Aspect-Ratio Sensitive Network (ARS Net)

Jianxiong Cai

Anqi Pang

Lei Jin

Ruijian Li

Peijia Xu

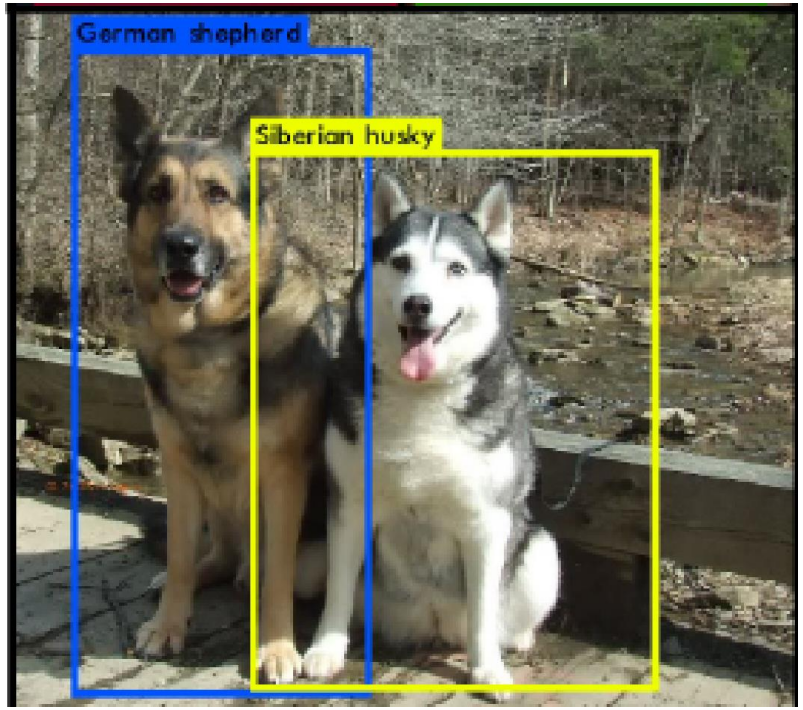
Contents

Contents

Object Detection

- Object Detection?
  - Identify bounding boxes
  - Classify the objects

- Two-stage Network (Faster RCNN)
  - Feature Extractor
  - RPN
  - Classifier + Regressor



An object detection result from YOLOv2

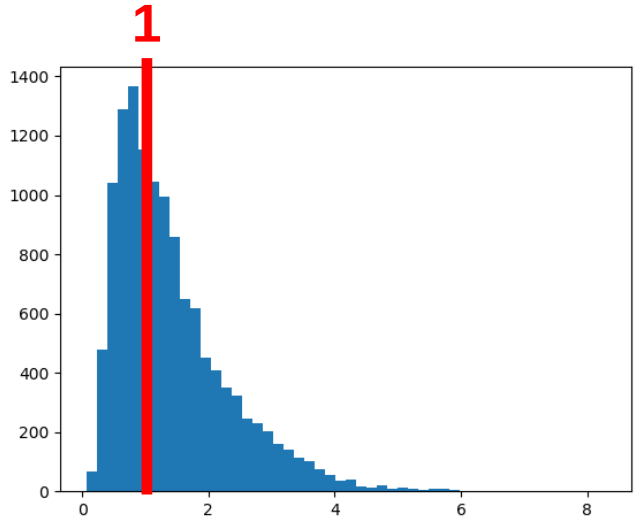
Redmon, Joseph, and Ali Farhadi. "YOLO9000: Better, Faster, Stronger." *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*. IEEE, 2017.

