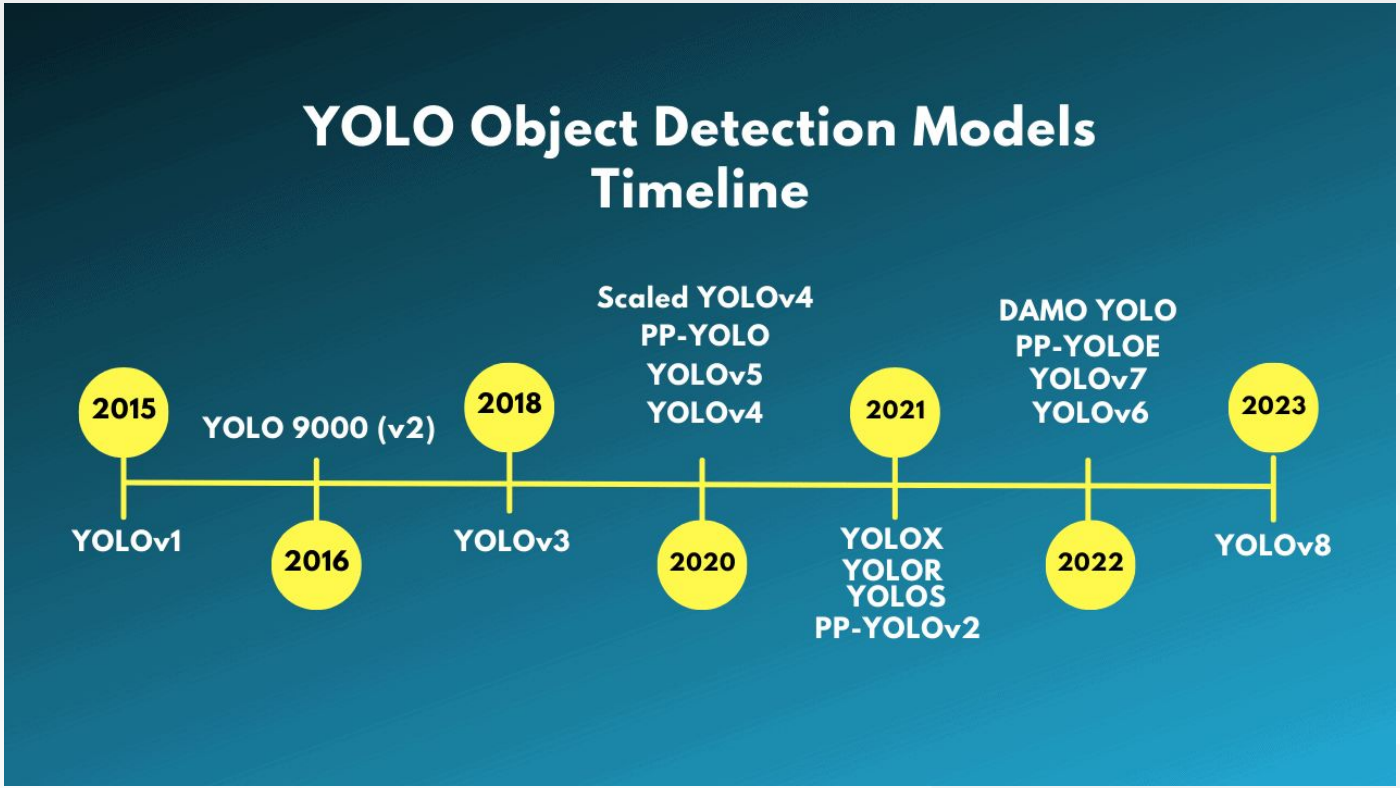
YOLO

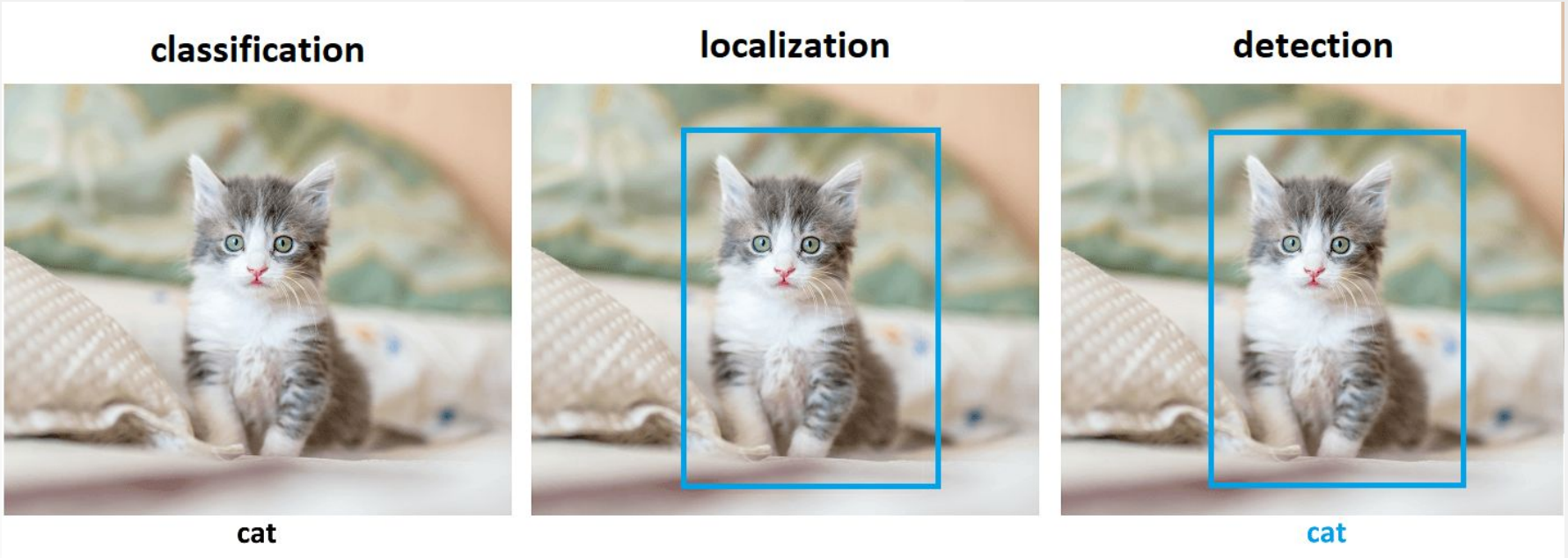
11.06.2024 Computer Vision Seminar 23/24

https://medium.datadriveninvestor.com/yolov8-the-evolution-of-real-time-object-detection-7c158948e0de



