net(**FPN**) → RON → DCN → DeNet → CoupleNet → **RetinaNet** → DSOD →  
CVPR' 17 CVPR' 17 ICCV' 17 ICCV' 17 ICCV' 17 ICCV' 17 ICCV' 17

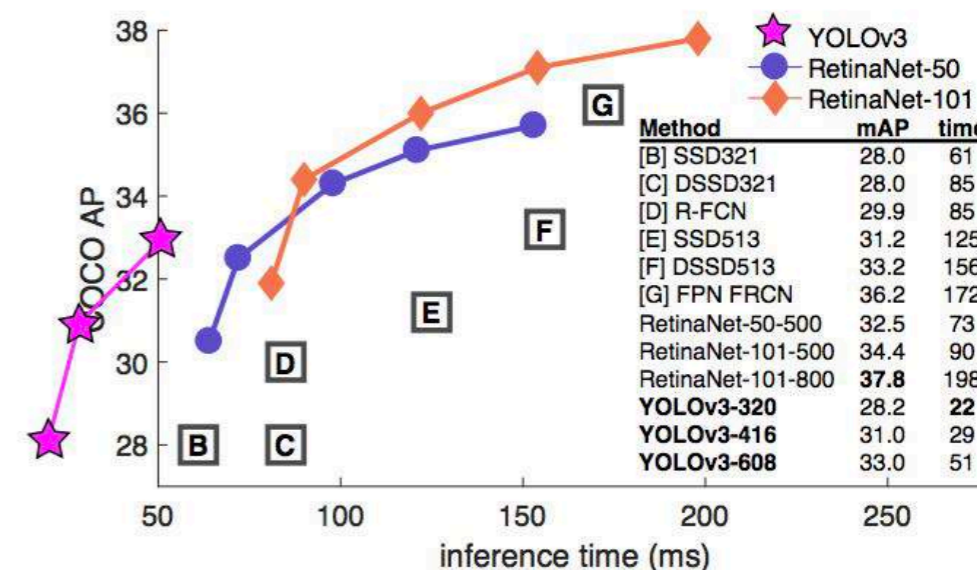
**Mask R-CNN** → SMN → **YOLO v3** → SIN → STDN → **RefineDet** → MLKP → Relation-Net →  
ICCV' 17 ICCV' 17 arXiv' 18 CVPR' 18 CVPR' 18 CVPR' 18 CVPR' 18 CVPR' 18

Cascade R-CNN → RFBNet → CornerNet → PFPNet → Pelee → HKRM → R-DAD → **M2Det** ...  
CVPR' 18 ECCV' 18 ECCV' 18 ECCV' 18 NIPS' 18 NIPS' 18 AAAI' 19 AAAI' 19

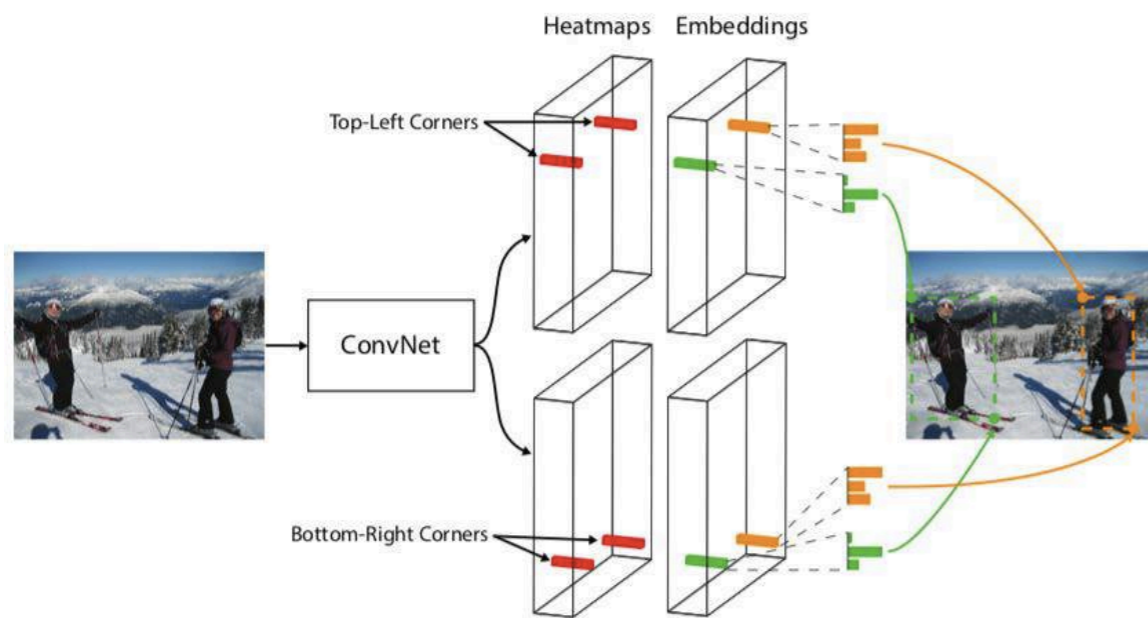
- ベースネットワーク
  - DarkNet19 -> DarkNet53
- 矩形の予測
  - アンカーボックスからのオフセット予測



- マルチスケール
  - 3つスケールから物体検出を行う
  - 各スケールの特徴マップをアップサンプリング



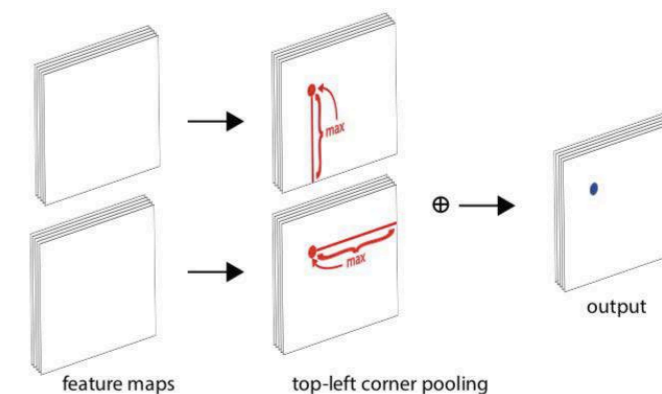
- 物体の左上座標と右下座標を推定  
→ アンカーボックスを必要としない物体検出手法
- 各CornerのヒートマップからCorner poolingによりオフセットを算出



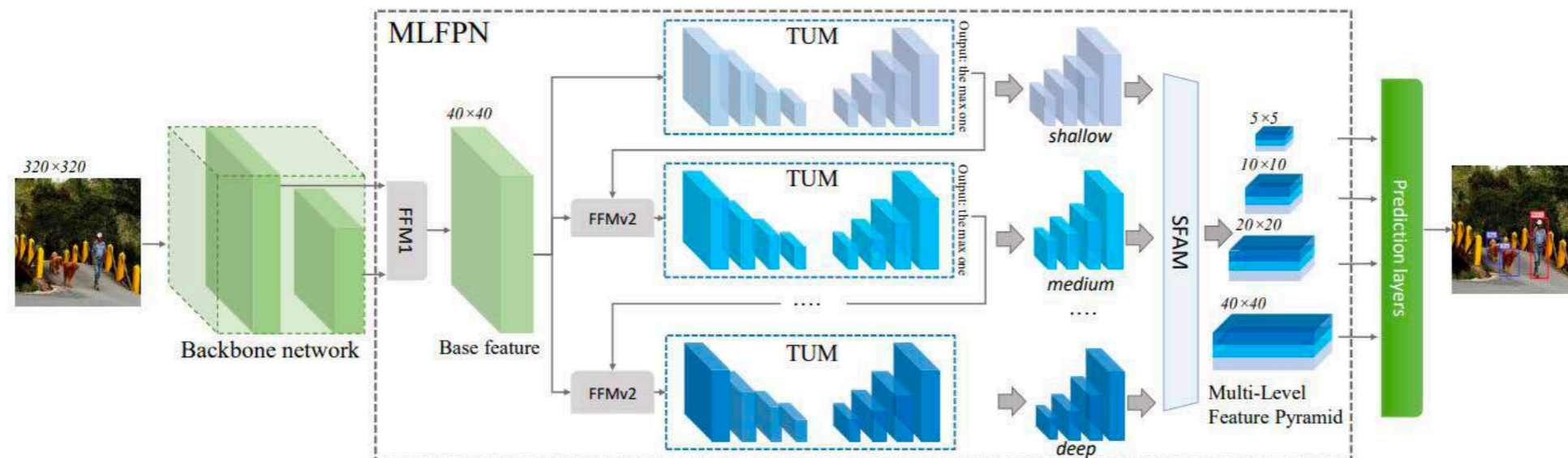
垂直方向: 
$$t_{ij} = \begin{cases} \max(f_{t_{ij}}, t_{(i+1)j}) & \text{if } i < H \\ f_{t_{Hj}} & \text{otherwise} \end{cases}$$

水平方向: 
$$l_{ij} = \begin{cases} \max(f_{l_{ij}}, l_{i(j+1)}) & \text{if } j < W \\ f_{l_{iW}} & \text{otherwise} \end{cases}$$

top-left cornerに対するcorner pooling

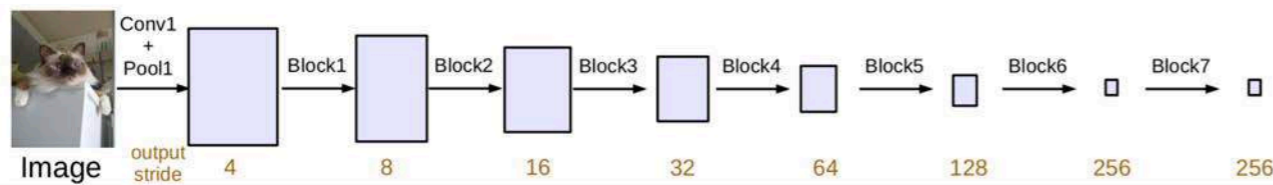


- One-shot ベースの最新手法
  - Backbone network : VGG
  - TUM : Feature Pyramid NetworkやDSSDのようなアップサンプリングをマルチスケールで行う
  - SFAM : 複数スケールの特徴マップをSqueeze & Excitationして集約

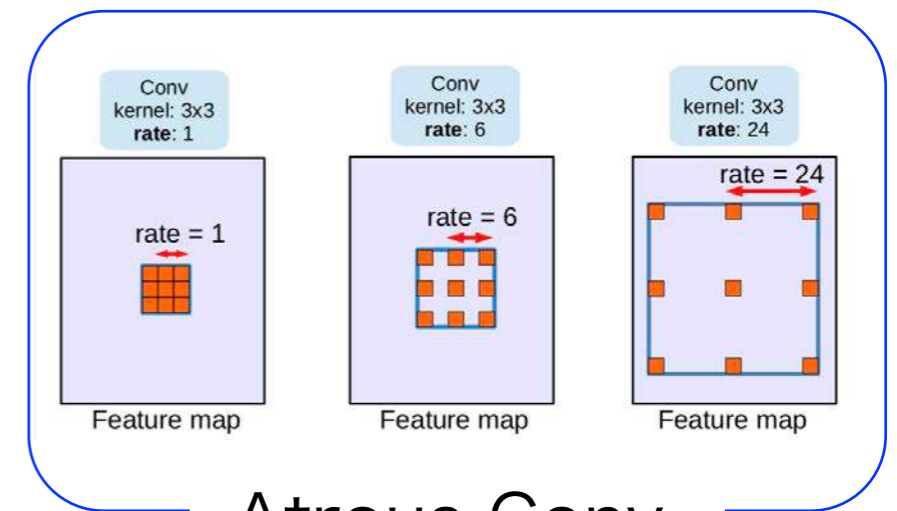
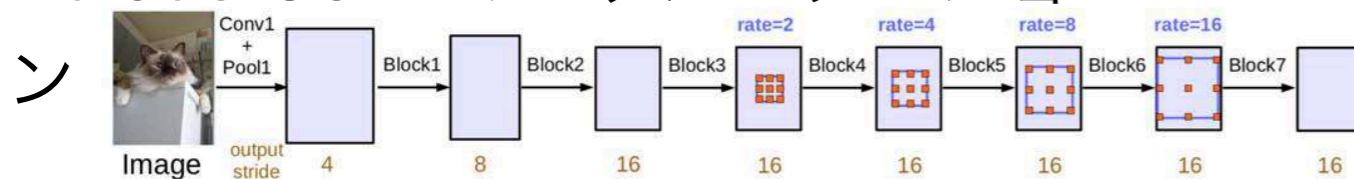


- Atrous Convolutionにより特徴マップのサイズを小さくしない
- 特徴マップが大きいため、詳細なセグメンテーションが可能

## 一般的なセグメンテーション

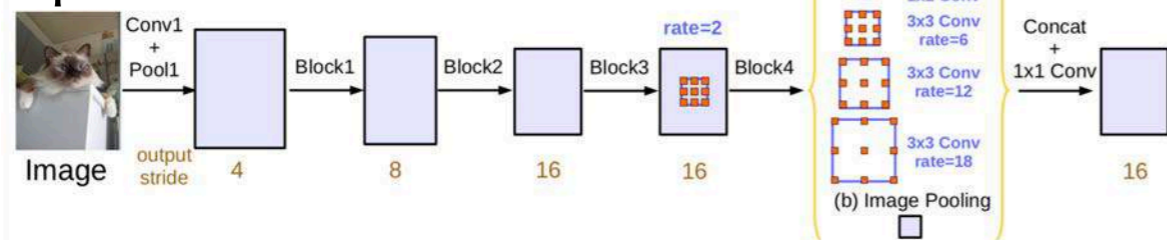


## Atrous Conv.のセグメンテーション

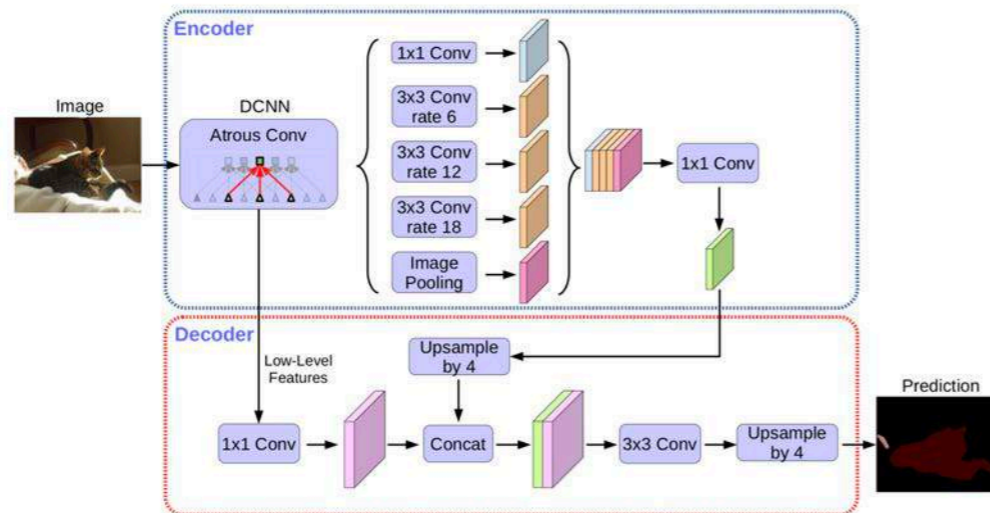
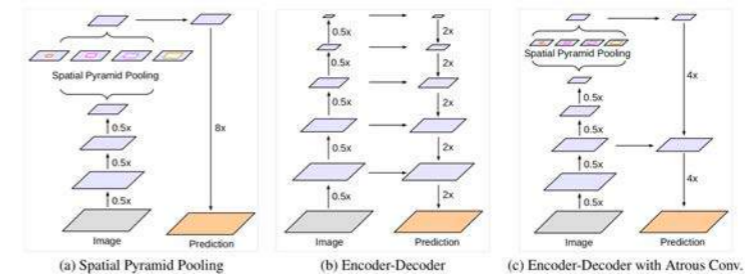
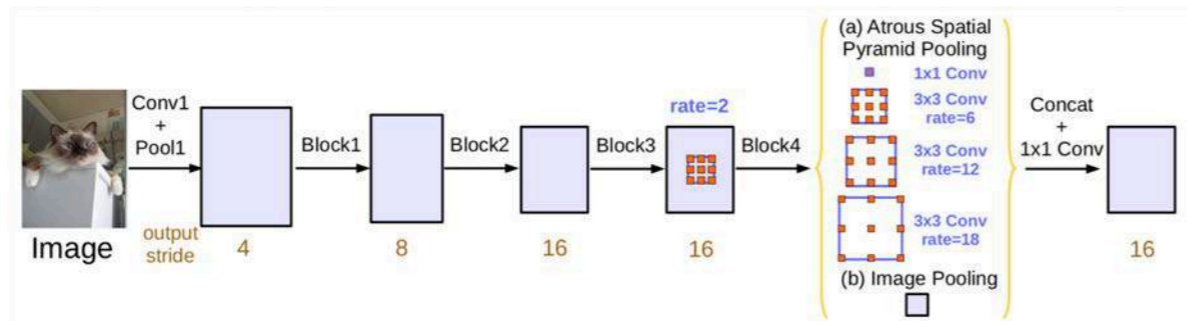


Atrous Conv.