Motivation

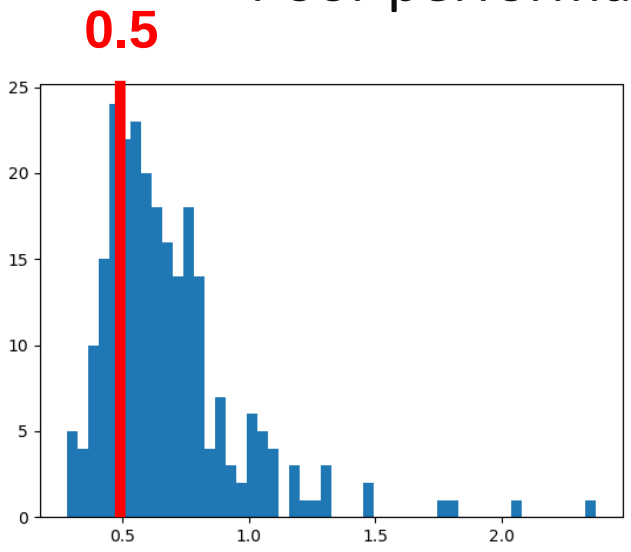
- Object tends to have certain shape (Aspect Ratio)

Motivation 1: Aspect Ratio

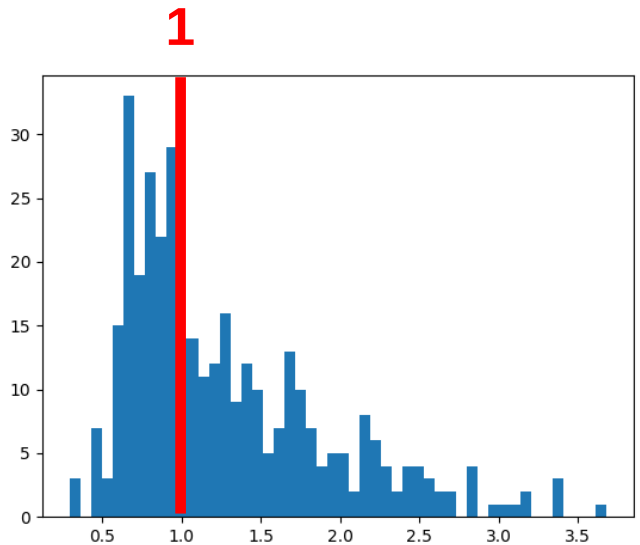
- VOC 2007 training set
  - Aspect ratio differ
  - Median ≈ 1
  - Poor performance