https://www.baeldung.com/cs/yolo-algorithm

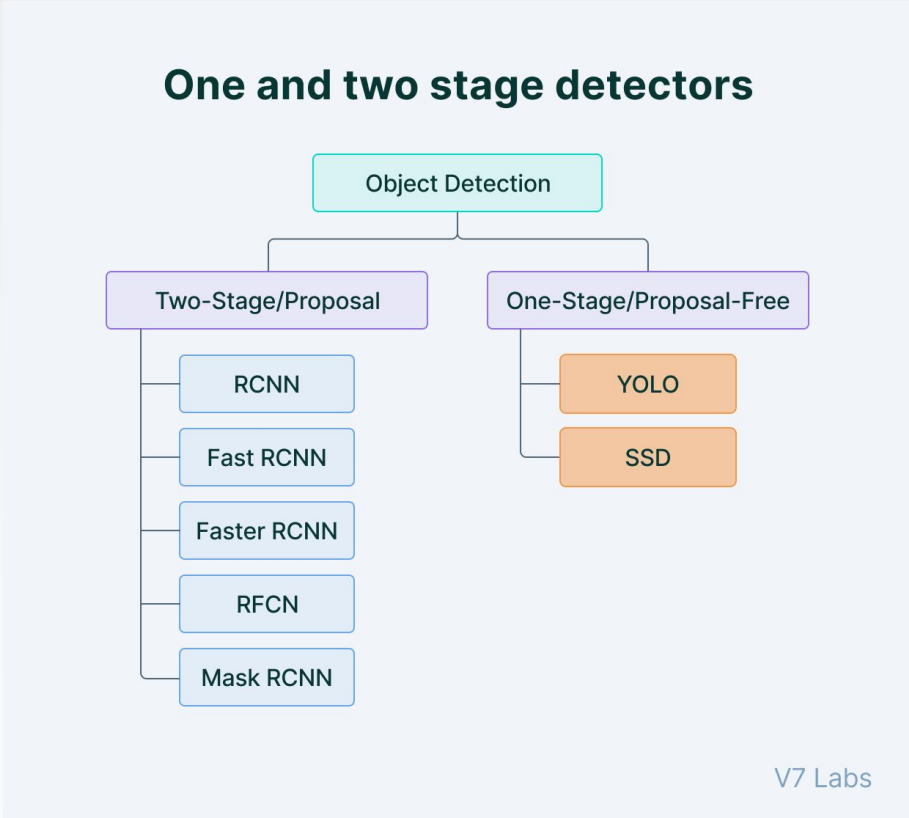
Detekcja obiektów



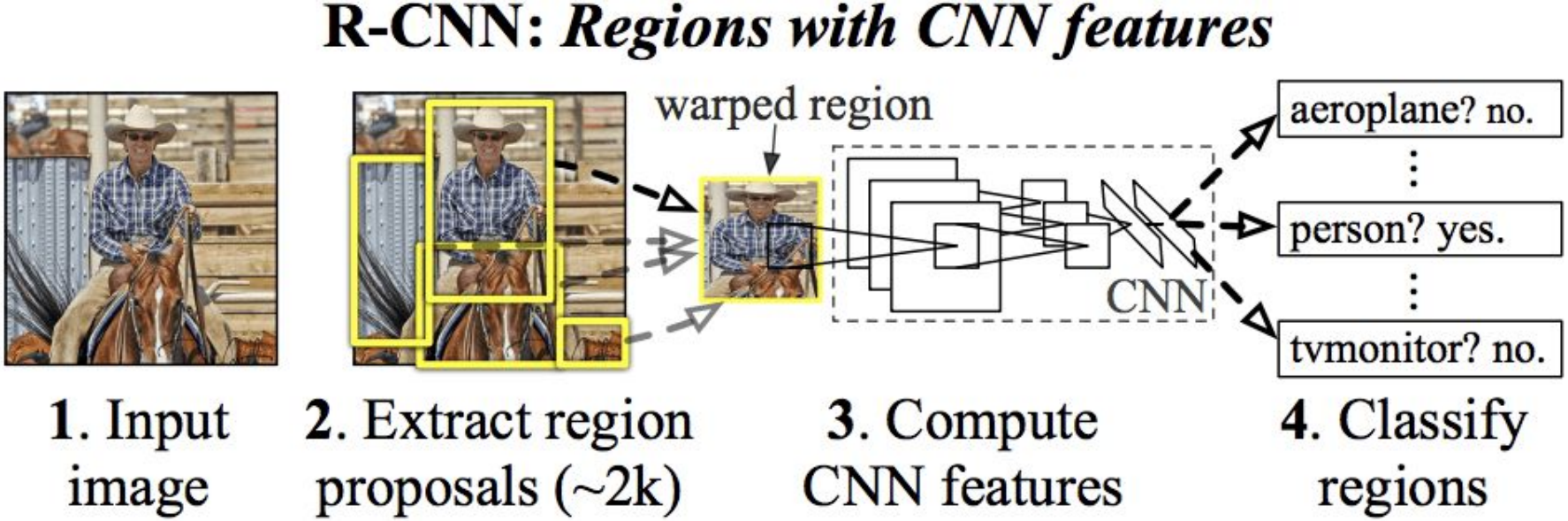
https://www.v7labs.com/blog/object-detection-guide

Podział:

