
物体検出手法の解説

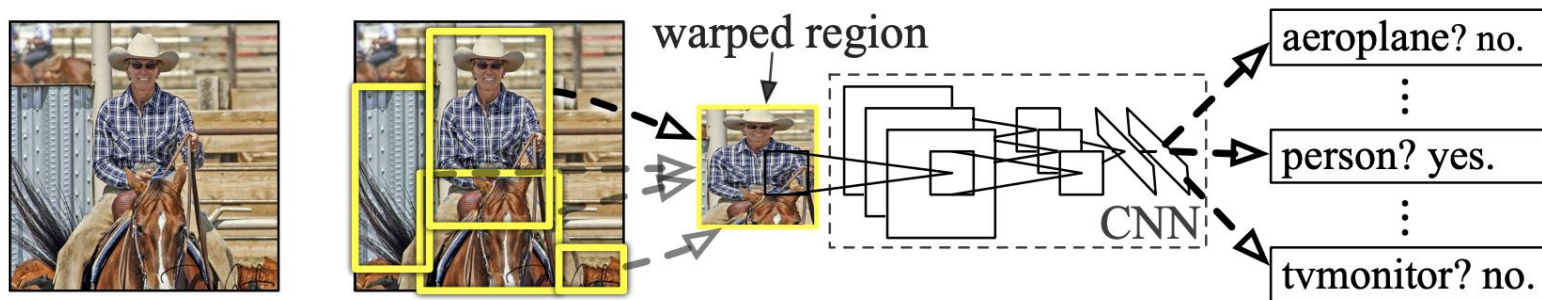
Faster-RCNN, YOLO

簡単なおさらい (1/2)

従来手法 (R-CNN)

- パート 1 : 物体候補領域の提案
- パート 2 : 提案された候補領域における物体のクラス分類

R-CNNのネットワーク



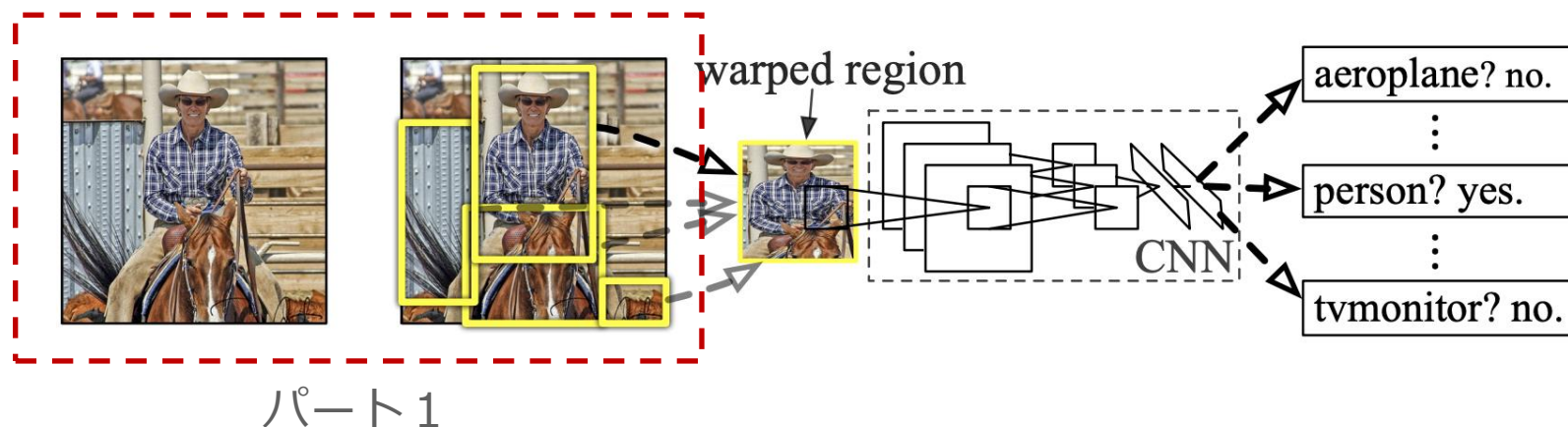
[1] Girshick, Ross B. et al. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." 2014 IEEE Conference on Computer Vision and Pattern Recognition (2014): 580-587.