## DeepLab V3



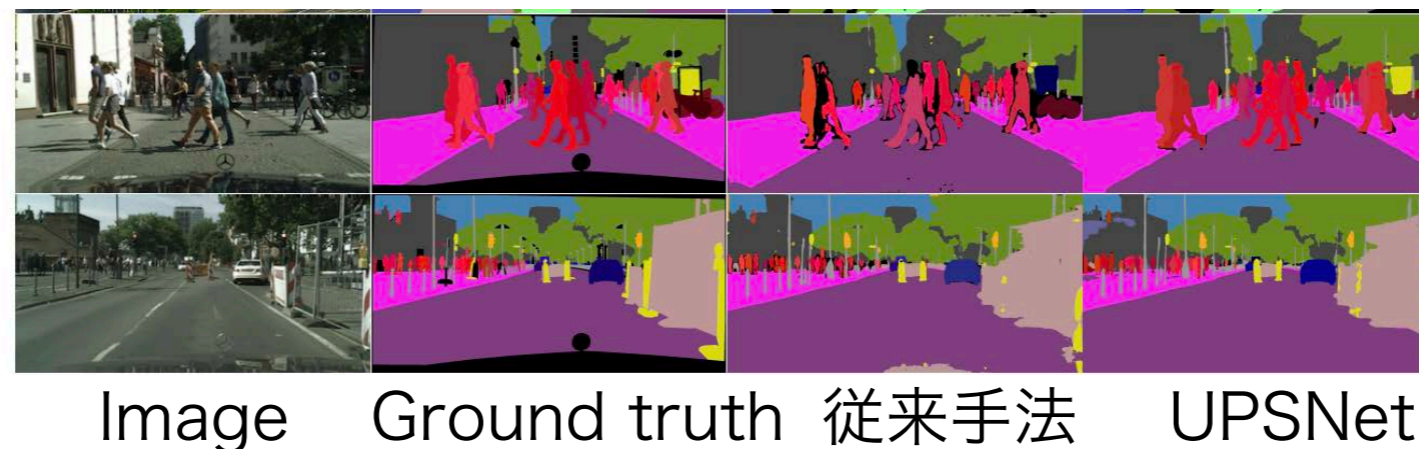
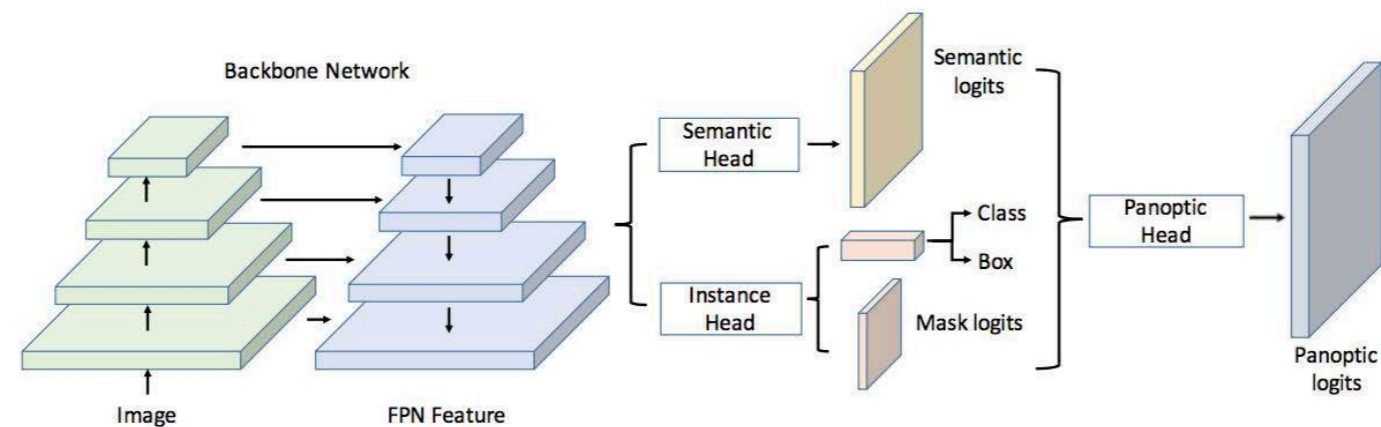
- エンコーダ・デコーダ構造を導入
- Atrous Conv.にDepthwise Convolutionを利用





# UPSNNet: A Unified Panoptic Segmentation Network [Xiong+, ArXiv2019]

- ベースネットワーク：Feature Pyramid Network
- デコーダの各特徴マップからセマンティック/インスタンスセグメンテーション
- Panoptic部で統合



# 大規模データセット


## IMAGENET



<http://image-net.org>



<http://places.csail.mit.edu>

 **COCO**  
Common Objects in Context

info@cocodataset.org

**Home** People Dataset Tasks Evaluate

### What is COCO?



COCO is a large-scale object detection, segmentation, and captioning dataset. COCO has several features:

- ✓ Object segmentation
- ✓ Recognition in context
- ✓ Superpixel stuff segmentation
- ✓ 330K images (>200K labeled)
- ✓ 1.5 million object instances
- ✓ 80 object categories
- ✓ 91 stuff categories
- ✓ 5 captions per image
- ✓ 250,000 people with keypoints

### Collaborators

- Tsung-Yi Lin Cornell Tech
- Genevieve Patterson MSR
- Matteo R. Ronchi Caltech
- Yin Cui Cornell Tech
- Michael Maire TTI-Chicago
- Serge Belongie Cornell Tech
- Lubomir Bourdev WaveOne, Inc.
- Ross Girshick FAIR
- James Hays Georgia Tech
- Pietro Perona Caltech
- Deva Ramanan CMU
- Larry Zitnick FAIR

### Sponsors



<http://cocodataset.org>

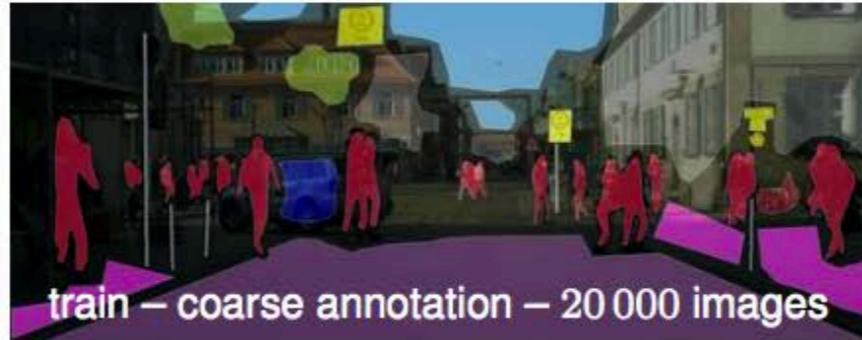
## Open Images Dataset V5 + Extensions

15,851,536 boxes on 600 categories  
2,785,498 instance segmentations on 350 categories  
36,464,560 image-level labels on 19,959 categories  
391,073 relationship annotations of 329 relationships  
Extension - 478,000 crowdsourced images with 6,000+ categories

<https://storage.googleapis.com/openimages/web/index.html>



5000枚の高精細なセグメンテーションデータセット

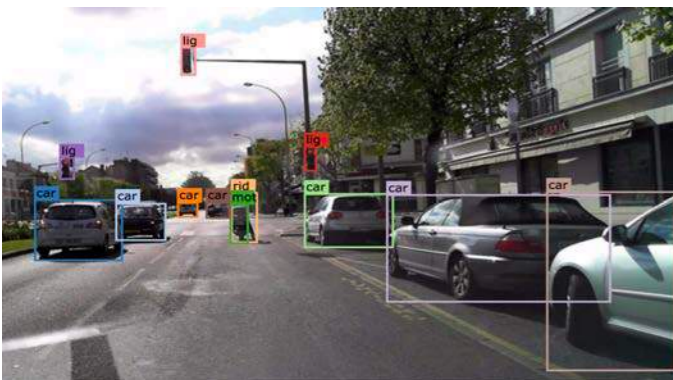


<https://www.cityscapes-dataset.com/>

M. Cordts, The Cityscapes Dataset for Semantic Urban Scene Understanding, CVPR2016



1100時間以上の運転シーン(IMU, GPSあり)



<http://bdd-data.berkeley.edu>

セグメンテーションラベル, 物体検出用バウンディングボックス  
白線, 走行可能領域などのアノテーションあり

- 中部大学・MPRGで保有している環境



GPU搭載ワークステーション：97台  
(Quadro P5000)