- Jednoetapowe
- Dwuetapowe



https://arxiv.org/abs/1311.2524



https://arxiv.org/pdf/1506.02640v5

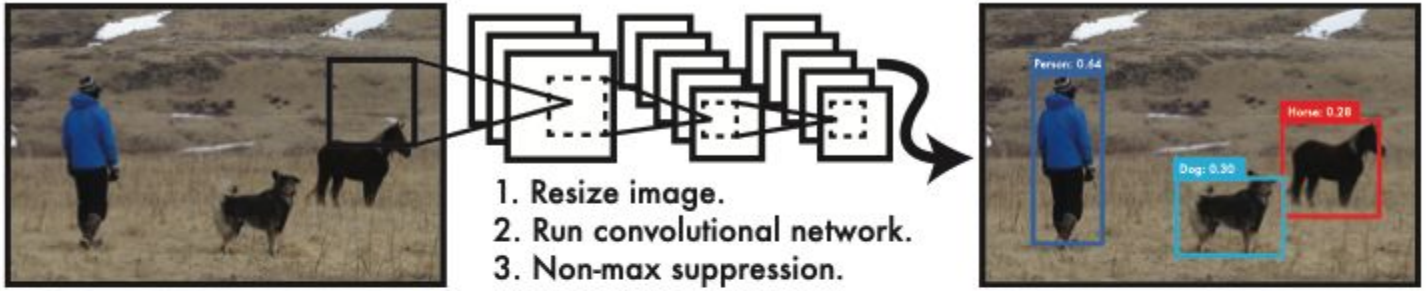
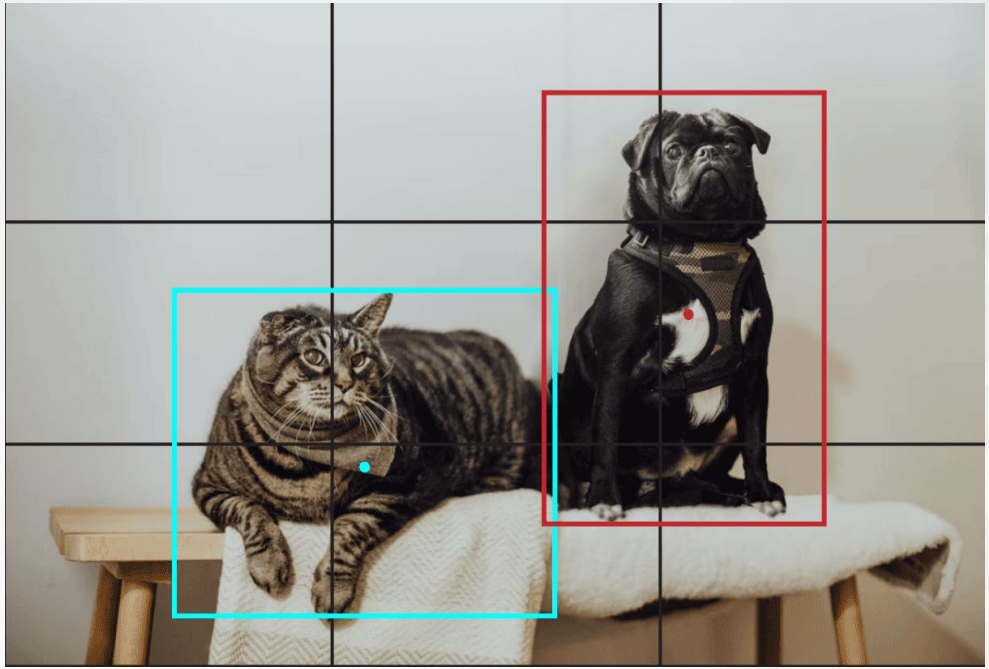


