

YOLO - Unified, Real-Time Object Detection

Stanisław Wilczyński

Uniwersytet Wrocławski

8 maja 2018

Po co nam wykrywanie obiektów w czasie rzeczywistym?

- Autonomiczne samochody

Na zachętę

Filmik promujący YOLO

Wykrywanie obiektów - techniki

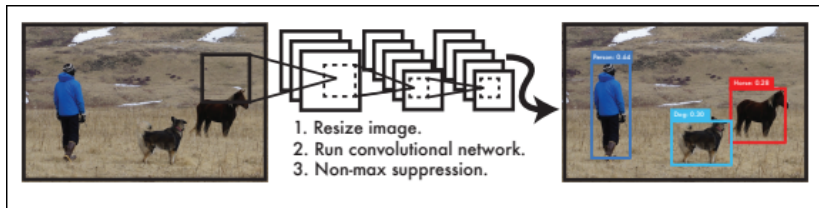
- Aplikowanie sieci klasyfikującej w różnych rejonach obrazka
- Okno przesuwne
- Bounding boxes jak np. w R-CNN

- Aplikowanie sieci klasyfikującej w różnych rejonach obrazka
- Okno przesuwne
- Bounding boxes jak np. w R-CNN

Problemy

Zbyt skomplikowane \Rightarrow zbyt wolne, żeby działać w czasie rzeczywistym

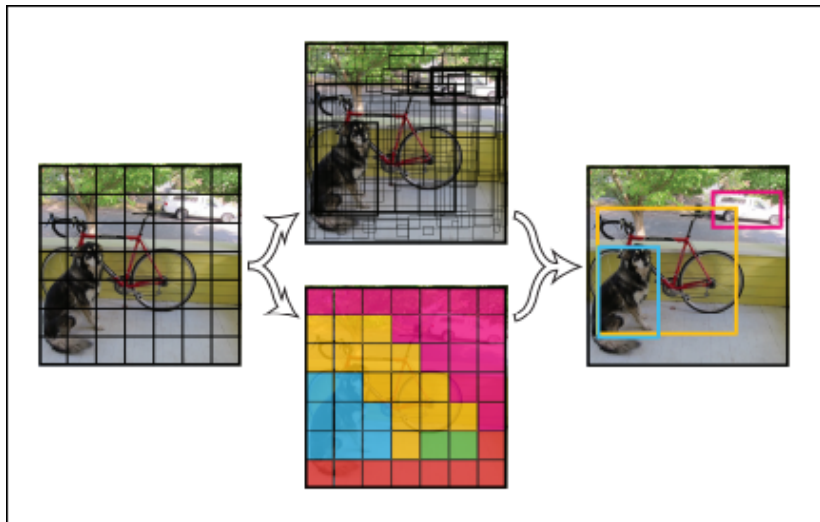
- Bounding boxes i klasyfikacja naraz - jedno przejście przez sieć
- Globalna analiza obrazka



Rysunek: Schemat działania YOLO

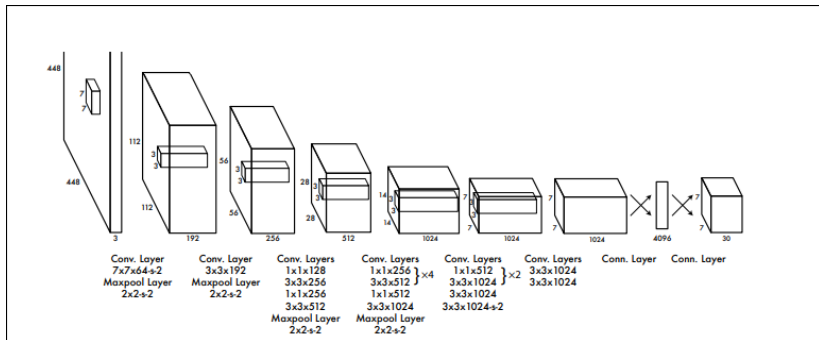
Wstęp w rozdziale 2

Bounding boxes



Rysunek: Bounding boxes

Architektura sieci



Rysunek: Architektura sieci

$$\begin{aligned} & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \\ & + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} \left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \\ & + \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{obj}} (C_i - \hat{C}_i)^2 \\ & + \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B \mathbb{1}_{ij}^{\text{noobj}} (C_i - \hat{C}_i)^2 \\ & + \sum_{i=0}^{S^2} \mathbb{1}_i^{\text{obj}} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2 \end{aligned}$$

Rysunek: Kara

Trening (2.2, 2.3)

Ograniczenia (2.4)

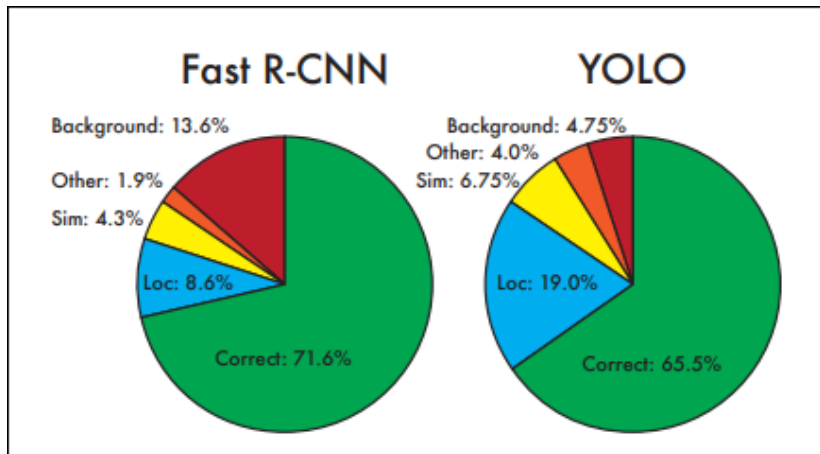
- mAP - mean average precision?
- Wykresy

Porównanie z innymi metodami

Real-Time Detectors	Train	mAP	FPS
100Hz DPM [30]	2007	16.0	100
30Hz DPM [30]	2007	26.1	30
Fast YOLO	2007+2012	52.7	155
YOLO	2007+2012	63.4	45
Less Than Real-Time			
Fastest DPM [37]	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[27]	2007+2012	73.2	7
Faster R-CNN ZF [27]	2007+2012	62.1	18

Rysunek: YOLO vs inne metody

YOLO vs R-CNN



Rysunek: YOLO vs R-CNN

YOLO - generalizacja



Rysunek: YOLO na obrazkach z bardziej różnorodnych zbiorów

YOLO9000 oraz YOLOv3:

- Jak starczy czasu



Joseph Redmon, Santosh Kumar Divvala, Ross B. Girshick, and Ali Farhadi.

You only look once: Unified, real-time object detection.

CoRR, abs/1506.02640, 2015.



Joseph Redmon and Ali Farhadi.

YOLO9000: better, faster, stronger.

CoRR, abs/1612.08242, 2016.



Joseph Redmon and Ali Farhadi.

Yolov3: An incremental improvement.

CoRR, abs/1804.02767, 2018.