AIスーパーコンピュータ DGX-1  
GPU搭載サーバなど GPU数：>100

アルゴリズム

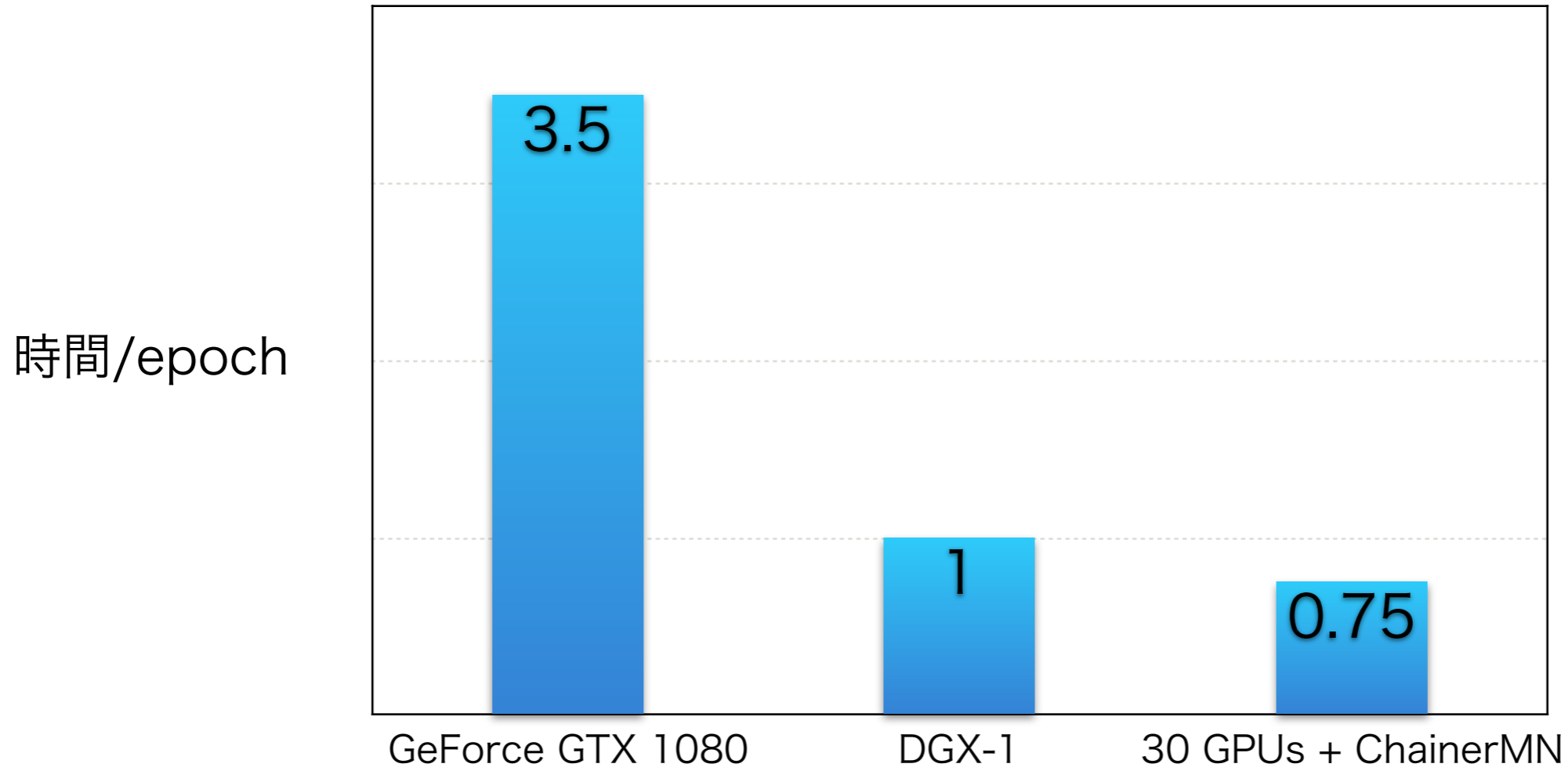
**最新技術をキャッチアップ**

データセット

**大量の良質なデータを用意**

計算機リソース

**設備投資・クラウド活用**

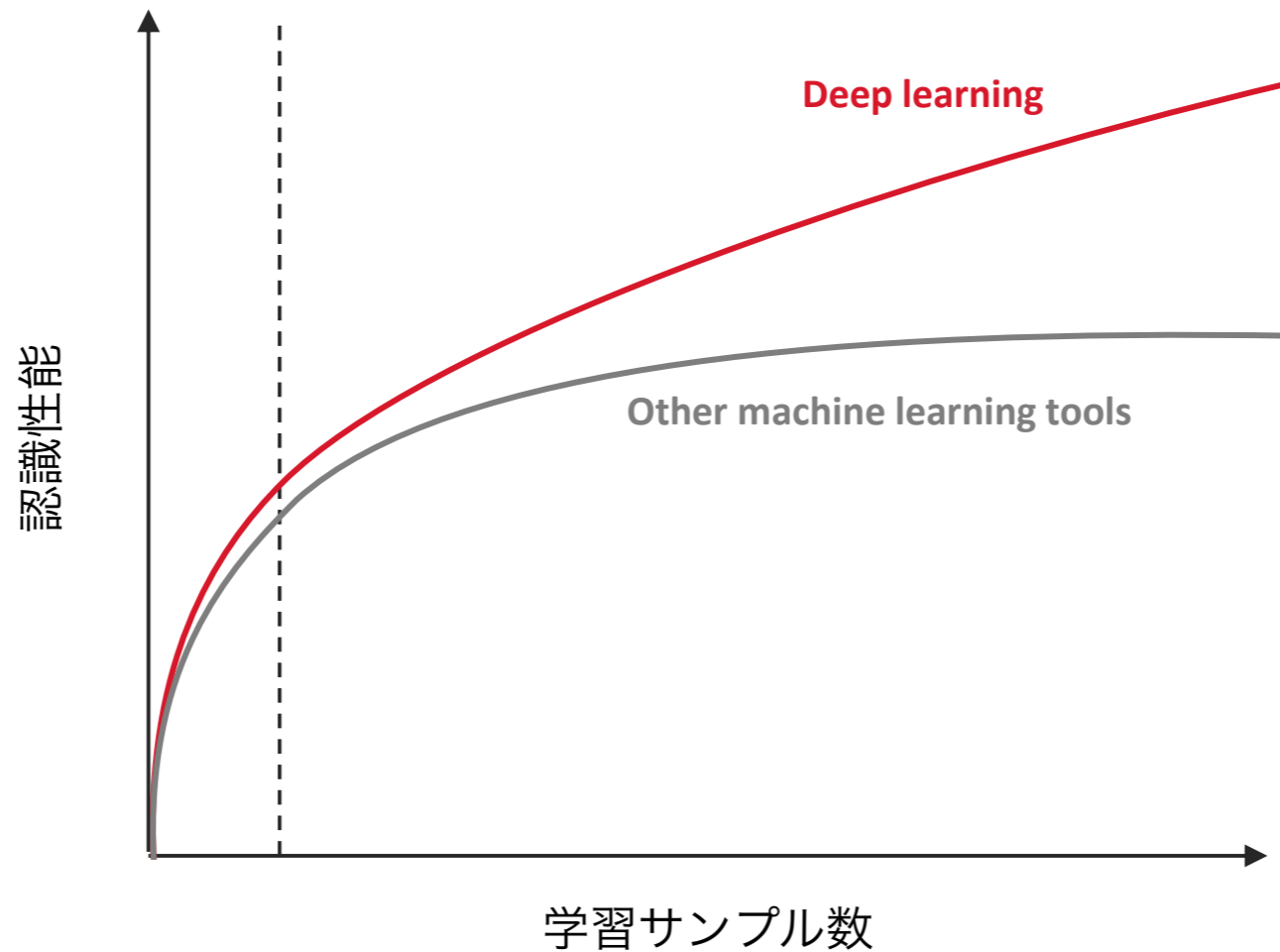


Quadro P5000

150 epoch → 525 時間

150 時間

112 時間



## 深層学習による画像認識

- 学習により特徴表現を自動獲得
- 学習データ数が増えるほど認識性能を向上

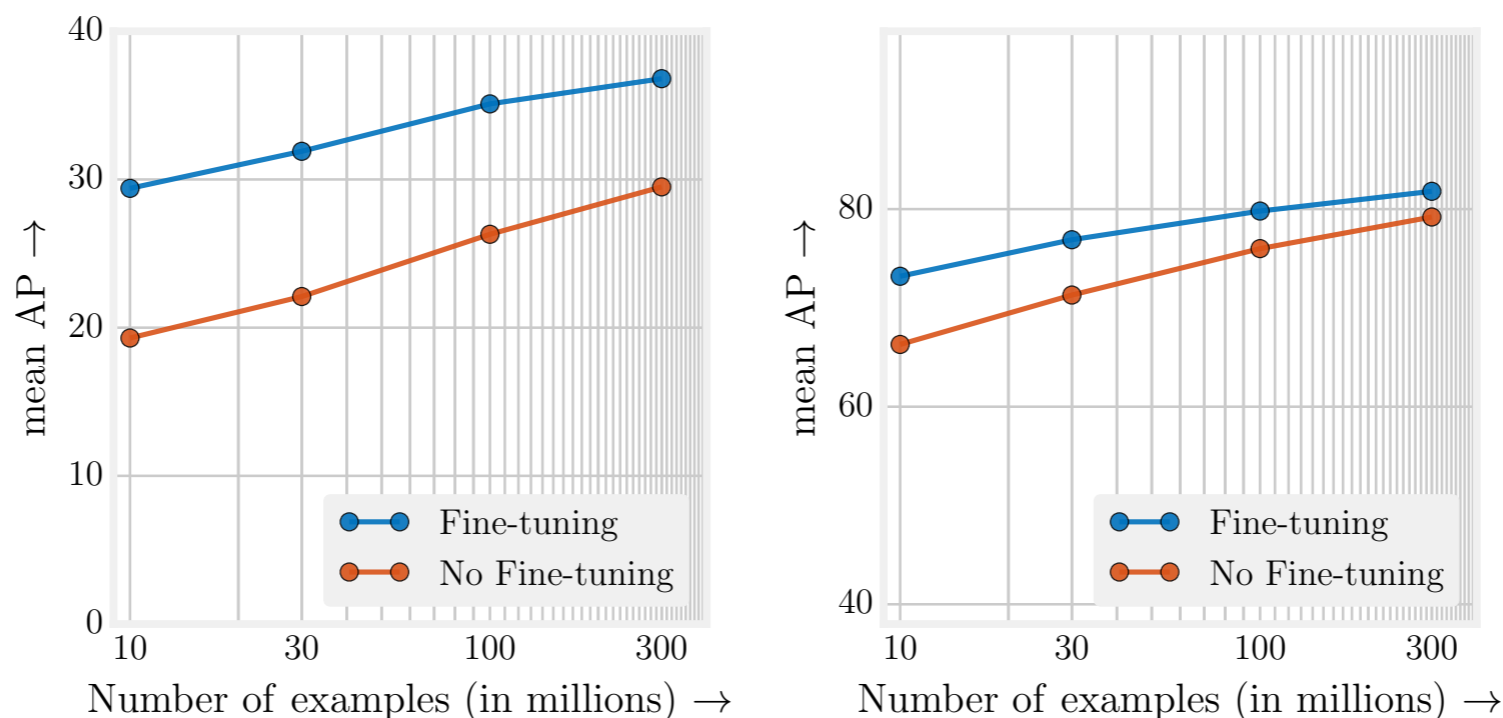


Figure 4. Object detection performance when initial checkpoints are pre-trained on different subsets of JFT-300M from scratch. x-axis is the data size in log-scale, y-axis is the detection performance in mAP@[.5,.95] on COCO minival\* (left), and in mAP@.5 on PASCAL VOC 2007 test (right).

ImageNet : 100万枚→訓練データを3億枚に増加

50GPUで2ヶ月かけ学習した結果, 精度はデータ数の対数に比例して向上

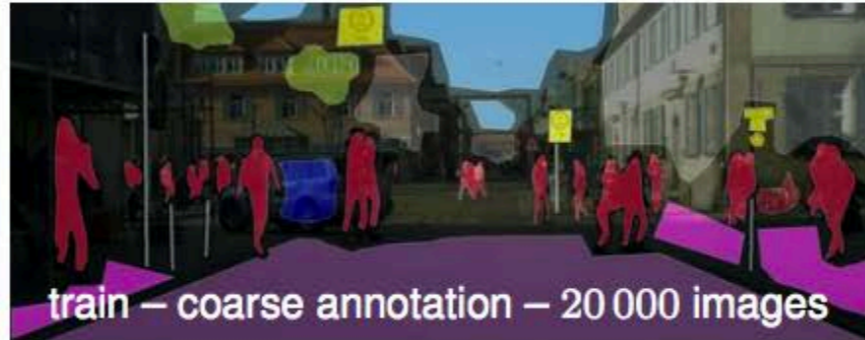


# セマンティックセグメンテーション : MNet [山下+,ICIP'18] MPRG



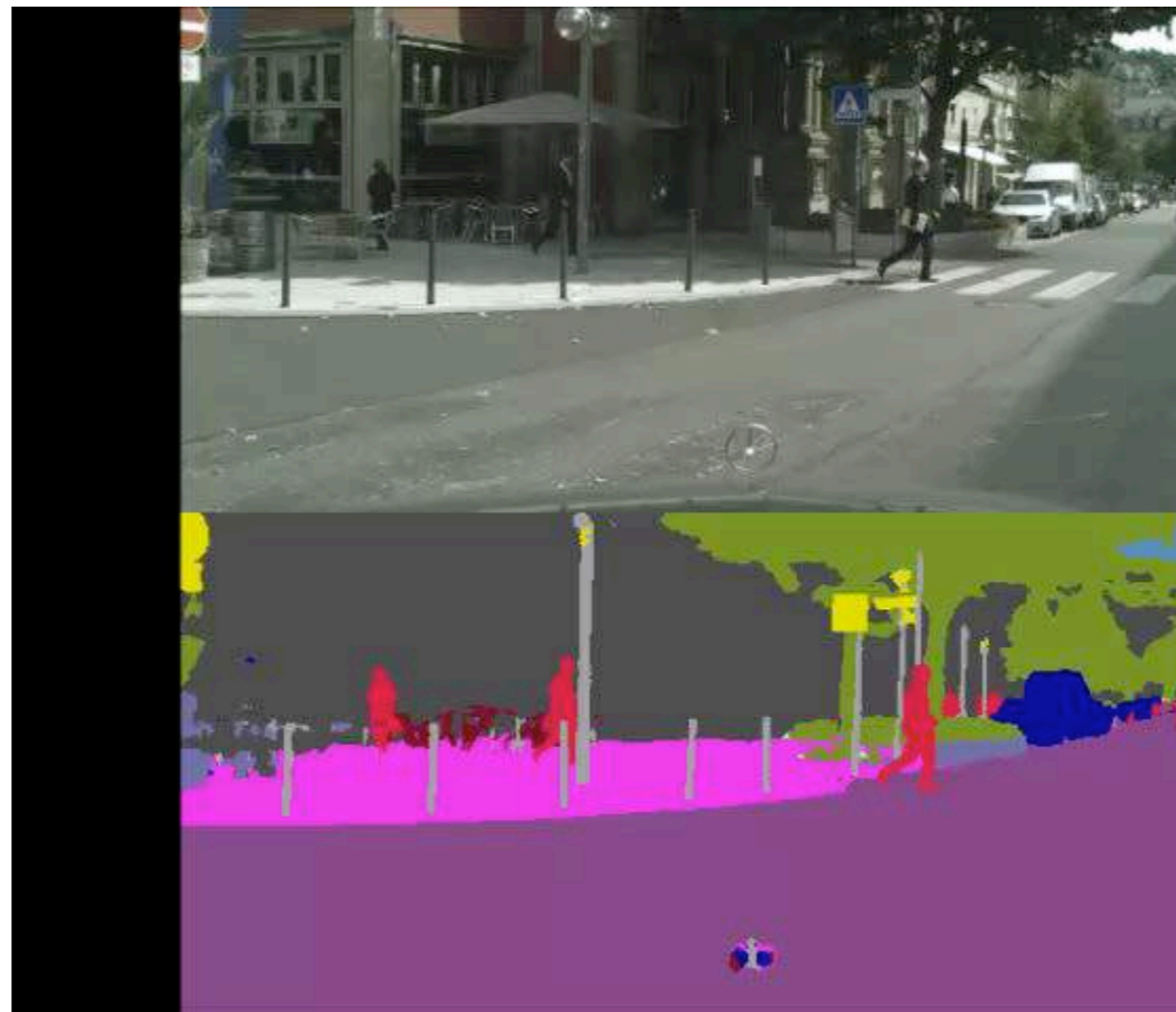
**CITYSCAPES**  
DATASET

5000枚の高精細なセグメンテーションデータセット



<https://www.cityscapes-dataset.com/>

M. Cordts, The Cityscapes Dataset for Semantic Urban Scene Understanding, CVPR2016



- human
- car
- motorcycle
- bicycle
- road
- sidewalk
- pole
- traffic sign
- traffic signal
- sky
- vegetation
- terrain

MPRG  
MACHINE PERCEPTION AND ROBOTICS GROUP





## データ+正解情報

アノテーション作業（正解情報を付与する作業）が重要



学習サンプル数 : **11,960** 枚

総アイテム数 : **137** 種類

(既知アイテム数 : 40, 未知アイテム数 : 97)