簡単なおさらい (1/2)

従来手法 (R-CNN)

- パート 1 : 物体候補領域の提案
- パート 2 : 提案された候補領域における物体のクラス分類

R-CNNのネットワーク



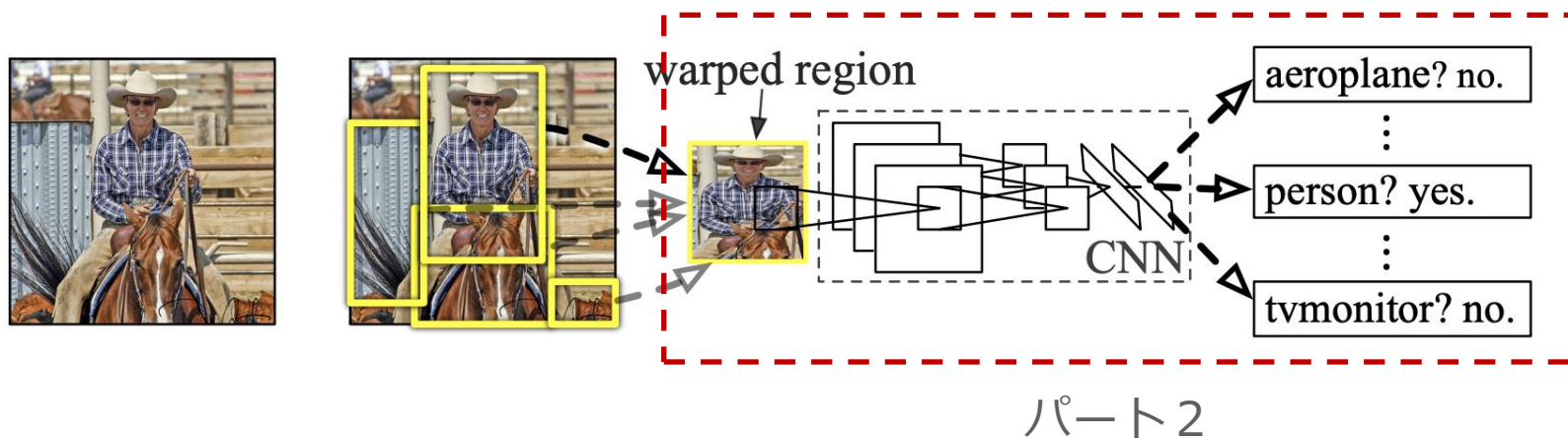
[1] Girshick, Ross B. et al. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." 2014 IEEE Conference on Computer Vision and Pattern Recognition (2014): 580-587.

簡単なおさらい (1/2)

従来手法 (R-CNN)

- パート1：物体候補領域の提案
- パート2：提案された候補領域における物体のクラス分類

R-CNNのネットワーク



[1] Girshick, Ross B. et al. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." 2014 IEEE Conference on Computer Vision and Pattern Recognition (2014): 580-587.

簡単なおさらい (2/2)

R-CNNの課題

- パート 1 の処理に Selective Searchを使用しているため、処理速度が遅い
- Fast-R-CNNではパート 2 の処理は改良されたものの、パート 1 にはSelective Searchを使用