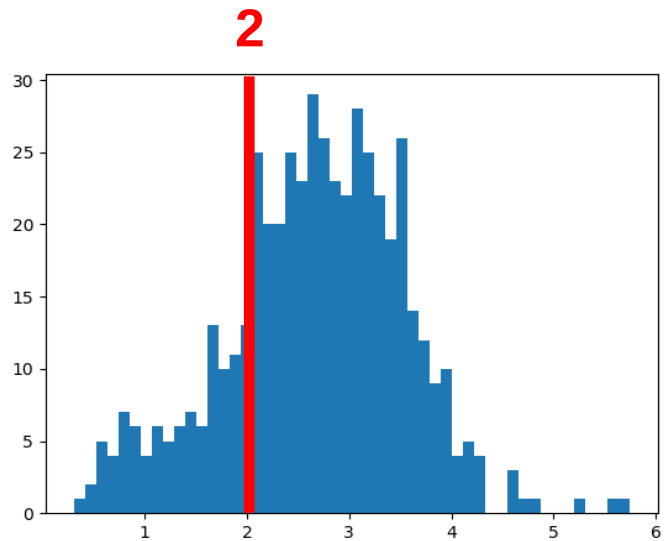
Overall distribution



car



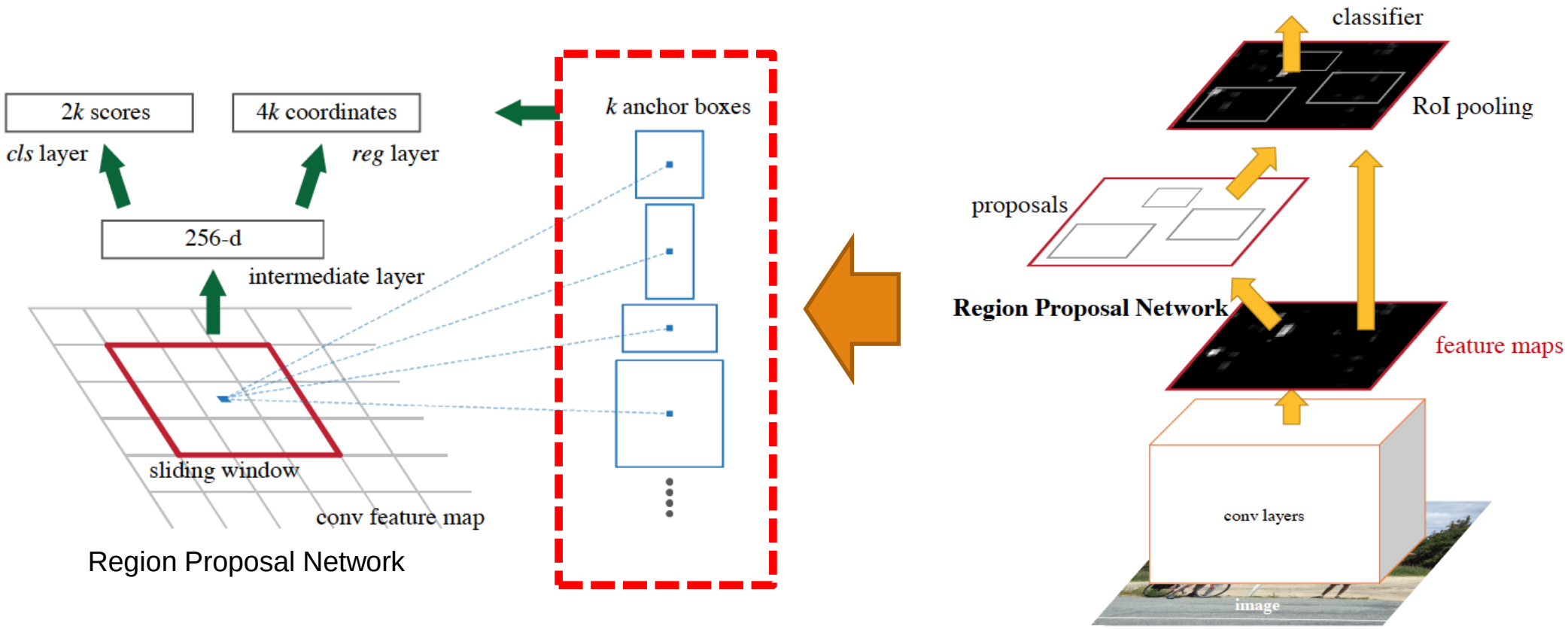
horse



bottle

M. Everingham, L. Van Gool, C. Williams, J. Winn, A. Zisserman, *The PASCAL Visual Object Classes Challenge 2007 (VOC2007)*.

Motivation 2: RPN



Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. 2015.