人手によるアノテーション : **48,590個** → 約1週間

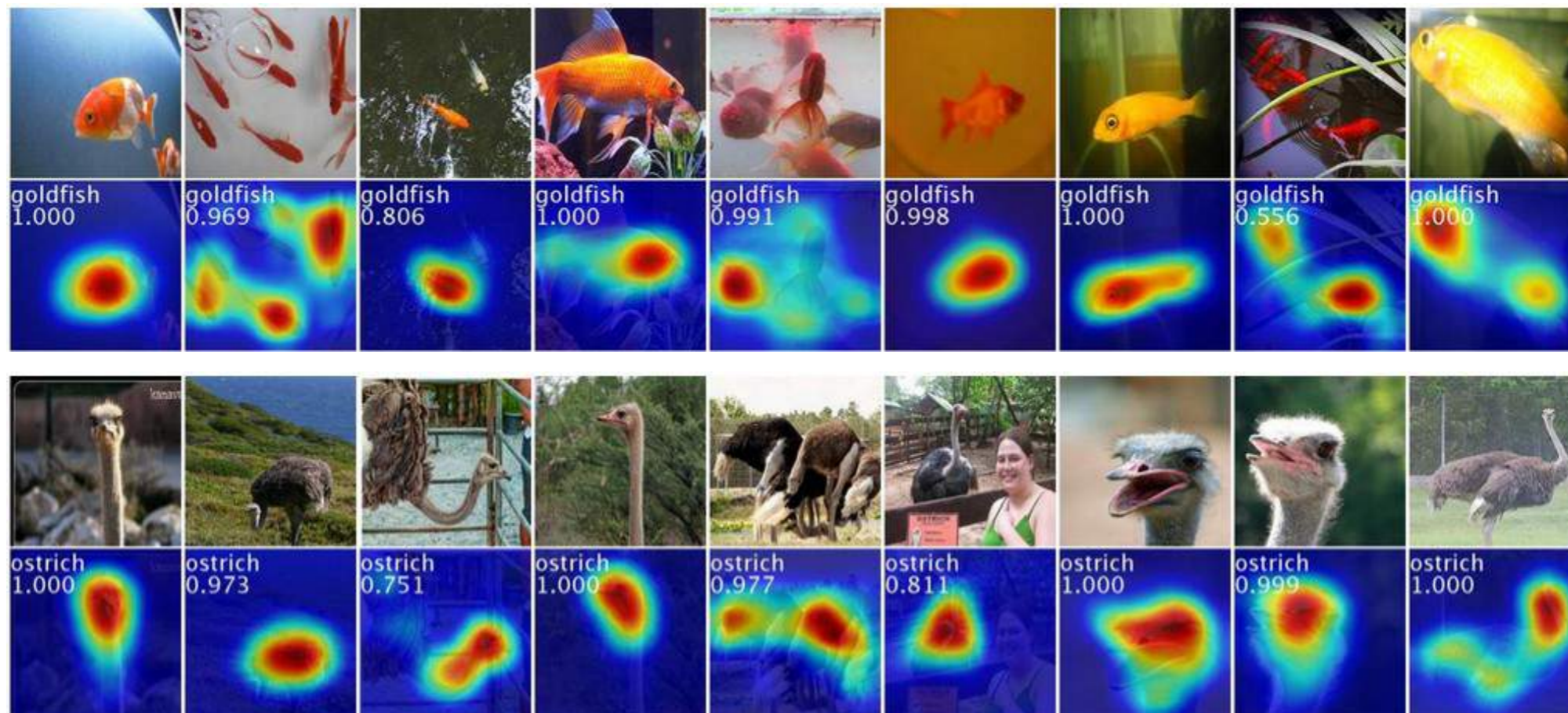
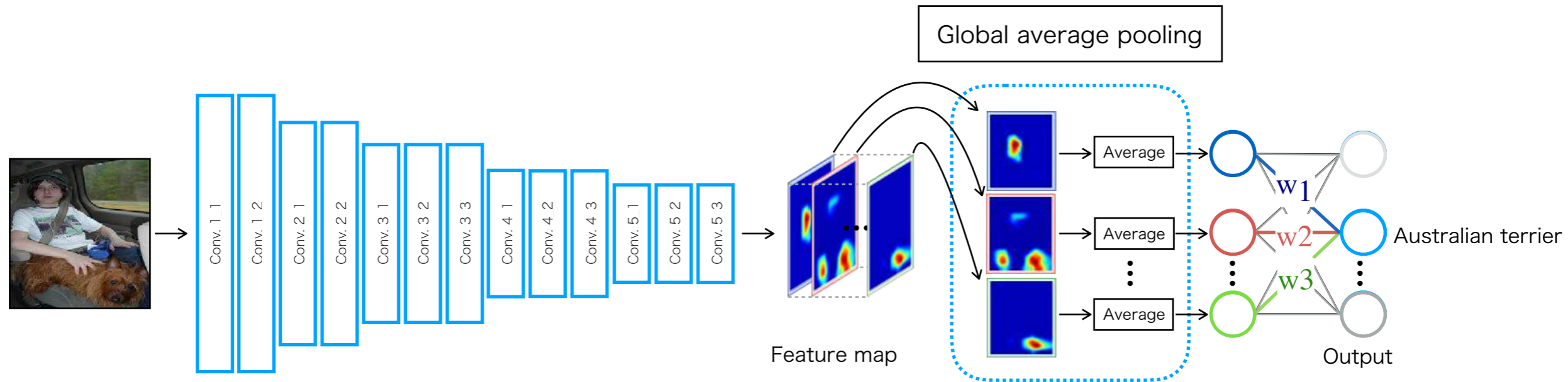


- ・ アノテーションミス
  - 主観的で定量的でないアノテーション
- ・ アノテーションコスト
  - 学習に必要なデータ数の見積もりが不可能
- ・ データの偏り
  - 発生頻度が極端に異なる異常データ
- ・ データの経年変化
  - 工程の変動、素材やラインによる変動

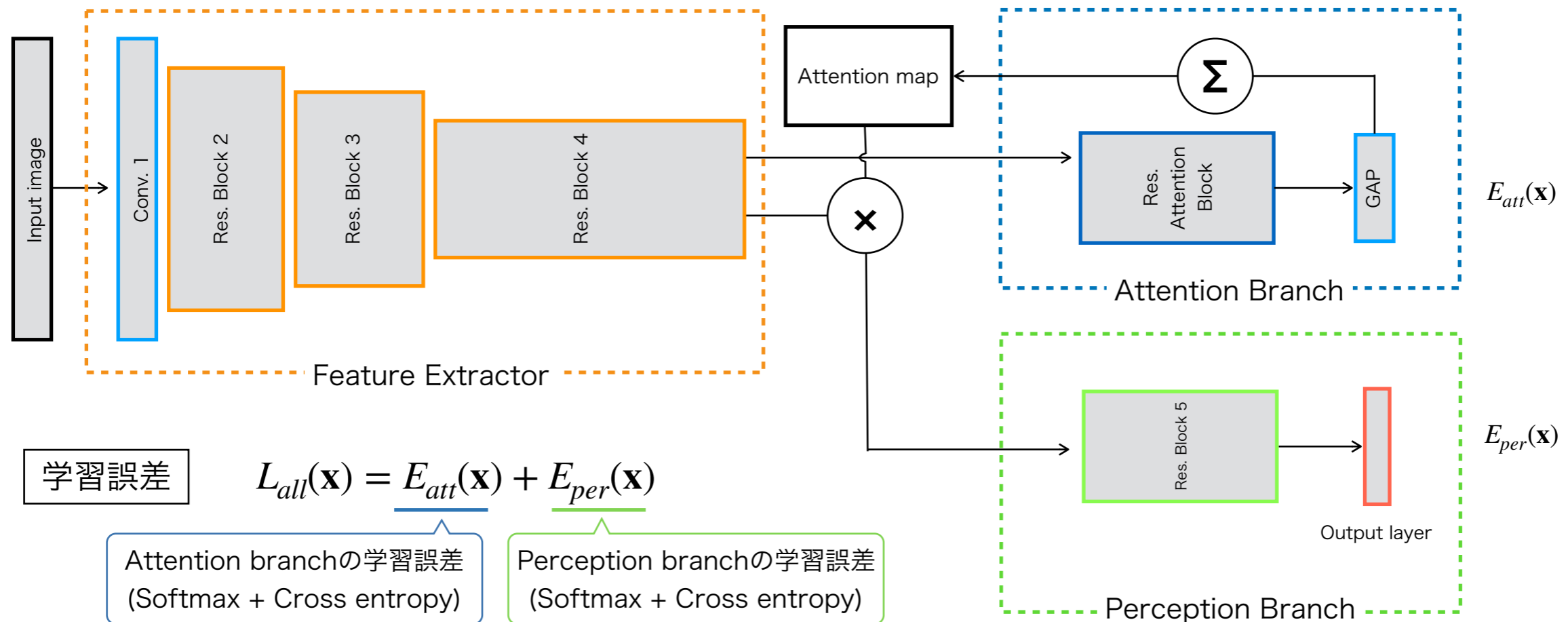
**ディープラーニングは何を見ている？**



- Class Activation Mapping(CAM) [Zhou+, CVPR'16]

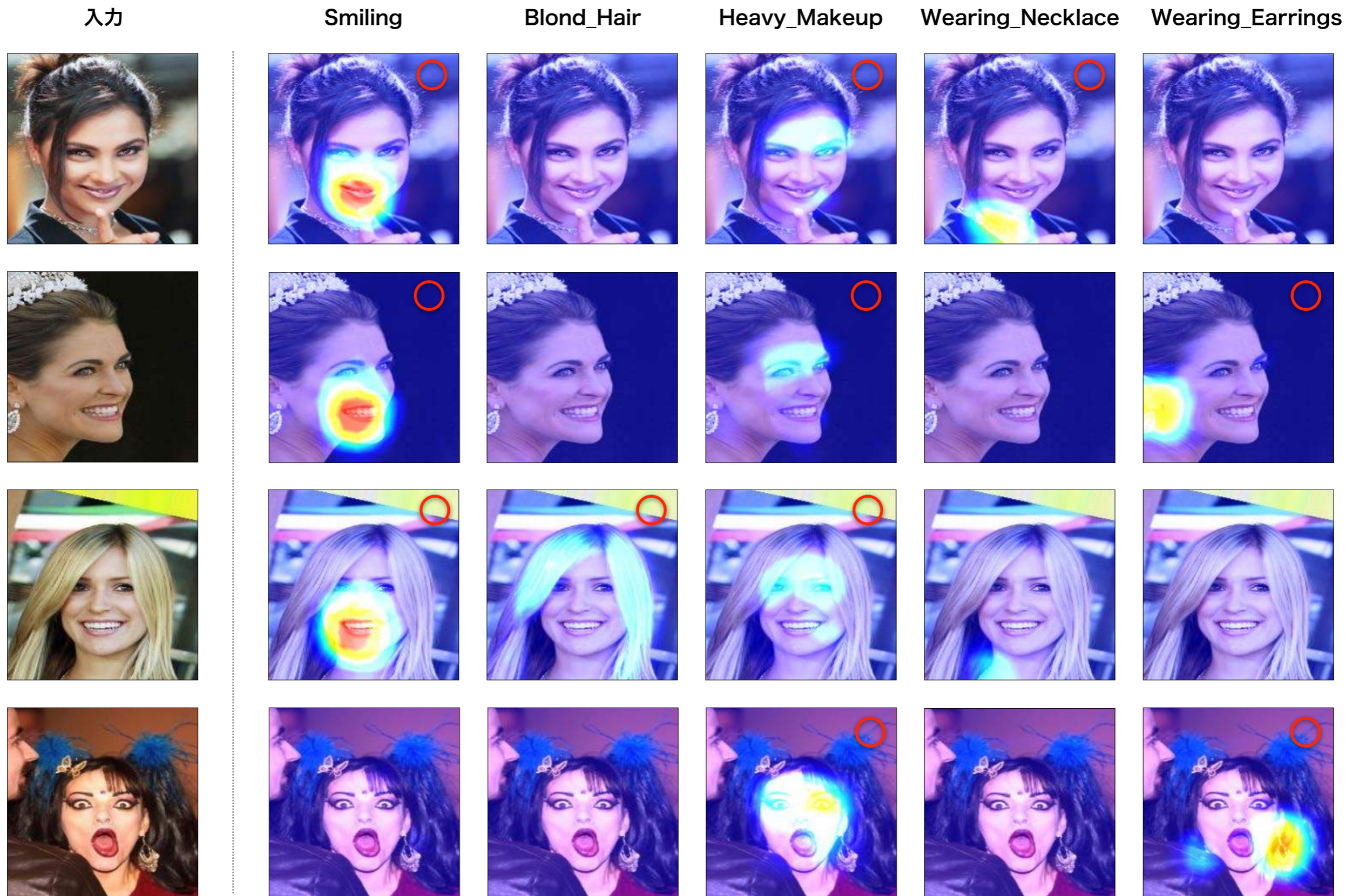


- GAPからAttention mapを生成するネットワーク
  - Feature extractor : 特徴マップを出力
  - **Attention Branch** : Attention mapを生成
  - **Perception Branch** : Attention mapを用いて最終的な推論結果を出力



→Attention BranchとPerception BranchをEnd-to-endで学習

# 顔属性を認識する時の判断根拠



## ものづくり

- ・自動車
- ・産業ロボット
- ・家電
- ・農業機械
- ・建設機械
- ・医療機器

...

×

ディープラーニング  
「眼の技術」

=

## 作業の自動化

- ・自動運転
- ・組立加工の自動化
- ・外観検査の自動化
- ・建設作業の自動化
- ・農作業の自動化
- ・画像診断の自動化
- ・物流の自動化
- ・介護の自動化
- ・陳列廃棄の自動化

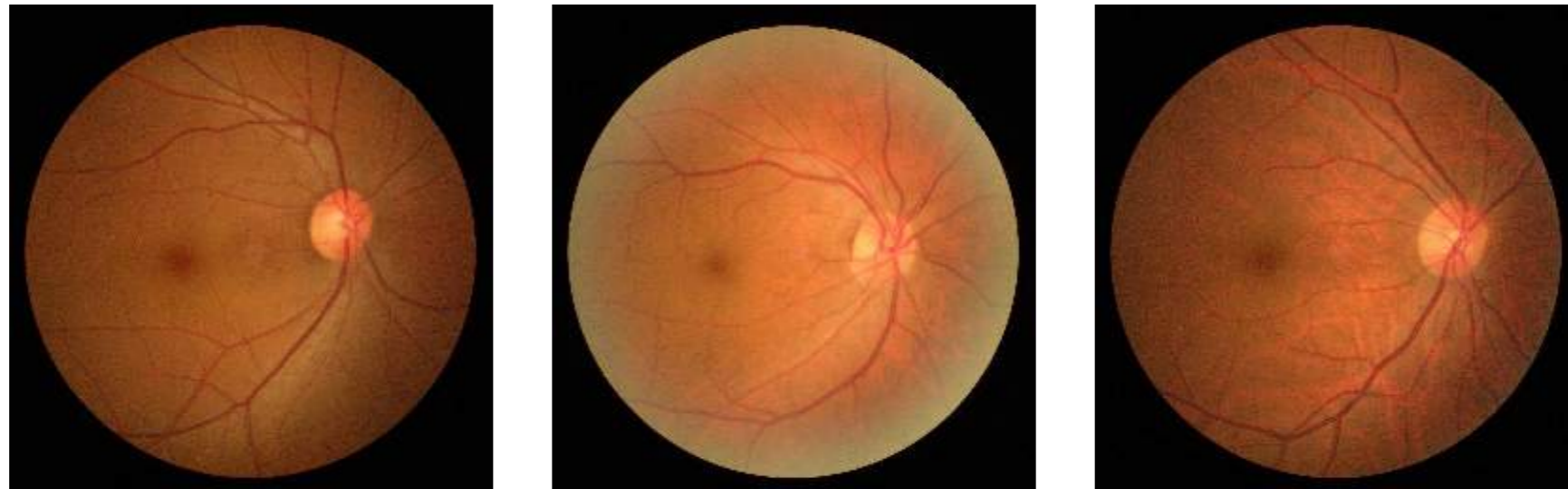
...