→ 画像 1 枚の処理時間は、**パート 1 : 1 秒**, パート 2 : 0.22秒



----- 課題解決のために -----

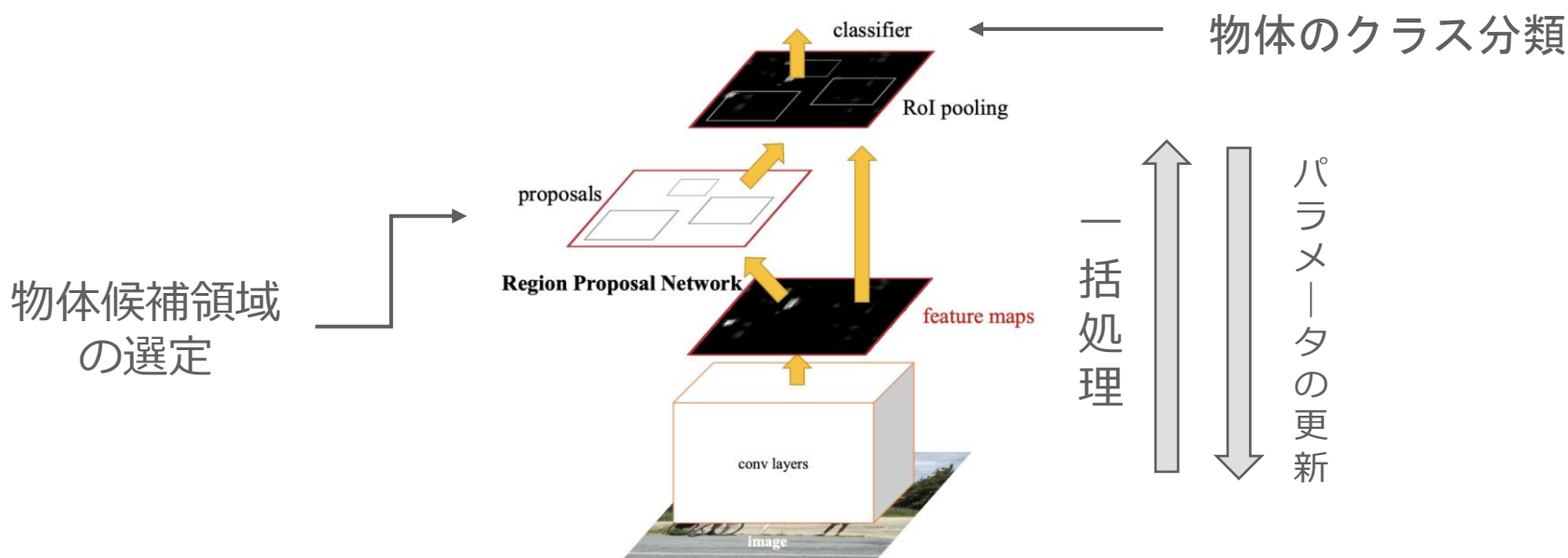
高速な処理を行う手法を提案する

→動画などのリアルタイム処理が可能になる

Faster-RCNN | アイデア

- 物体候補領域の提案の処理にCNNを使用する**RPN**
(Region Proposal Network)を提案
 - **End-to-Endな処理**が可能になった