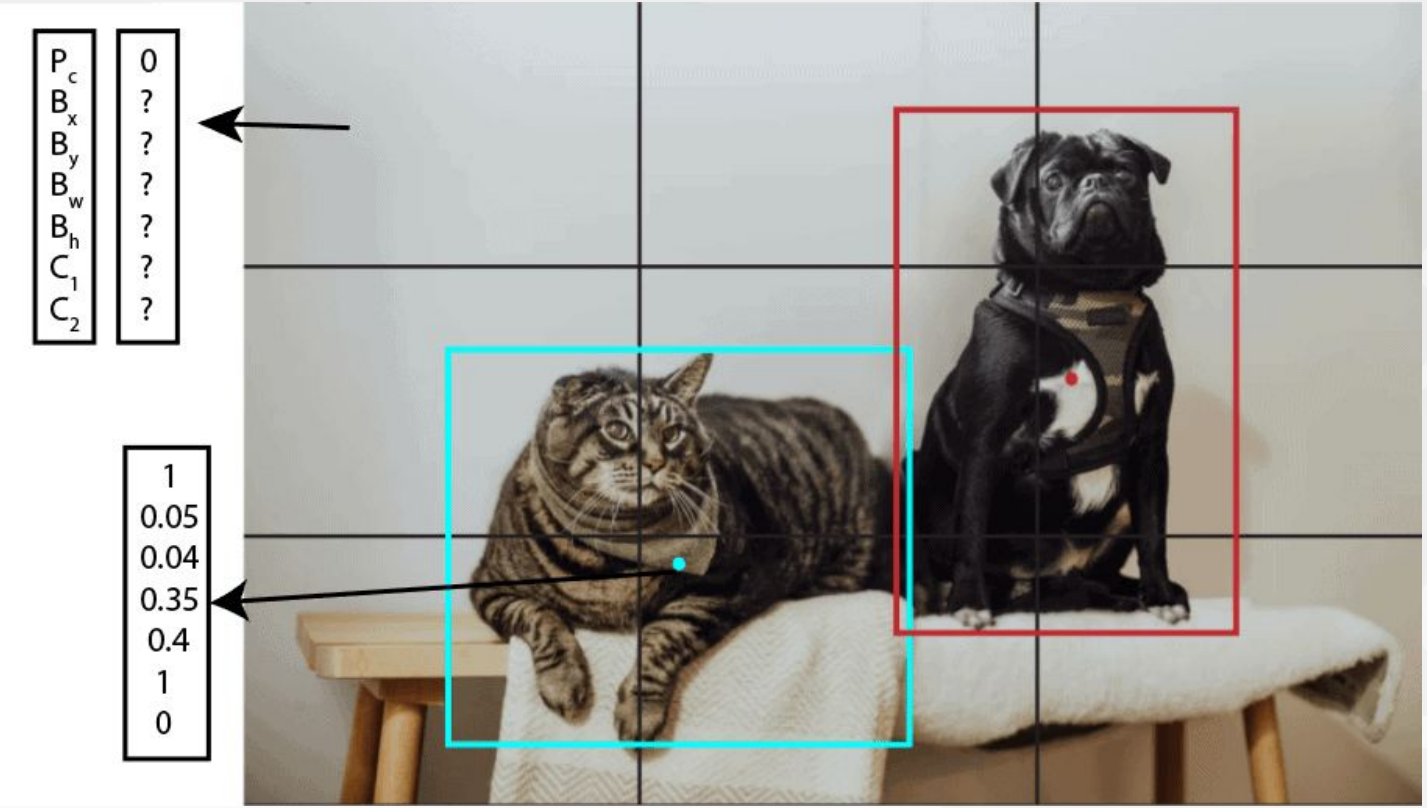
Figure 1: The YOLO Detection System. Processing images with YOLO is simple and straightforward. Our system (1) resizes the input image to 448 × 448, (2) runs a single convolutional network on the image, and (3) thresholds the resulting detections by the model's confidence.