MPRGの取り組み：

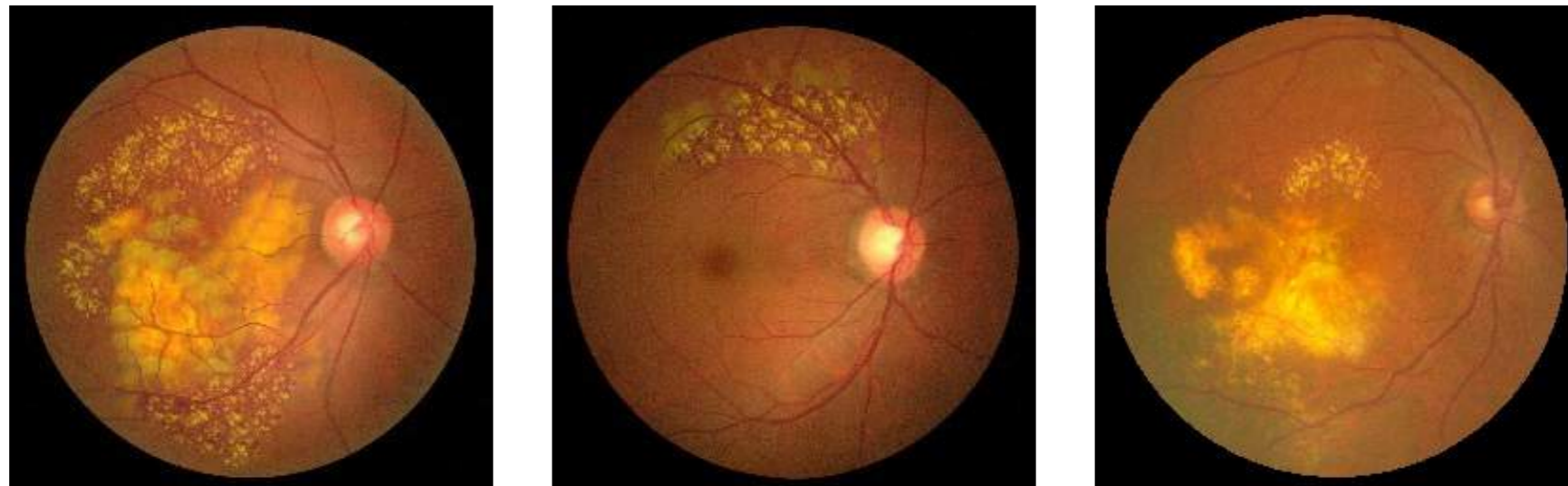
画像診断における判断根拠の獲得

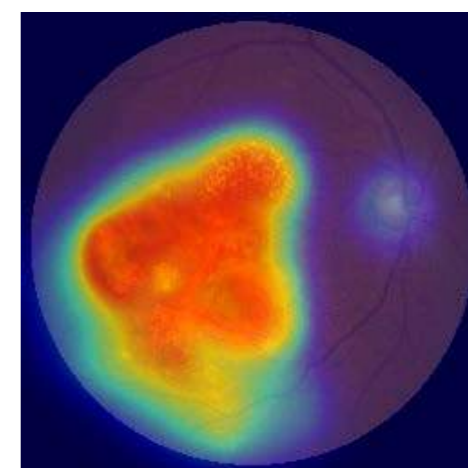
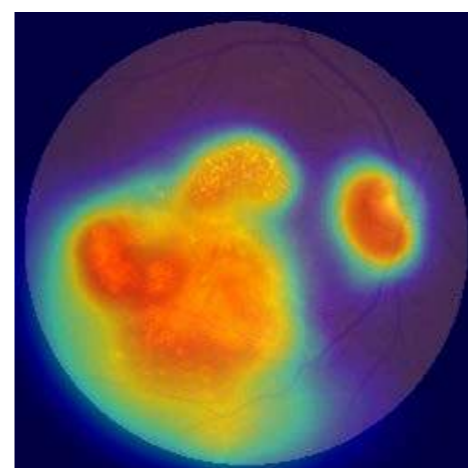
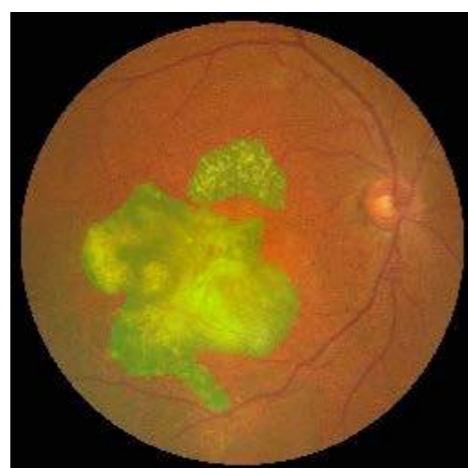
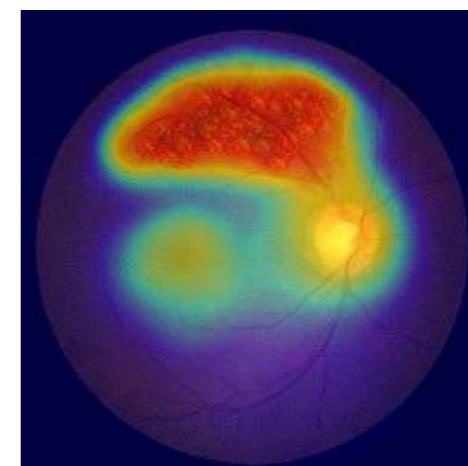
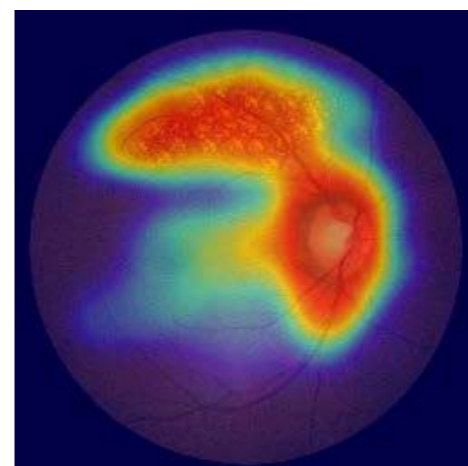
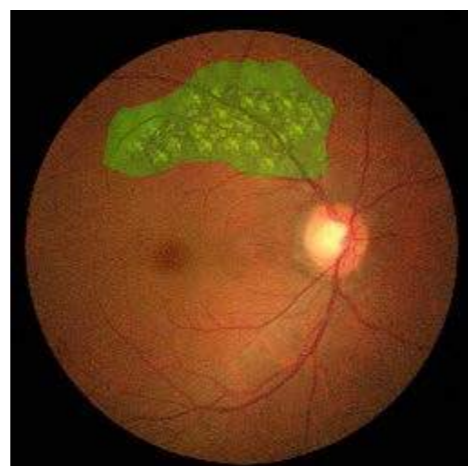
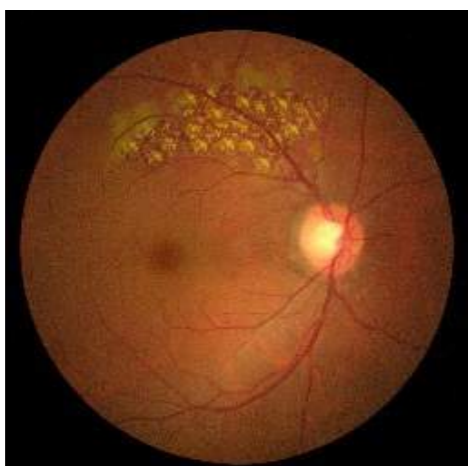
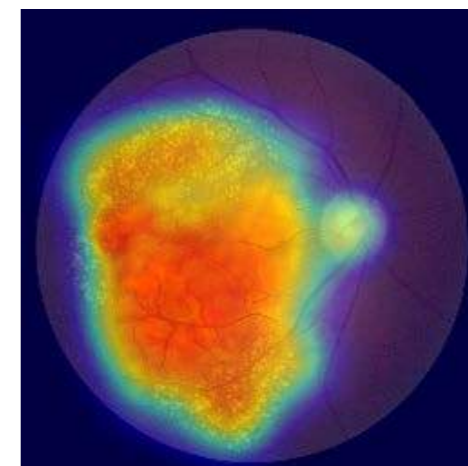
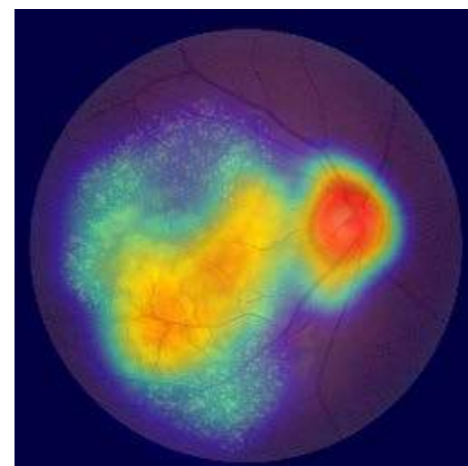
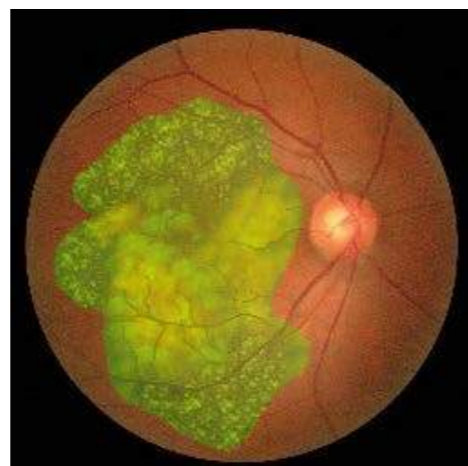
- 眼底画像からの疾患判別
  - ポリープ状脈絡膜血管症：加齢黄斑変性 (AMD) によって起こる脈絡膜部分の疾患

正常



疾患





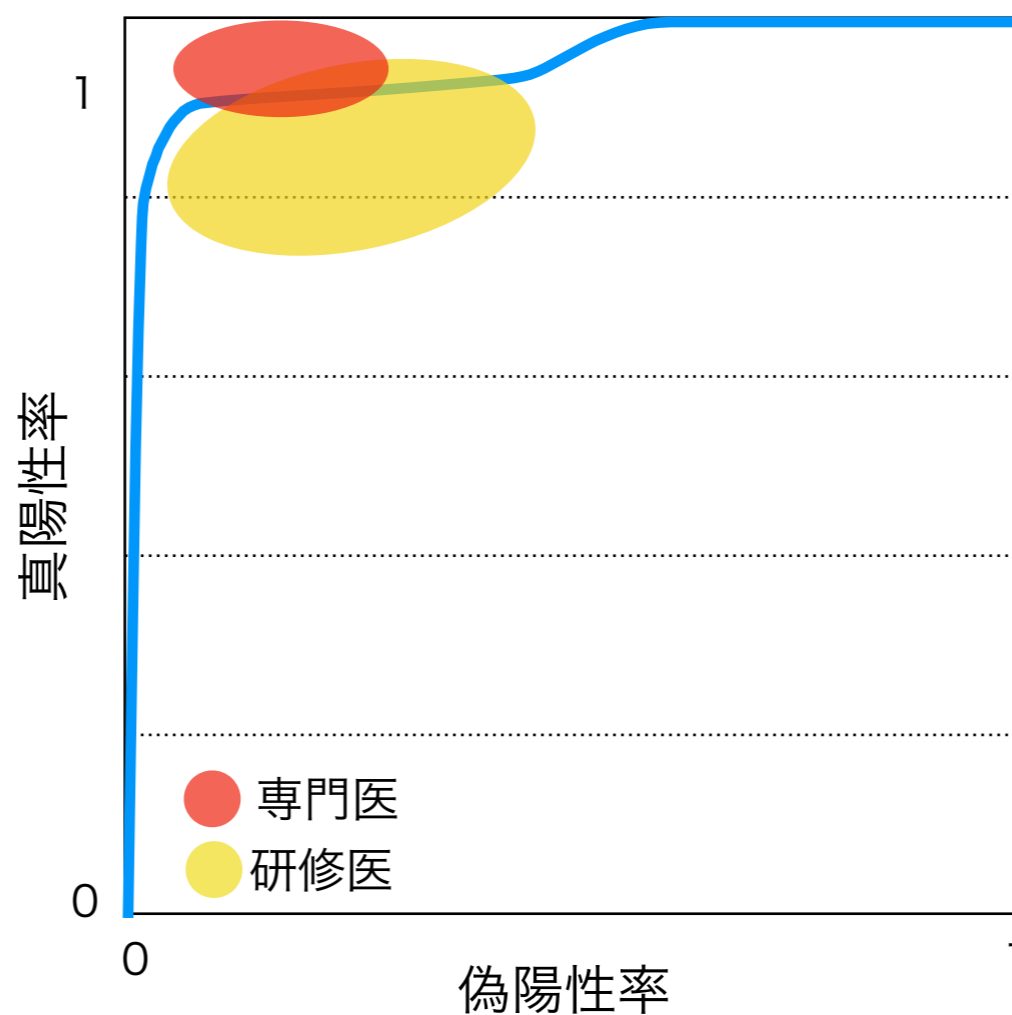
疾患眼底画像

専門家

従来手法

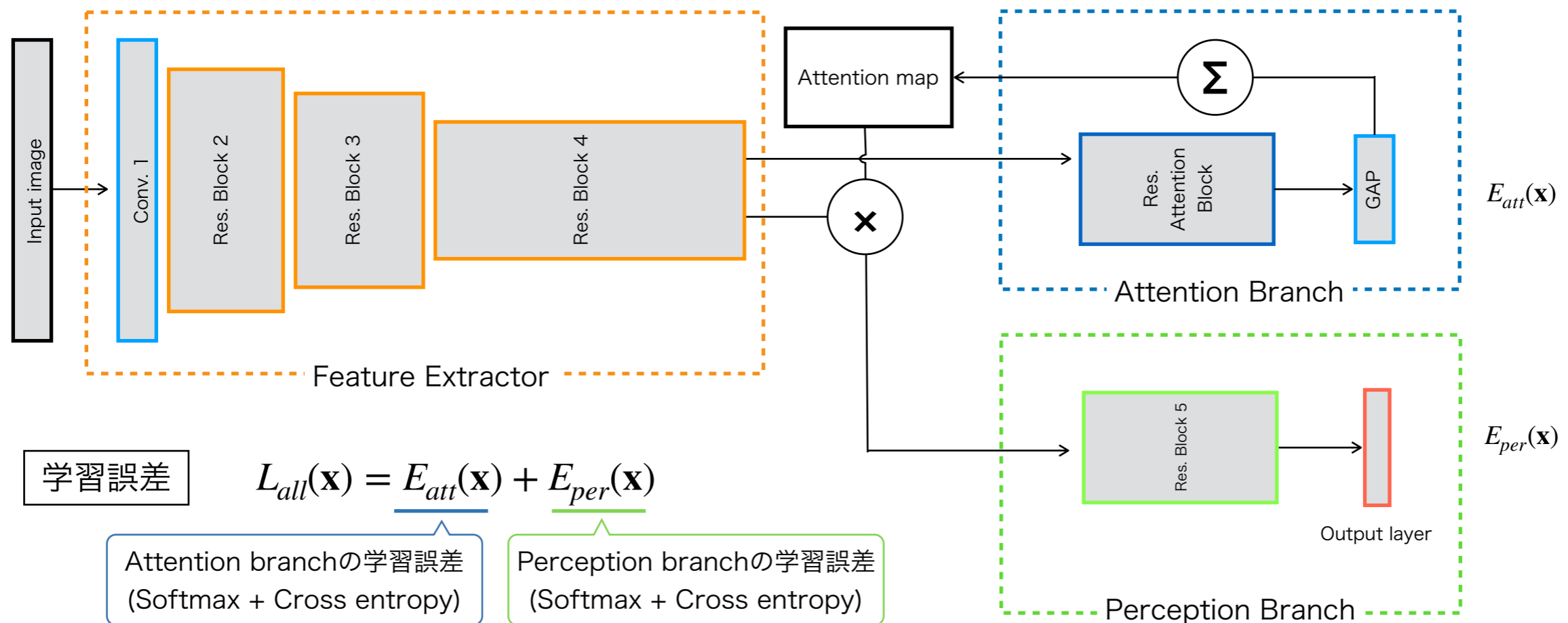
ABN(提案手法)

- 眼底画像からの疾患判別
  - Attention Branch Network なし AUC : ~0.94
  - Attention Branch Network あり AUC : 0.95~



専門医 > AI (提案手法ABN) > 研修医

- GAPからAttention mapを生成するネットワーク
  - Feature extractor : 特徴マップを出力
  - **Attention Branch** : Attention mapを生成
  - **Perception Branch** : Attention mapを用いて最終的な推論結果を出力



→Attention BranchとPerception BranchをEnd-to-endで学習



## ものづくり

- ・ 自動車
- ・ 産業ロボット
- ・ 家電
- ・ 農業機械
- ・ 建設機械
- ・ 医療機器
- ...