Faster-RCNNのネットワーク

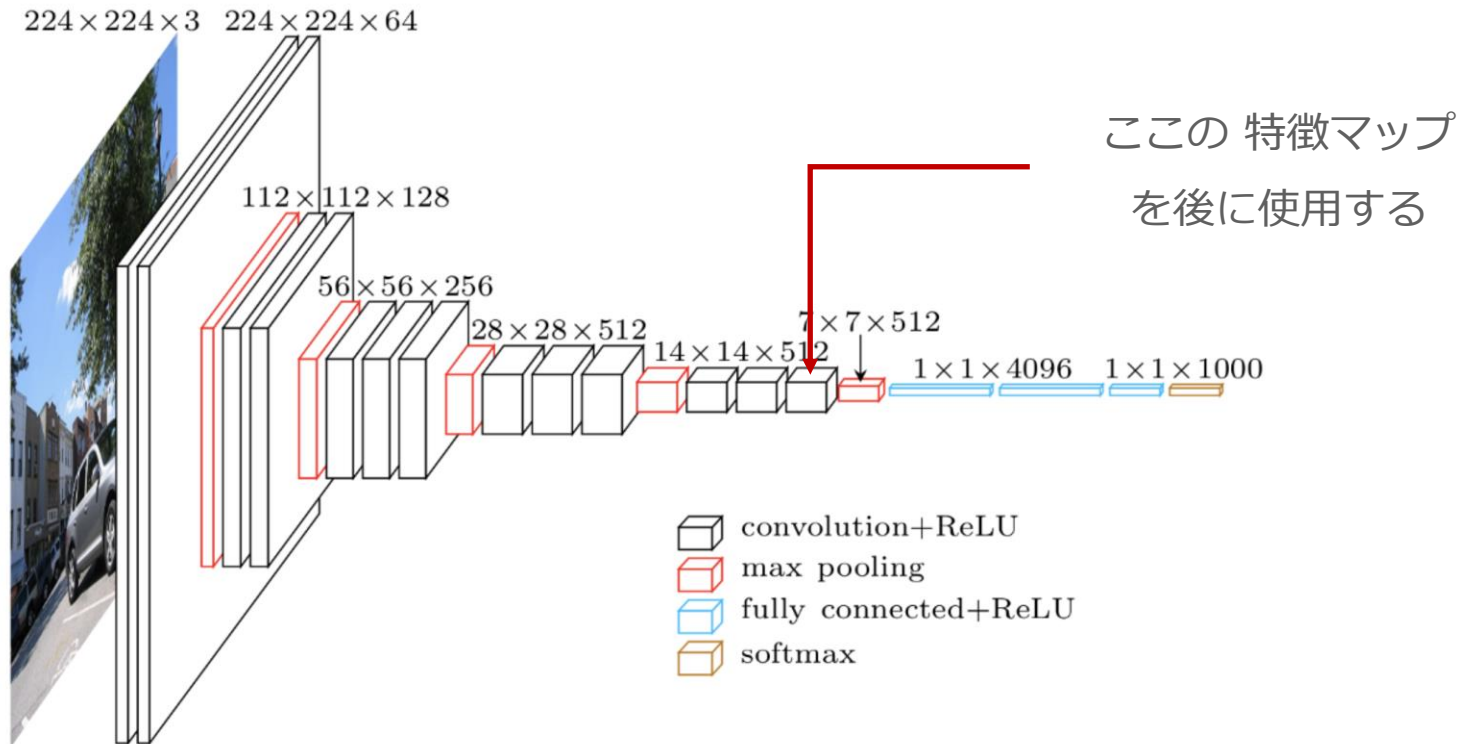


[2] Ren, Shaoqing et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (2015): 1137-1149.

Faster-RCNN | RPN (1/2)

■ 特徴マップ

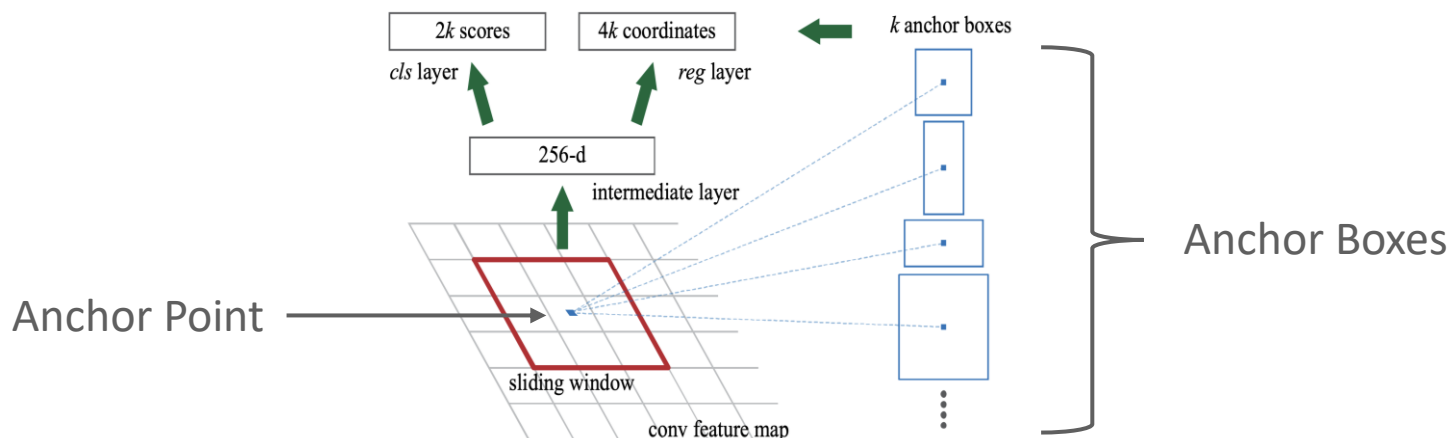
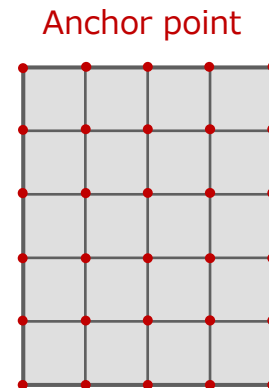
- 入力された画像は, **VGG16**により特徴マップに変換される



Faster-RCNN | RPN (2/2)

■ Anchors

- 特徴マップに Anchor Points を仮視し, PointごとにAnchor Boxesを作成する
- Anchor Points の個数 : $H \times W$
- 1つのAnchor PointあたりのAnchor Boxの個数 : S
- Anchor Boxes の個数 : $H \times W \times S$