https://www.baeldung.com/cs/yolo-algorithm

Jak działa YOLO?

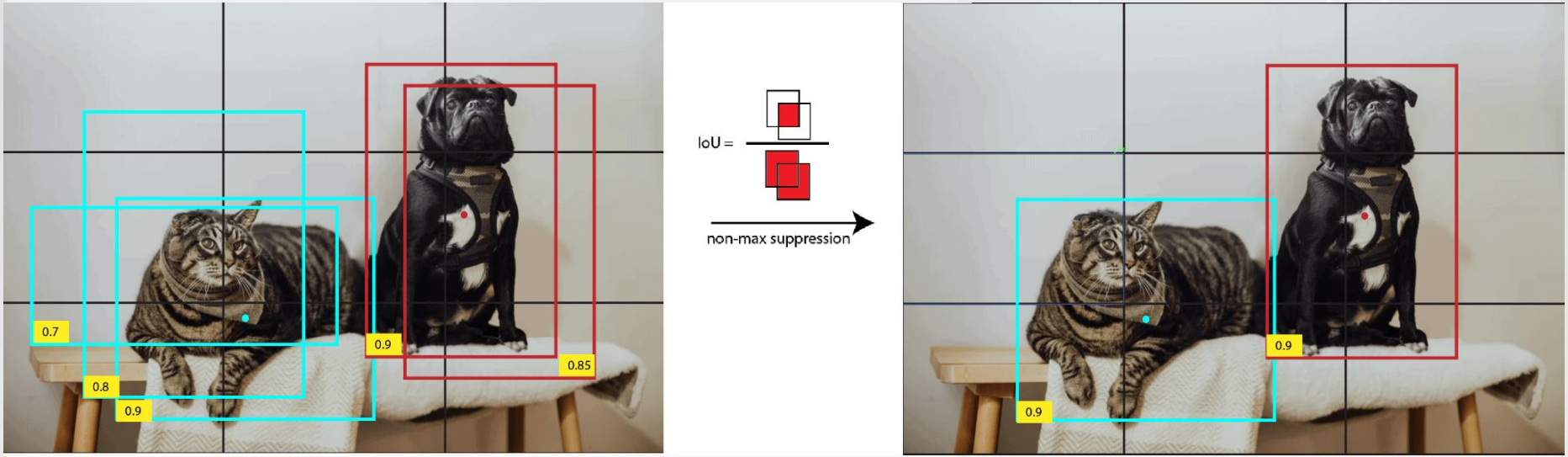


https://www.baeldung.com/cs/yolo-algorithm



https://www.baeldung.com/cs/yolo-algorithm

$$IoU = \frac{\text{area of the intersection between } B_1 \text{ and } B_2}{\text{area of the union between } B_1 \text{ and } B_2},$$