×

ディープラーニング  
「眼の技術」

=

## 作業の自動化

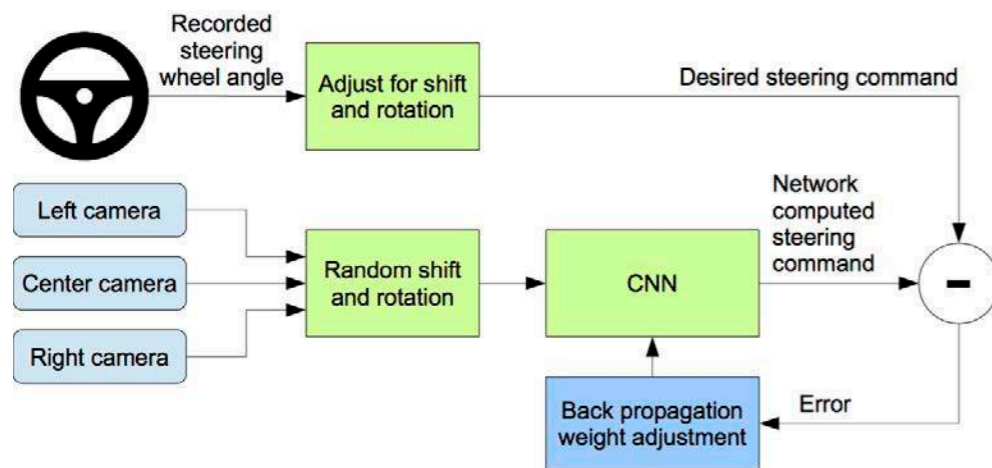
- ・ 自動運転
- ・ 組立加工の自動化
- ・ 外観検査の自動化
- ・ 建設作業の自動化
- ・ 農作業の自動化
- ・ 画像診断の自動化
- ・ 物流の自動化
- ・ 介護の自動化
- ・ 陳列廃棄の自動化
- ...

MPRGの取り組み：

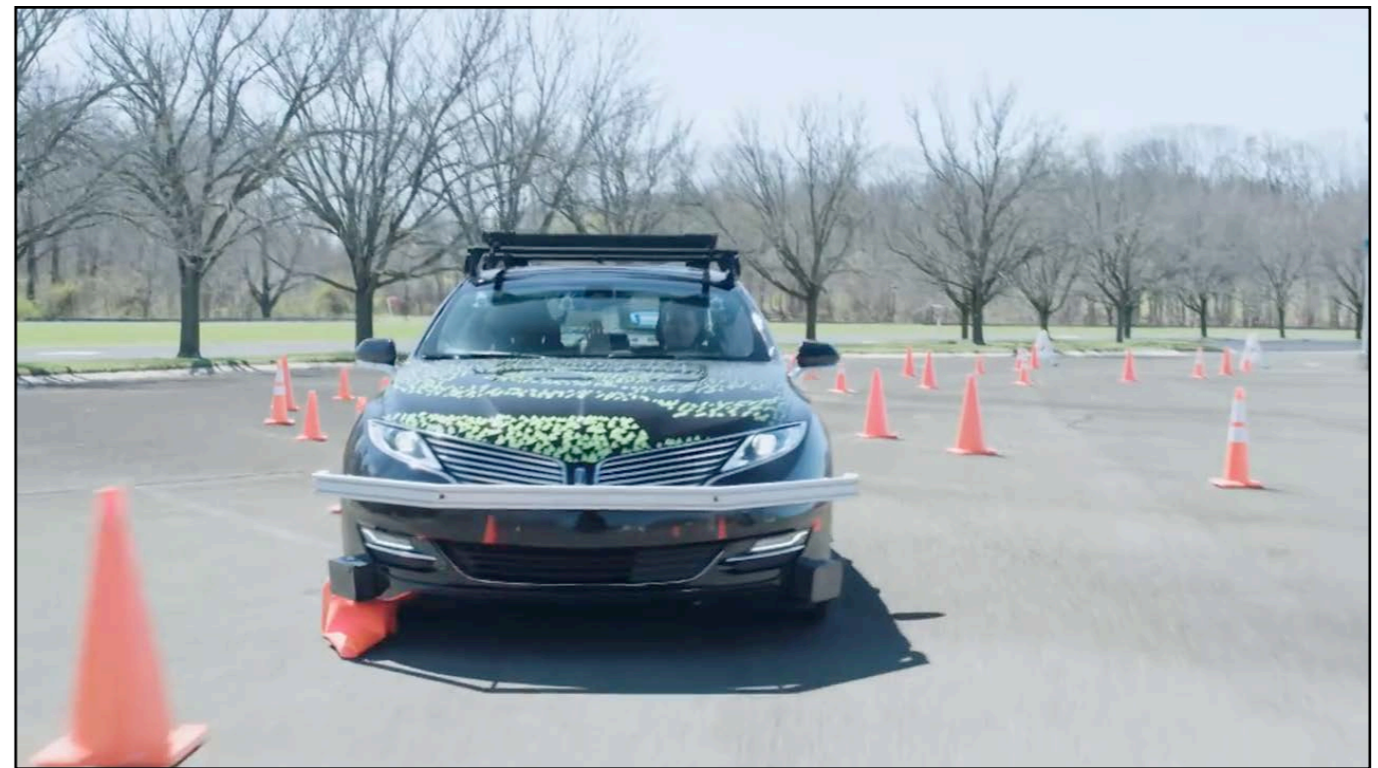
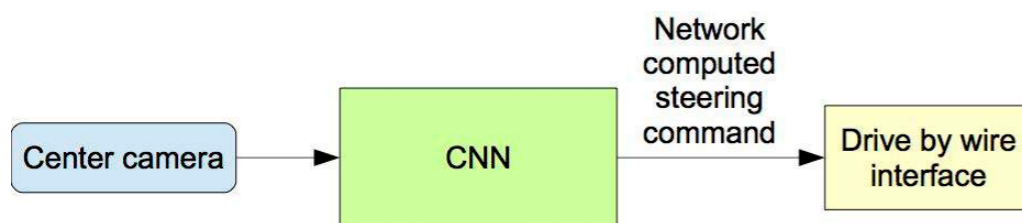
自動運転における判断根拠の獲得

- End-to-End学習による自動運転
  - NVIDIA Self-Driving Cars [Borjaski+, arXiv'16]

## ニューラルネットの学習



## 学習済みのニューラネットが操舵角を出力

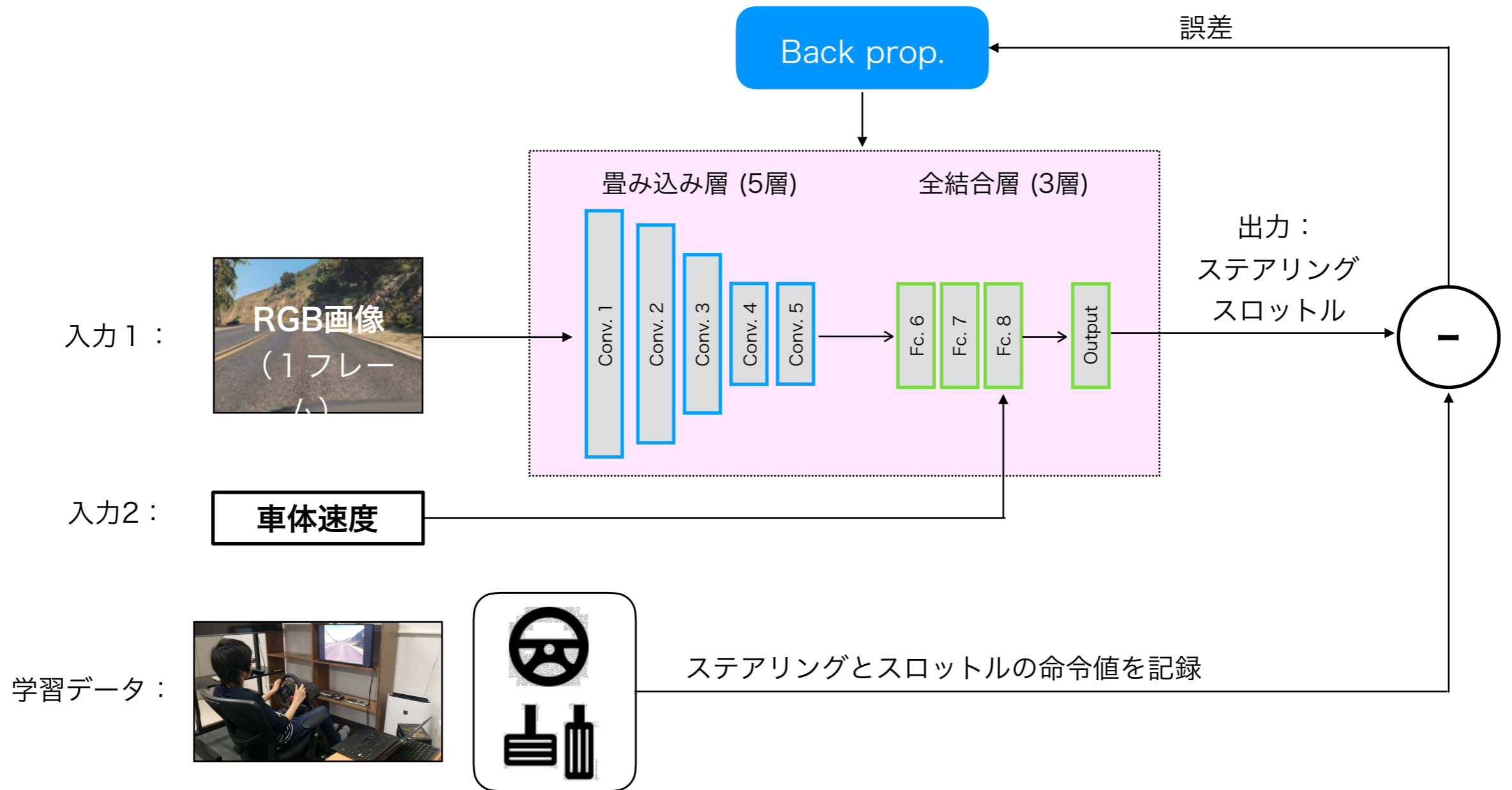


<https://www.youtube.com/watch?v=qhUvQiKec2U>

- ・ 人が実際にGrand Theft Auto V (GTAV)を走行してデータを収集
  - 取得データ：フレーム, ステアリング, スロットル, 車体速度



- 画像と車体速度を入力しステアリングとスロットルの制御値を回帰



# 一貫学習による自動走行例（上り坂）

カメラ映像のみを入力



```
('steer = ', 0.0)  
( 'throttle = ', 0.0)
```

自己状態を付与

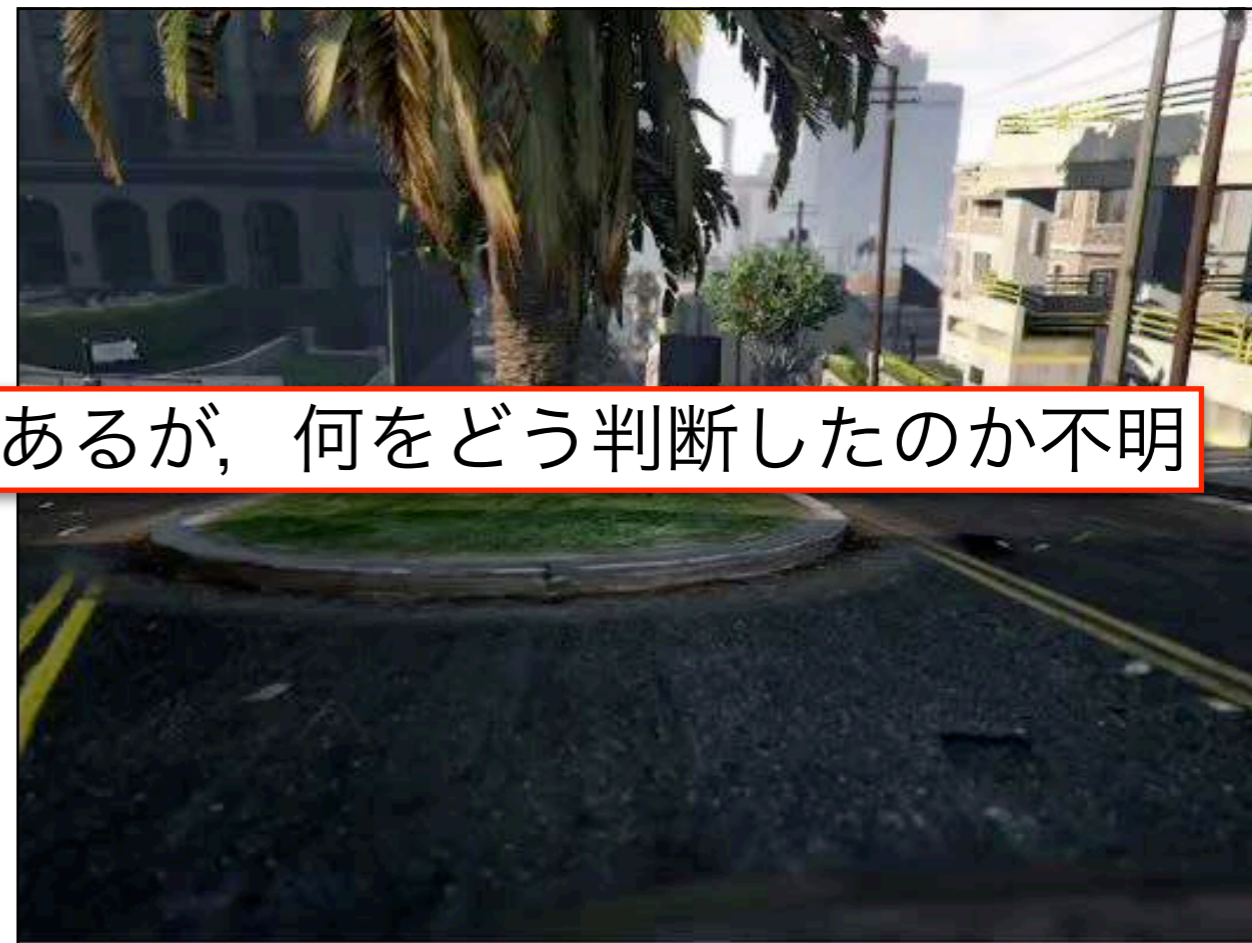


```
('steer = ', 0.0)  
( 'throttle = ', 0.0)
```