Goal

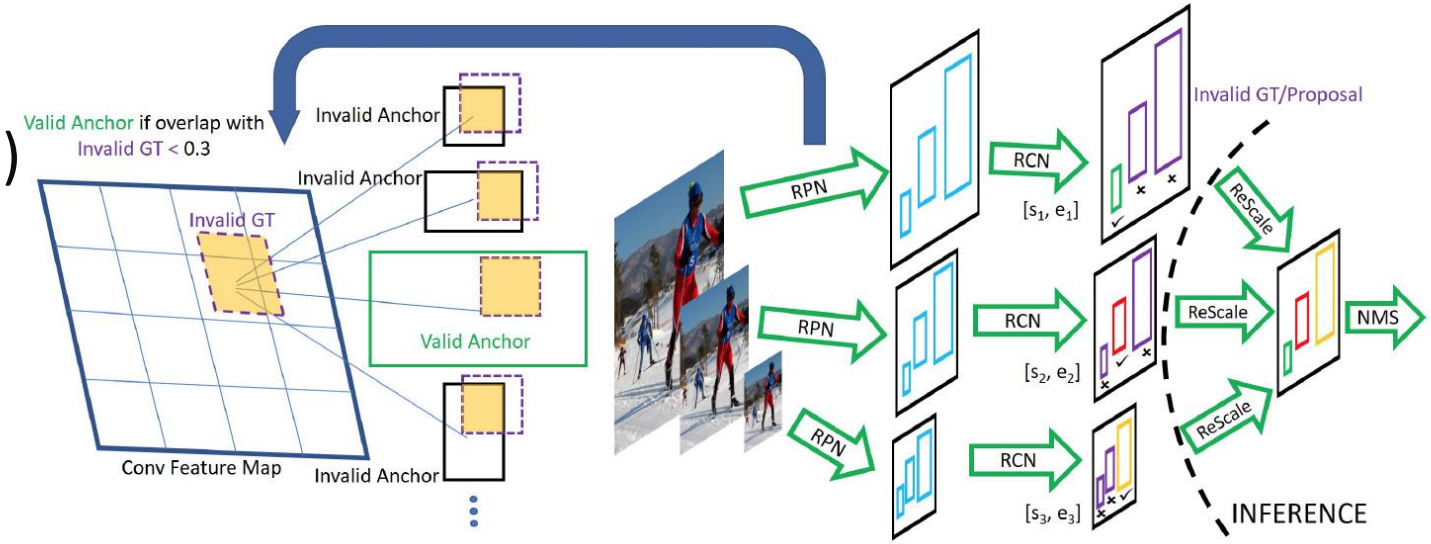
- RPN
  - Better ROI Proposal
  - Fewer ROI Proposal (less false negative)

- Aspect Ratio

Related work

- SNIP
  - Feature pyramid
  - Multi RPN

  (based on multi-scale)



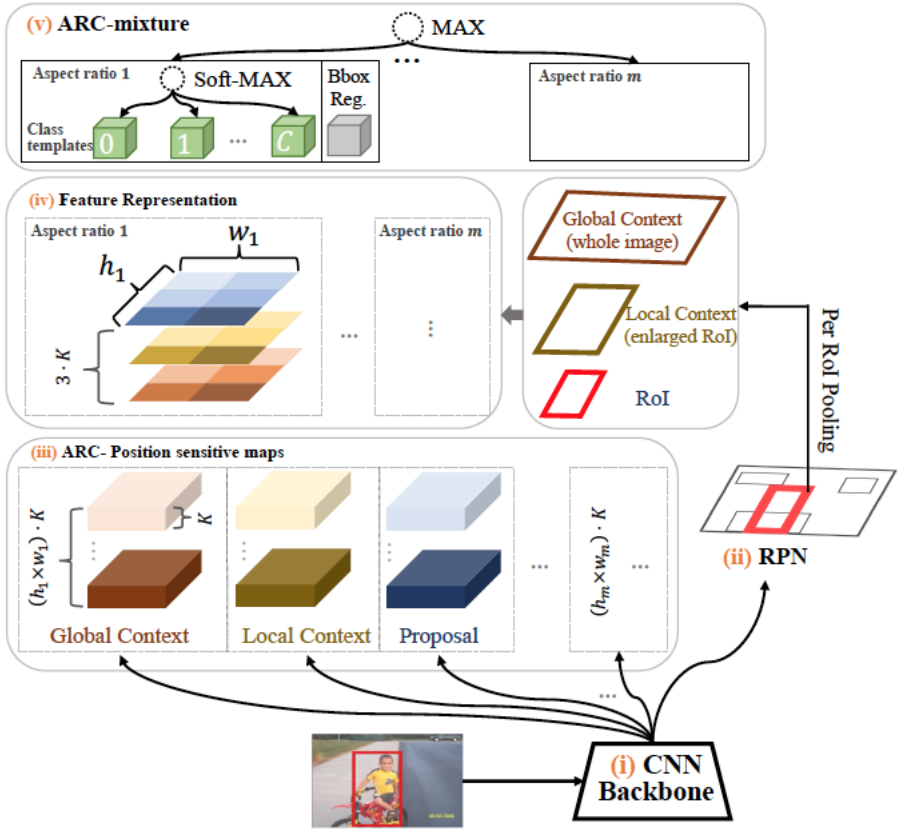
Singh, Bharat, and Larry S. Davis. "An Analysis of Scale Invariance in Object Detection–SNIP.", 2017.