https://arxiv.org/pdf/1506.02640v5

Architektura YOLOv1

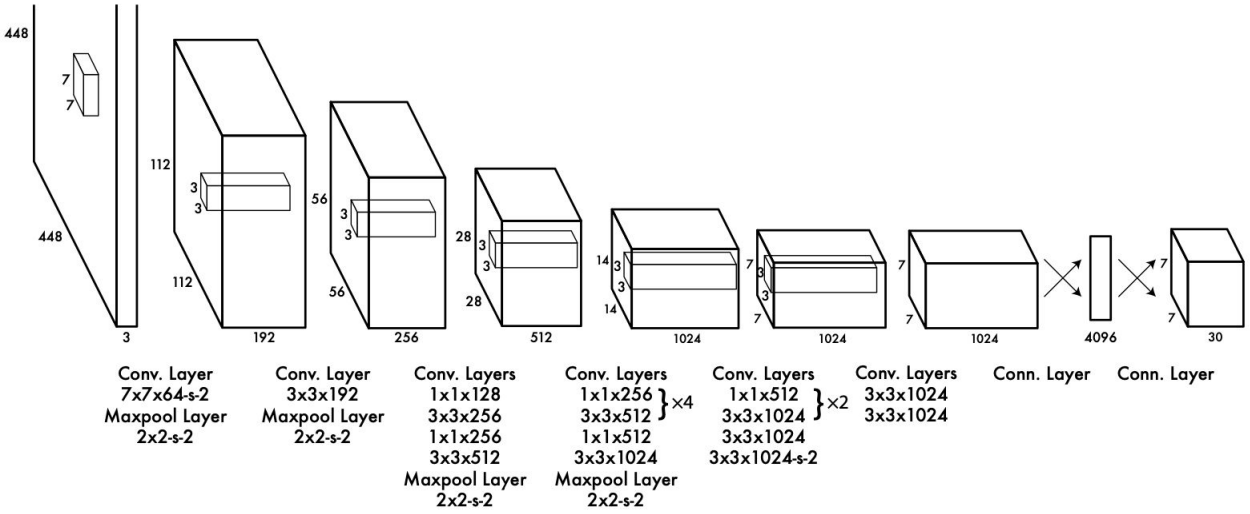
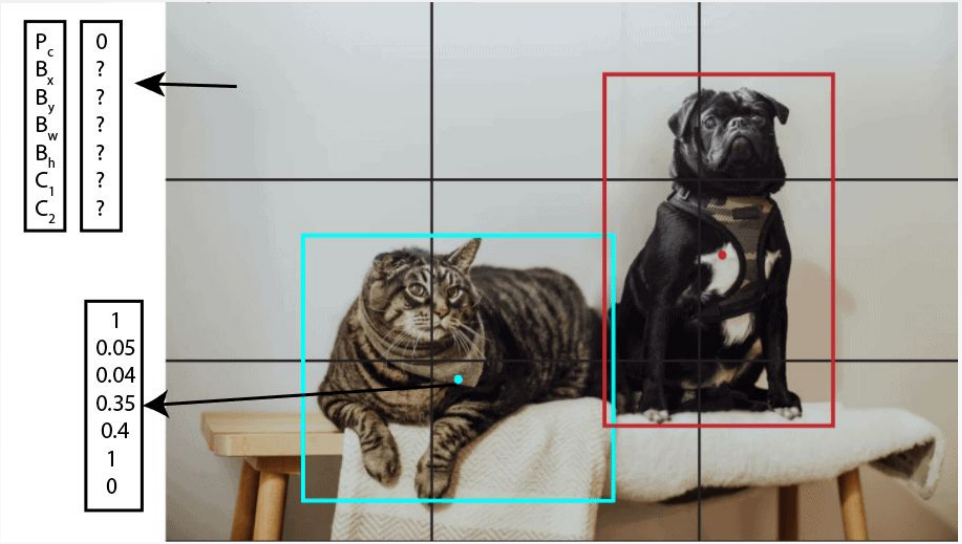