# 一貫学習による自動走行例（下り坂）

カメラ映像のみを入力

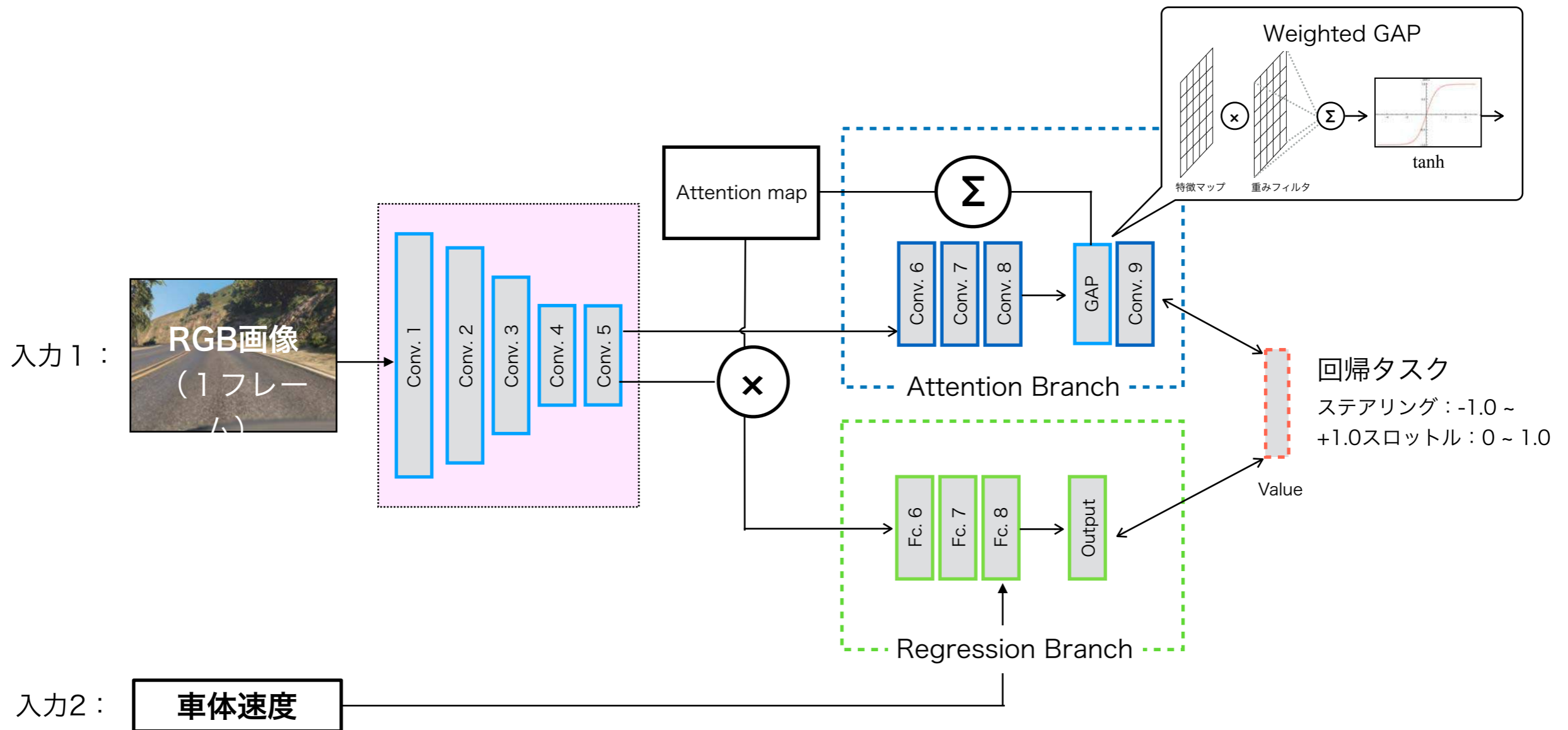
自己状態を付与



```
('steer = ', -0.048290835971592366)  
( 'throttle = ', 0.0)
```

```
('steer = ', 0.0)  
( 'throttle = ', 0.0)
```

- 回帰推定に対応したAttention branchを導入してABNを構築
  - 特徴マップに重み(カーネル)を畳み込んで連続値に変換



→Attention Branch, Regression Branchを回帰タスクとして学習

# ABNの適用による一貫学習ネットワークの解釈



入力画像



Attention map



# カーブ（ステアリング時）におけるアテンション



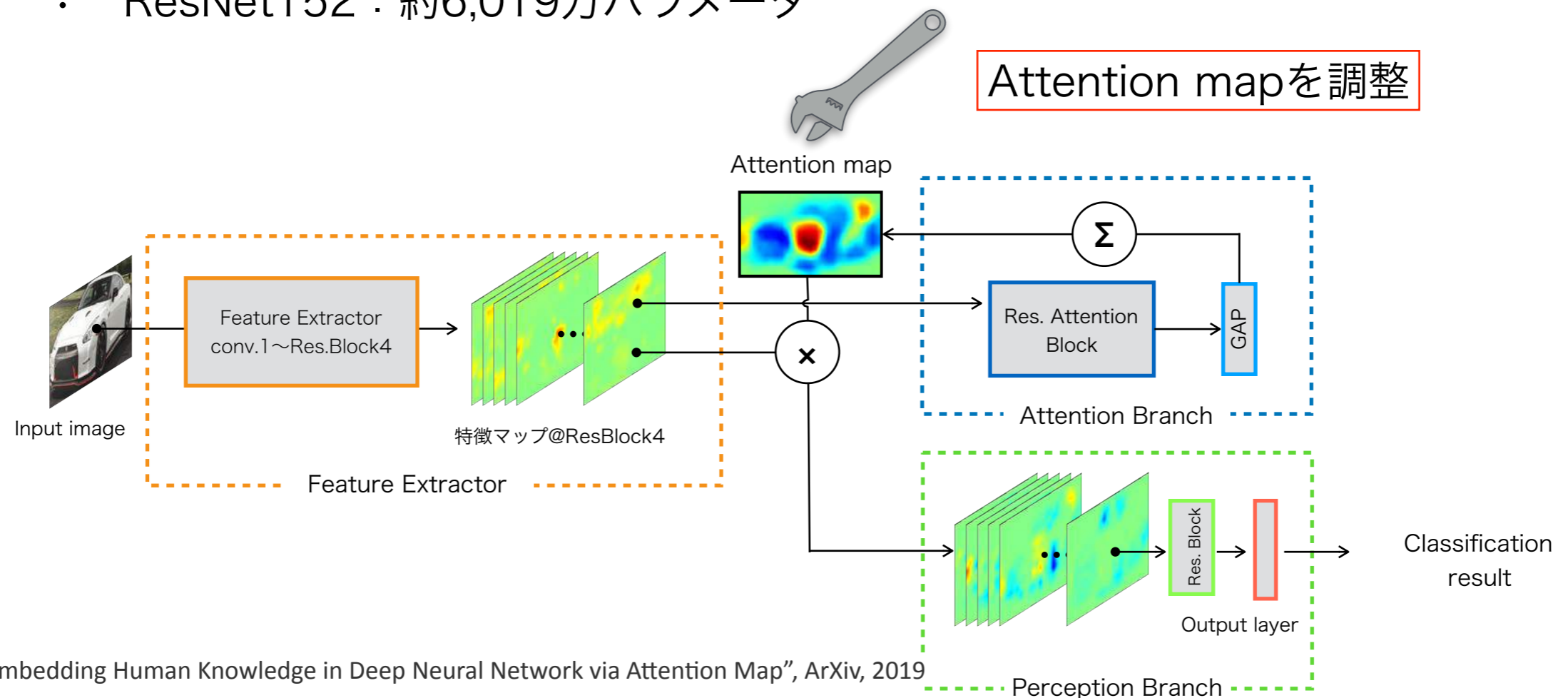
ステアリング : 0.18    スロットル : 0.07



ステアリング : -0.04    スロットル : 0.12

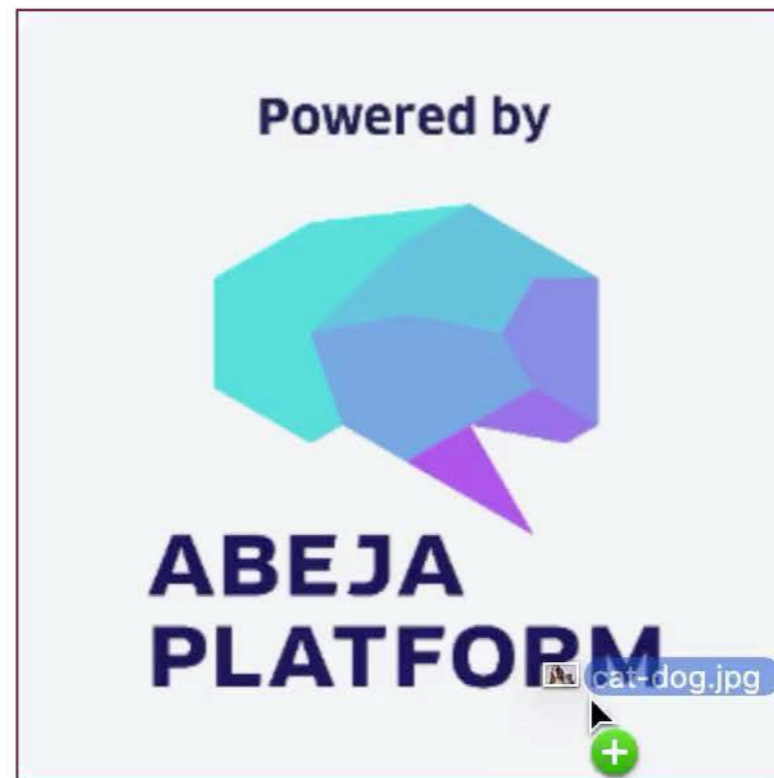
→ ステアリングと反対方向の車線に沿ってアテンションが発生

- ・ 誤認識したテストサンプルをうまく認識できるようにしたい
- ・ DNNのパラメータを手動で調整
  - ・ パラメータ数が多大なので、何をどのように変更すれば良いか
    - ・ AlexNet：約6,110万パラメータ
    - ・ VGG：約1.4億パラメータ
    - ・ ResNet152：約6,019万パラメータ



# 手動によるAttention mapの調整 (アプリ)

## Attention Editor



Choose File No file chosen

Upload Image Upload Mask Clear Mask change attention mode undo

Top 5 Results

index	name	score
-------	------	-------

# Attention mapの修正による誤認識率の削減

ImageNet Dataset (validation) の誤認識率の変化

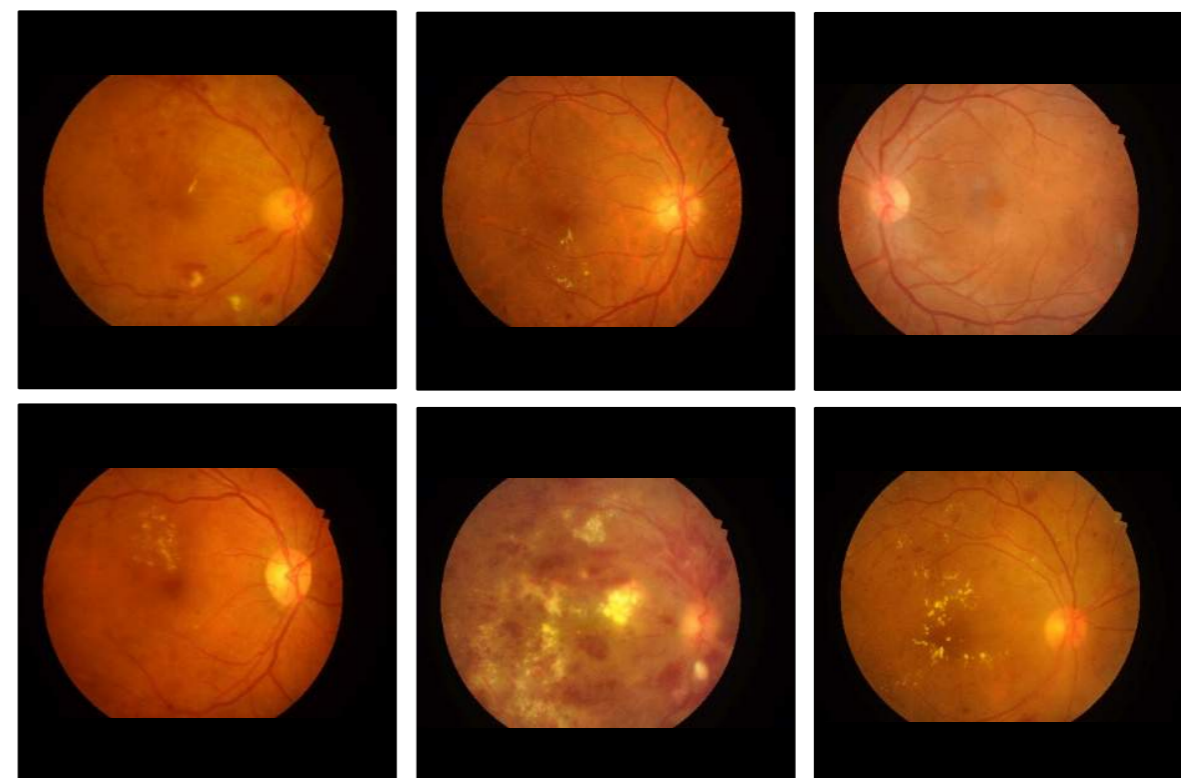
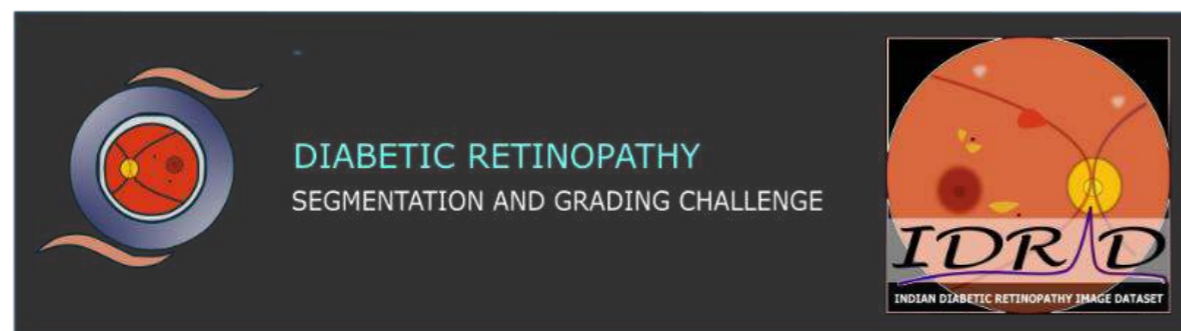
修正したサンプル数	1,000枚	
	Top-1 error [%]	Top-5 error [%]
修正前	100.0	19.0
修正後	83.2	15.8