Related work

- ARC-R-CNN

(Aspect Ratio and Context Aware)

  - Filter RPN output
  - Multiple Aspect Ratio

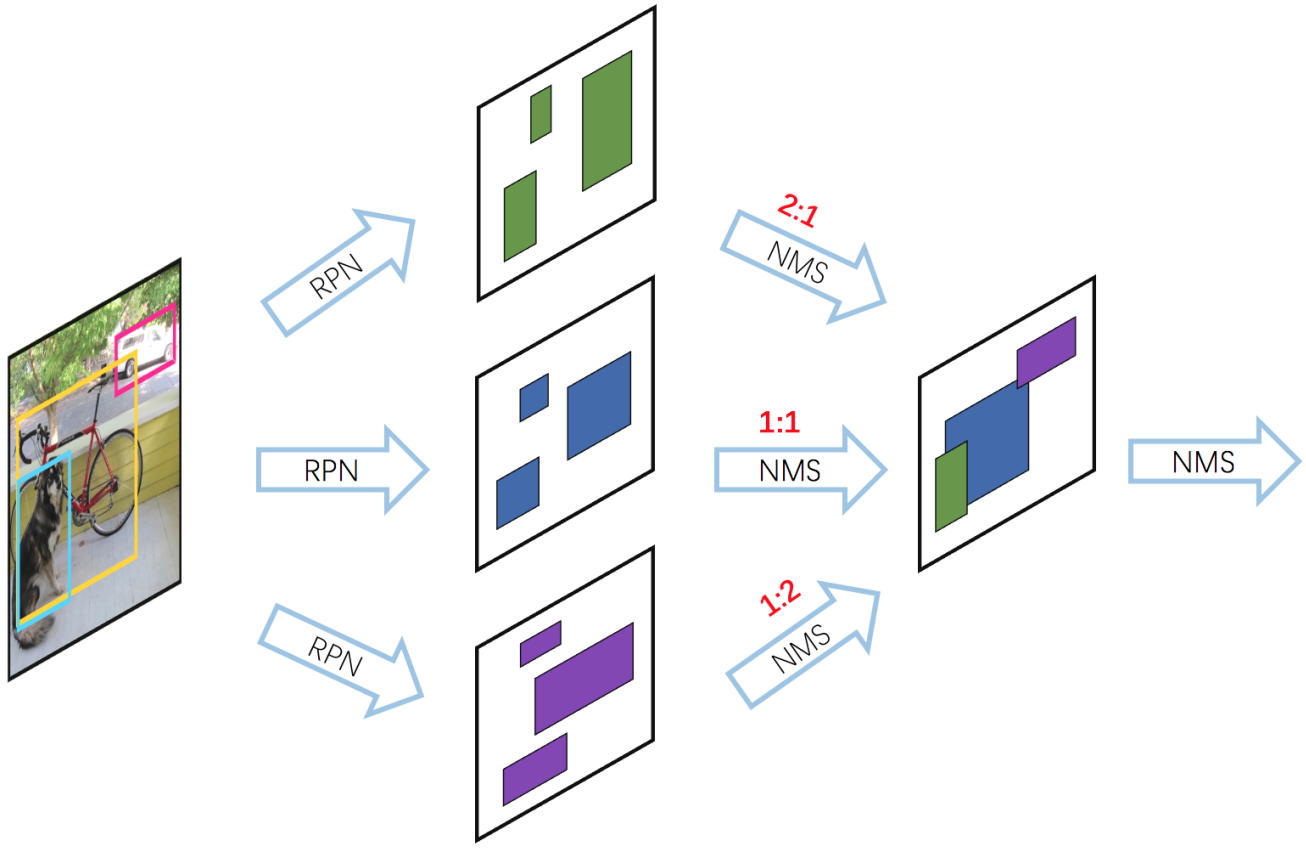


Li, Bo, et al. "Object Detection via Aspect Ratio and Context Aware Region-based Convolutional Networks.", 2016.

