

進捗報告

1 今週行ったこと

- YOLO の検出仕様

バウンディングボックスの予測として、4つの値 x, y, w, h を予測する必要がある。ここで x, y, w, h はそれぞれ中心の x 座標, 中心の y 座標, バウンディングボックスの幅, バウンディングボックスの高さを表す。さらにロジスティック回帰でバウンディングボックスのスコアを予測する。このスコアが 0.5（しきい値）を超えた場合物体があると認識する。

- YOLO の特徴抽出ネットワーク

YOLOv3 は YOLOv2 からの変更点として, CNN のネットワークモデル (特徴抽出ネットワーク) が VGG ベースの Darknet19 から ResNet ベースの Darknet53 に変更された。53 層の畳み込み層をもっているためこう名付けられた。

表 1: Darknet53

	Type	Filters	Size	Output
	Convolutional	32	3×3	256×256
	Convolutional	64	$3 \times 3/2$	128×128
1 ×	Convolutional	32	1×1	
	Convolutional	64	3×3	
	Residual			128×128
	Convolutional	128	$3 \times 3/2$	64×64
2 ×	Convolutional	64	1×1	
	Convolutional	128	3×3	
	Residual			64×64
	Convolutional	256	$3 \times 3/2$	32×32
8 ×	Convolutional	128	1×1	
	Convolutional	256	3×3	
	Residual			32×32
	Convolutional	512	$3 \times 3/2$	16×16
8 ×	Convolutional	256	1×1	
	Convolutional	512	3×3	
	Residual			16×16
	Convolutional	1024	$3 \times 3/2$	8×8
4 ×	Convolutional	512	1×1	
	Convolutional	1024	3×3	
	Residual			8×8
	Avgpool		Global	
	Connected		1000	
	Softmax			

2 次回行うこと

- 一般物体認識を理解して実装．（出力がラベルから位置予測（矩形）にする方法がわかりません．）