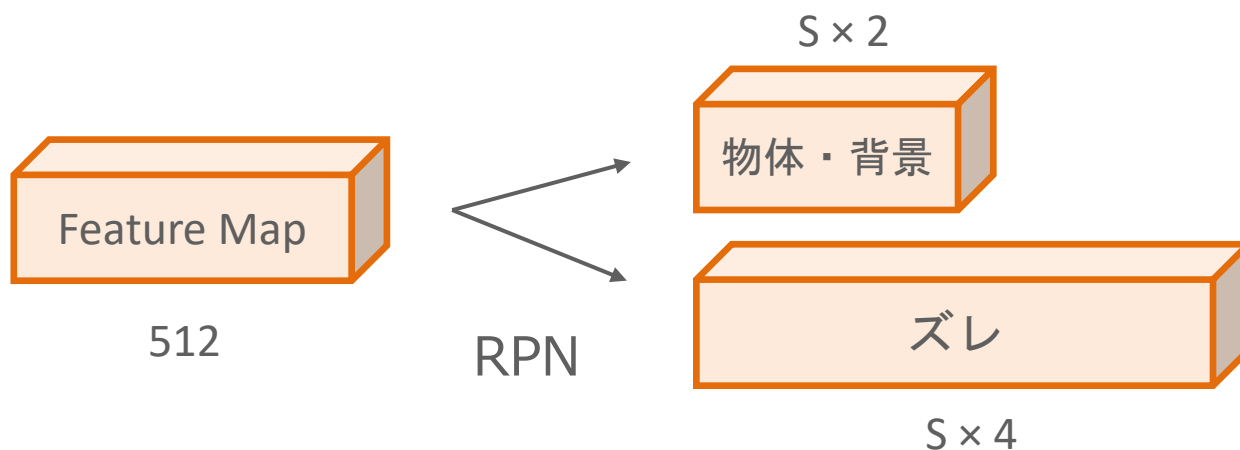


[2] Ren, Shaoqing et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39 (2015): 1137-1149.

Faster-RCNN | RPNの補足

■ RPNの出力

- 各 Anchor Boxes を Grand Truth の Boxesと比較し, 含まれているものが背景か物体か, どれくらいズレてるか出力
- 各Anchor Boxesにおいて背景か物体か : $H \times W \times S \times 2$
- 各Anchor Boxesにおいて正解Boxesとのズレ
(中心座標(x, y), 縦, 横) : $H \times W \times S \times 4$



Faster-RCNN | 実験

■ 従来手法と比較

- PASCAL VOC データで検証

	train	mAP	FPS
Fastest DPM [38]	2007	30.4	15
R-CNN Minus R [20]	2007	53.5	6
Fast R-CNN [14]	2007+2012	70.0	0.5
Faster R-CNN VGG-16[28]	2007+2012	73.2	7
Faster R-CNN ZF [28]	2007+2012	62.1	18

R-CNN, Fast R-CNN
より高速

R-CNNやFast R-CNNよりmAPスコアが高い

[2] Ren, Shaoqing et al. "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks."
IEEE Transactions on Pattern Analysis and Machine Intelligence 39 (2015): 1137-1149.

YOLO (V1) | アイデア

Faster R-CNN

物体候補領域の提案とクラス分類を
異なるネットワークで処理



YOLO (V1)^[3]

物体候補領域の提案とクラス分類を
一つのネットワークで処理

||

You Only Look Once

(一回だけ見れば良い)

[3] Redmon, Joseph et al. "You Only Look Once: Unified, Real-Time Object Detection." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 779-788.

YOLO (V1) | 利点と欠点

■ 利点

- 高速な処理
- 画像全体を一度に見るから、背景を物体と間違えることがない
- 汎化性が高い

■ 欠点

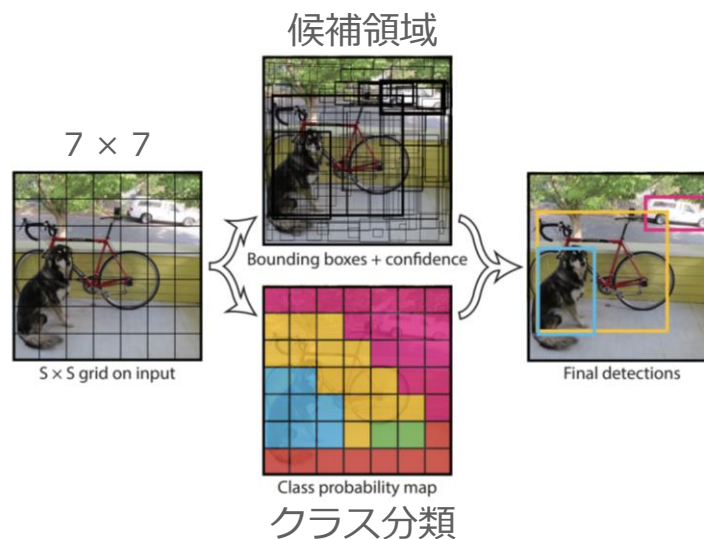
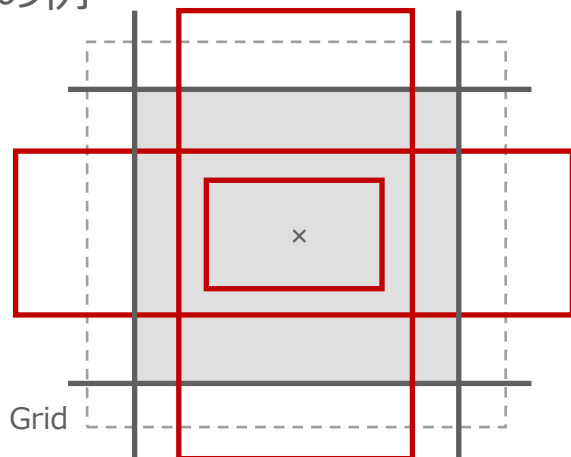
- 精度はFaster-RCNNに劣る