Figure 3: The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1×1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224×224 input image) and then double the resolution for detection.

https://arxiv.org/pdf/1506.02640v5

Funkcja straty



$$\lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right]$$

$$+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left[\left(\sqrt{w_i} - \sqrt{\hat{w}_i} \right)^2 + \left(\sqrt{h_i} - \sqrt{\hat{h}_i} \right)^2 \right]$$

$$+ \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{obj}} \left(C_i - \hat{C}_i \right)^2$$

$$+ \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \mathbb{1}_{ij}^{\text{noobj}} \left(C_i - \hat{C}_i \right)^2$$

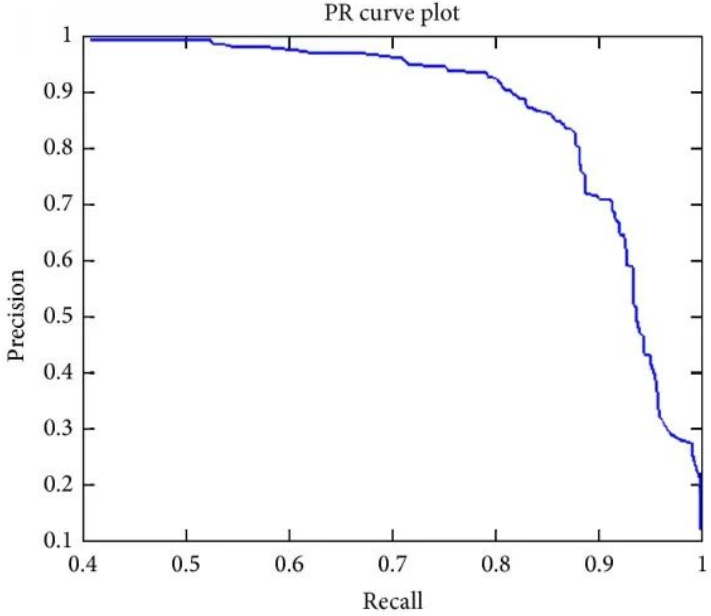
$$+ \sum_{i=0}^{S^2} \mathbb{1}_{i}^{\text{obj}} \sum_{c \in \text{classes}} \left(p_i(c) - \hat{p}_i(c) \right)^2$$

https://towardsdatascience.com/on-object-detection-metrics-with-worked-example-216f173ed31e

Metryki

- Średnia precyzja (AP)
- mAP – mean Average Precision

		Actual	
		Positive	Negative
Predicted	Positive	True Positive	False Positive
	Negative	False Negative	True Negative