-16.8%

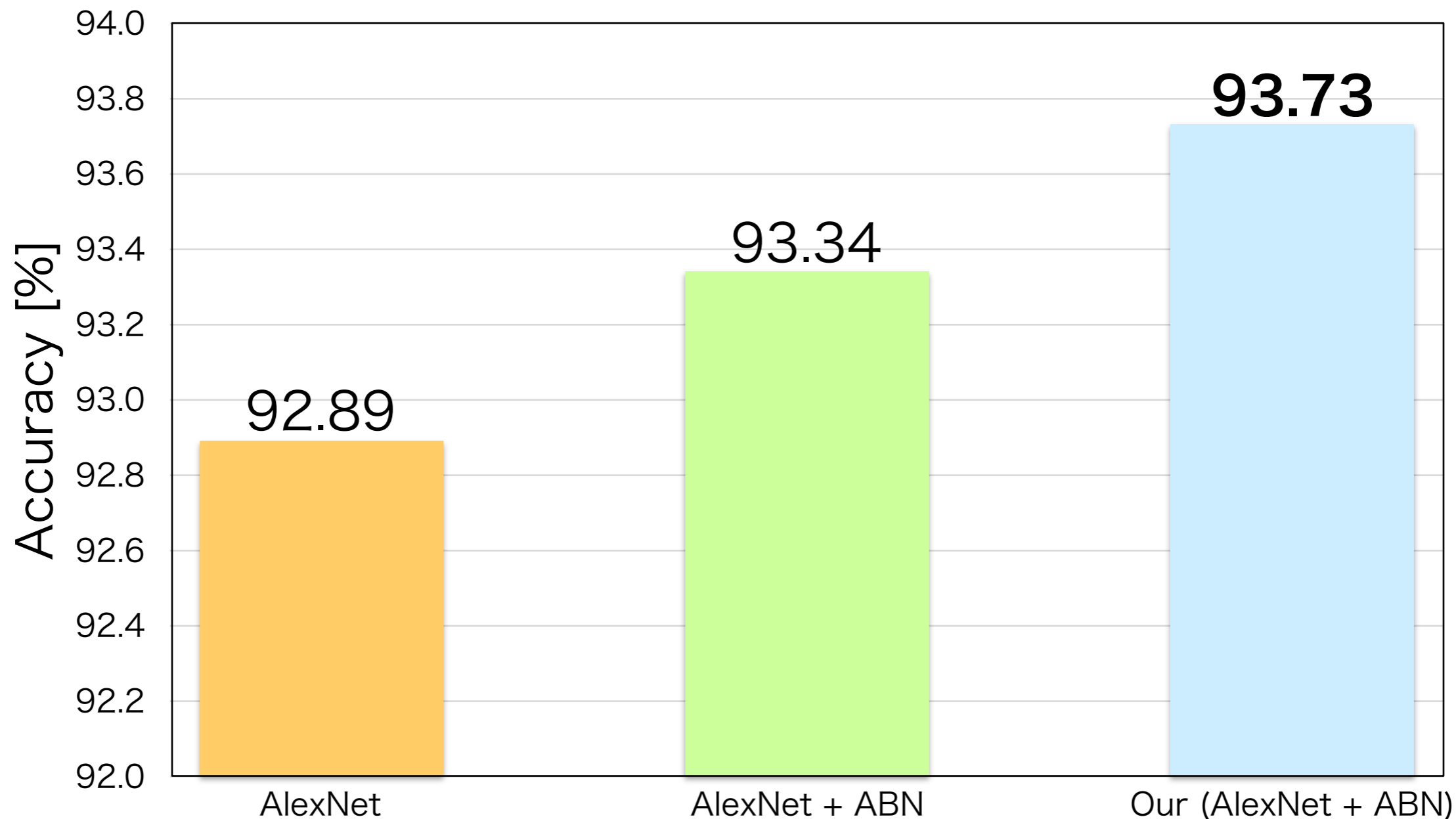


Attention mapの修正により認識結果の改善が可能

- INDIAN DIABETIC RETINOPATHY IMAGE DATASET (IDRiD) による評価
  - 眼底画像による疾患のグレード識別を対象としたデータセット
  - サンプル数：81

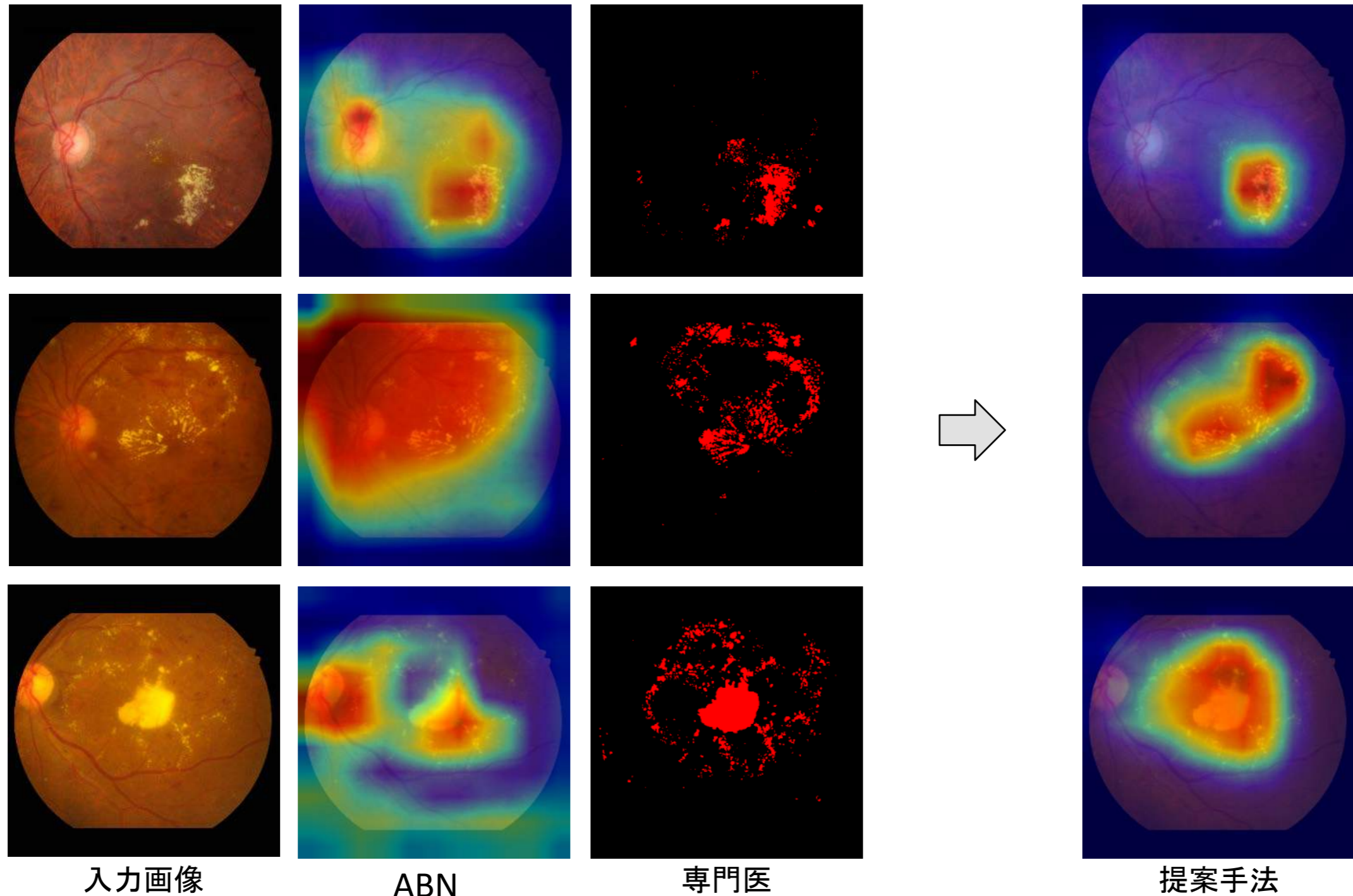


提案手法の導入により認識性能が向上



IDRiDの認識率の比較

# 提案手法とABNのAttention mapの比較



疾患領域のみに注視領域が反応

- ・ ディープラーニング協会
  - 理事長：松尾豊（東大）
  - 資格・検定を開催（年2-3回）



The screenshot shows the homepage of the Japan Deep Learning Association (JDLA). At the top, there is a navigation bar with links for '協会について', '資格試験', '資料室', '推薦図書', '会員について', and '新着情報'. Below the navigation bar is a large banner image with a neural network background. In the center of the banner, there is a white box containing the JDLA logo and the text: '一般社団法人日本ディープラーニング協会', 'Japan Deep Learning Association', and '本協会は、ディープラーニングを中心とする技術による日本の産業競争力の向上を目指します。' At the bottom of the page, there is a section for '新着情報' (NEWS & EVENTS).

## G 検定 ジェネラリスト

ディープラーニングの  
基礎知識を有し、  
適切な活用方針を決定して  
事業応用する能力を持つ人材

## E 資格 エンジニア

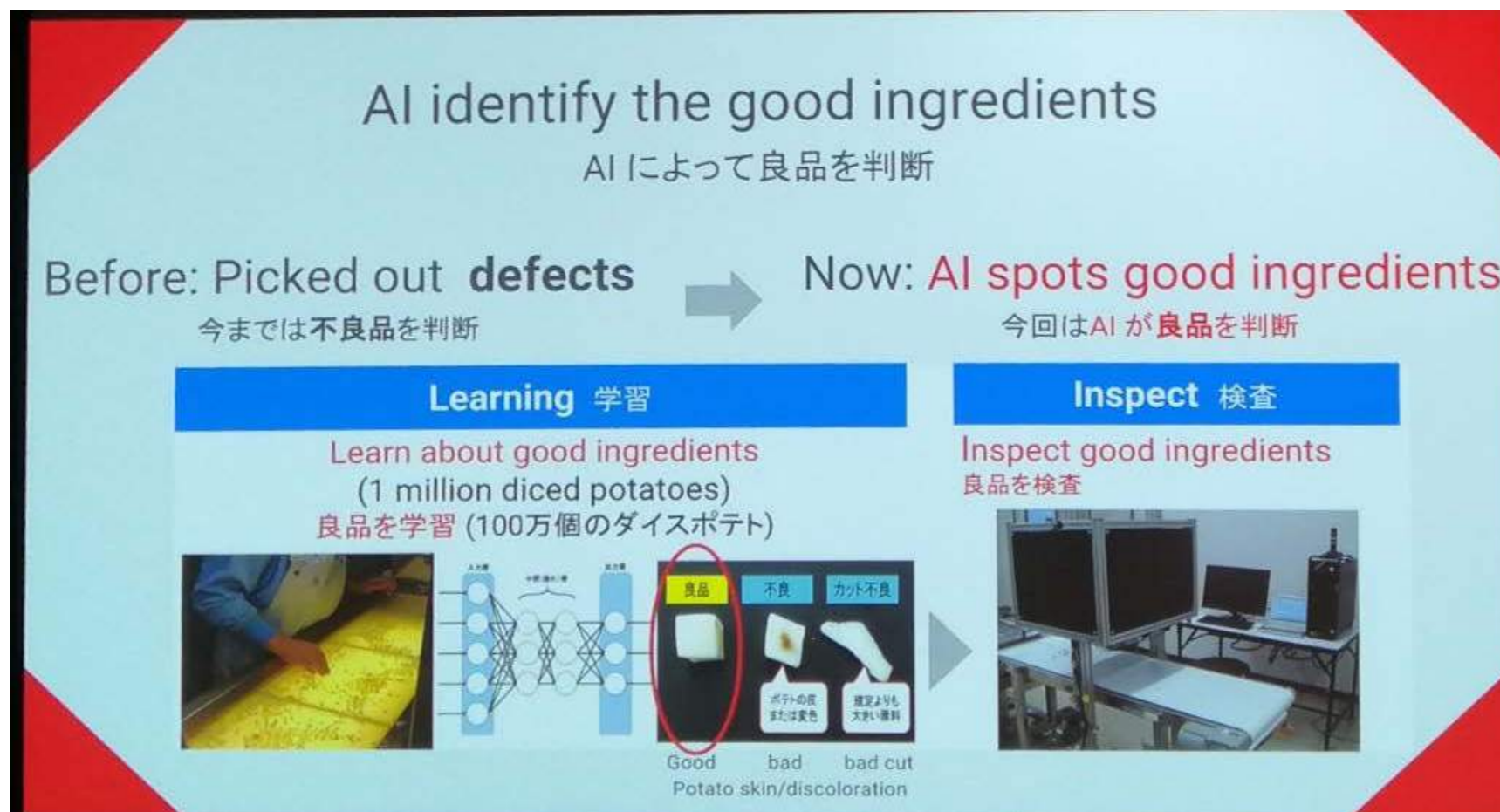
ディープラーニングの  
理論を理解し、  
適切な手法を選択して  
実装する能力を持つ人材



- ディープラーニング活用の教科書 実践編
  - 編：日経クロストレンド
  - 実用段階フェーズの26事例
  - ディープラーニングビジネス活用アワード受賞例



- ・ キューピー
  - 良品かどうかを判断（10万枚の事例をもとに）
  - 良品と違うものを不良品として排除



- ・ ディープラーニングにより『眼』となる技術が進化
  - ・ 日本の幅広い「ものづくり」への活用と自動化への期待
  - ・ いかに「良質な」データを集めるかが重要
  - ・ 「何を見ているか」を把握することも重要
- 
- ・ ご質問などは, [takayoshi@isc.chubu.ac.jp](mailto:takayoshi@isc.chubu.ac.jp)

ご静聴ありがとうございました



教授

**藤吉弘巨** Hironobu Fujiyoshi E-mail: [fujiyoshi@isc.chubu.ac.jp](mailto:fujiyoshi@isc.chubu.ac.jp)

1997年 中部大学大学院博士後期課程修了, 1997年 米カーネギーメロン大学ロボット工学研究所Postdoctoral Fellow, 2000年 中部大学工学部情報工学科講師, 2004年 中部大学准教授, 2005年 米カーネギーメロン大学ロボット工学研究所客員研究員(～2006年), 2010年 中部大学教授, 2014年名古屋大学客員教授.

計算機視覚, 動画像処理, パターン認識・理解の研究に従事.

ロボカップ研究賞(2005年), 情報処理学会論文誌CVIM優秀論文賞(2009年), 情報処理学会山下記念研究賞(2009年), 画像センシングシンポジウム優秀学術賞(2010, 2013, 2014年), 電子情報通信学会 情報・システムソサイエティ論文賞(2013年)他



准教授

**山下隆義** Takayoshi Yamashita E-mail: [takayoshi@isc.chubu.ac.jp](mailto:takayoshi@isc.chubu.ac.jp)

2002年 奈良先端科学技術大学院大学博士前期課程修了, 2002年 オムロン株式会社入社, 2009年 中部大学大学院博士後期課程修了(社会人ドクター), 2014年 中部大学講師, 2017年 中部大学准教授.

人の理解に向けた動画像処理, パターン認識・機械学習の研究に従事.

画像センシングシンポジウム高木賞(2009年), 電子情報通信学会 情報・システムソサイエティ論文賞(2013年), 電子情報通信学会PRMU研究会研究奨励賞(2013年)受賞.



特任助教

**平川 翼** Tsubasa Hirakawa E-mail: [hirakawa@mprg.chubu.ac.jp](mailto:hirakawa@mprg.chubu.ac.jp)

2013年 広島大学大学院博士課程前期終了, 2014年 広島大学大学院博士課程後期入学, 2017年 中部大学研究員(～2019年), 2017年 広島大学大学院博士後期課程修了. 2019年 中部大学特任助教. 2014年 独立行政法人日本学術振興会特別研究員DC1. 2014年 ESIEE Paris客員研究員(～2015年).

コンピュータビジョン, パターン認識, 医用画像処理の研究に従事.