YOLO (V1) | 工夫

■ Grid cell

- 入力画像を $S \times S$ の Gridsに分割
- 候補領域の提案：各 Grid において，その Gridの真ん中を中心とするB個のBounding Boxを生成
- クラス分類：各 Grid ごとに，含む物体のクラスを分類

B = 3の例



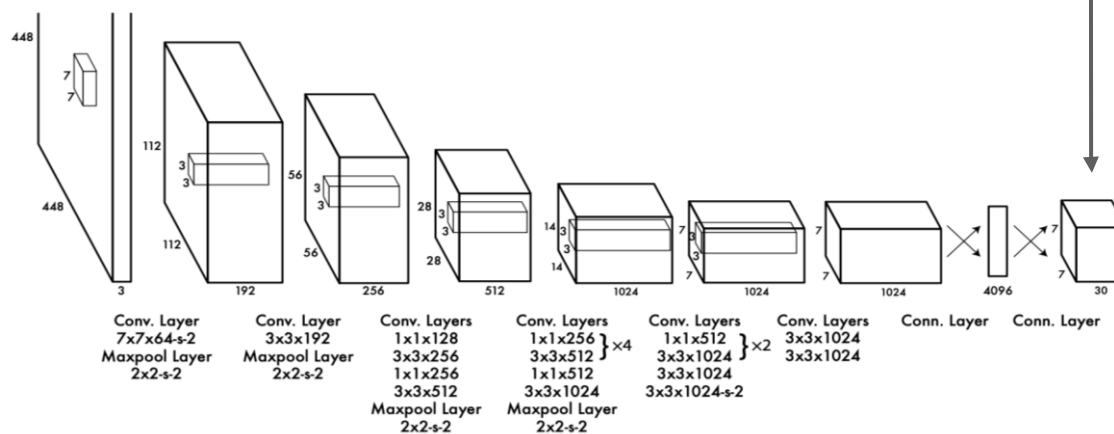
YOLO (V1) | ネットワーク

■ 何が出力されるか

- 各 Gridにおける各 バウンディングボックスの **中心, 高さ, 横** (x, y, w, h), **信頼度スコア**の5つと**各クラスに対応する** 特徴マップを同時に出力

→ $S = 7$, $B = 2$, クラス数 = 20の例

→ $7 \times 7 \times (2 \times 5 + 20) = 7 \times 7 \times 30$



YOLO (V1) | 実験

■ 従来手法と比較

Real-Time Detectors	Train	mAP	FPS
100Hz DPM [31]	2007	16.0	100
30Hz DPM [31]	2007	26.1	30
Fast YOLO	2007+2012	52.7	155
YOLO	2007+2012	63.4	45
Less Than Real-Time			
Fastest DPM [38]	2007	30.4	15
R-CNN Minus R [20]	2007	53.5	6
Fast R-CNN [14]	2007+2012	70.0	0.5
Faster R-CNN VGG-16[28]	2007+2012	73.2	7
Faster R-CNN ZF [28]	2007+2012	62.1	18
YOLO VGG-16	2007+2012	66.4	21

YOLOが最速

mAPスコアは
従来手法より高いわけではない

まとめ

従来手法のSelective Searchによる処理を改良し
処理を高速化したい



- Faster R-CNN
 - CNNベースのRPNを提案し, **End-To-Endなネットワーク**を可能にした
- YOLO
 - 候補領域の提案とクラス分類を**同時**に行うネットワークの提案