https://arxiv.org/pdf/2405.14458

YOLOv10

https://arxiv.org/pdf/2405.14458

Dual label assignment

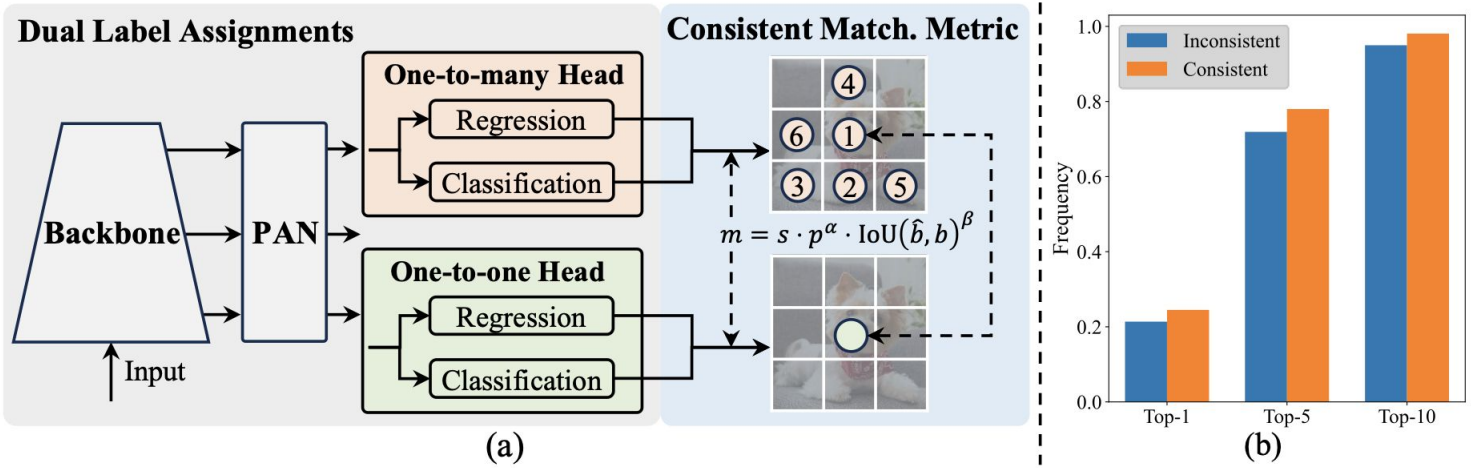
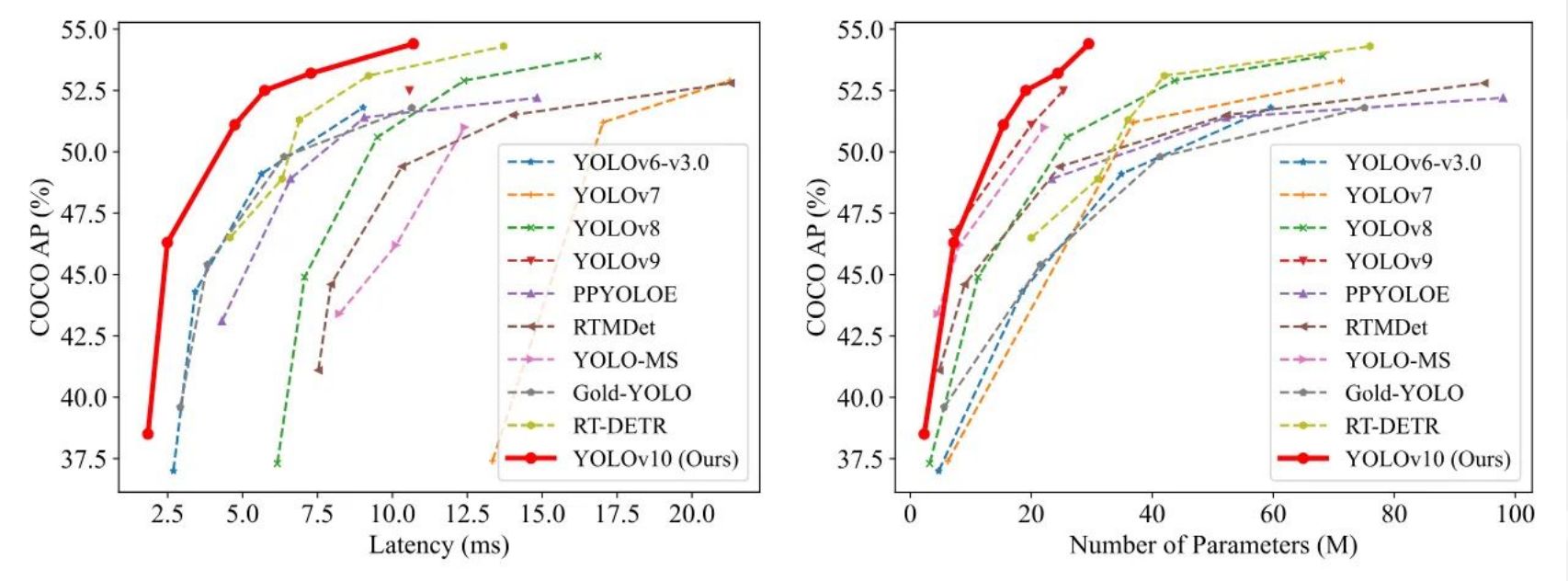


Figure 2: (a) Consistent dual assignments for NMS-free training. (b) Frequency of one-to-one assignments in Top-1/5/10 of one-to-many results for YOLOv8-S which employs α_{o2m}=0.5 and β_{o2m}=6 by default [20]. For consistency, α_{o2o}=0.5; β_{o2o}=6. For inconsistency, α_{o2o}=0.5; β_{o2o} =2.

https://arxiv.org/pdf/2405.14458

YOLOv10



Materiały

- YOLO Explained. What is YOLO? | by Ani Aggarwal | Analytics Vidhya | Medium
- https://colab.research.google.com/github/roboflow-ai/notebooks/blob/main/notebooks/train-yolov10-object-detection-on-custom-dataset.ipynb
- YOLOv10: Instant Real-Time Object Detection with YOLO (You Only Look Once) | by AI TutorMaster | May, 2024 | Level Up Coding