Contents

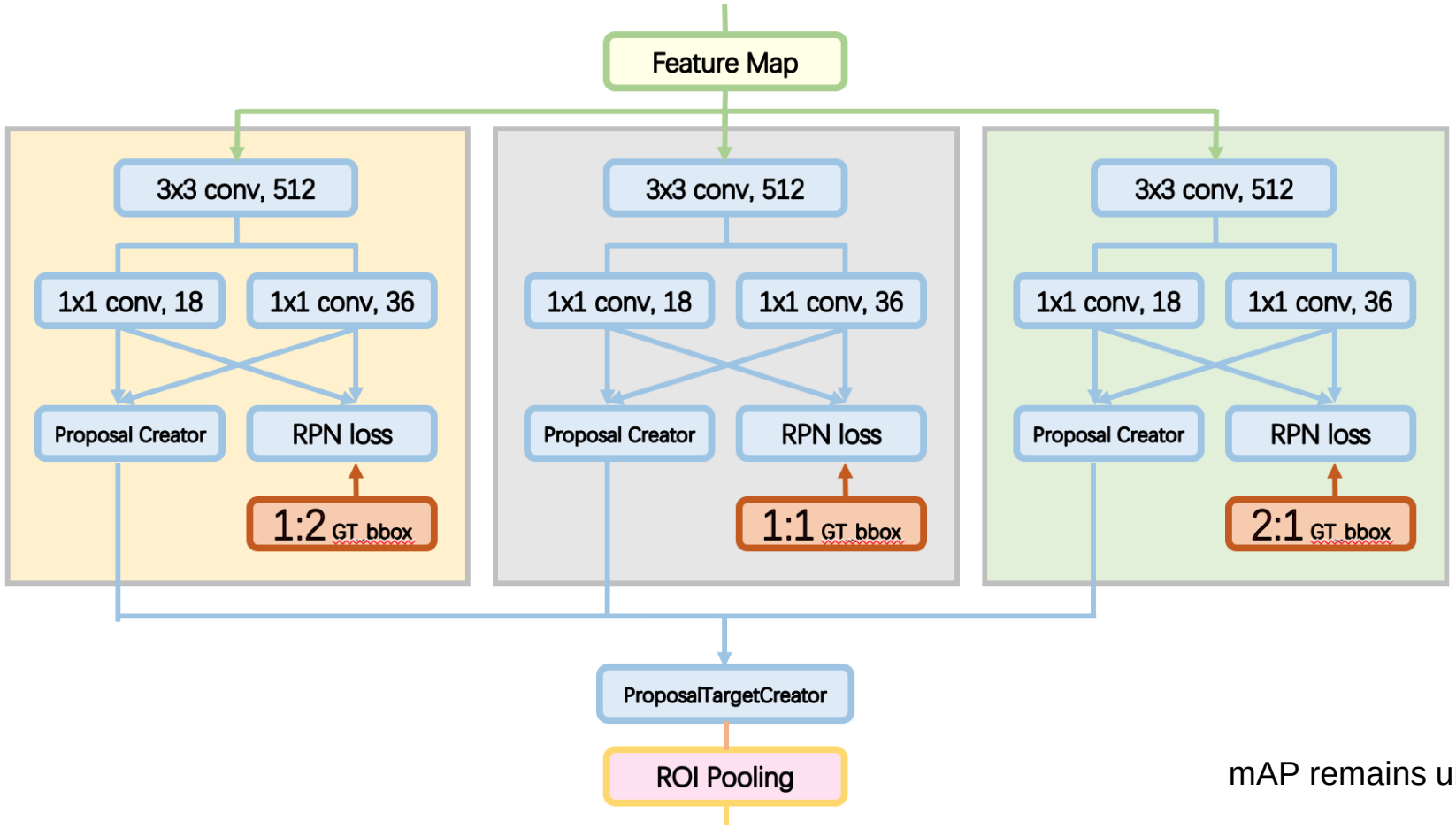
Our Approach

- Multiple RPNs
  - Aspect ratio differ
  - 3 anchor box each



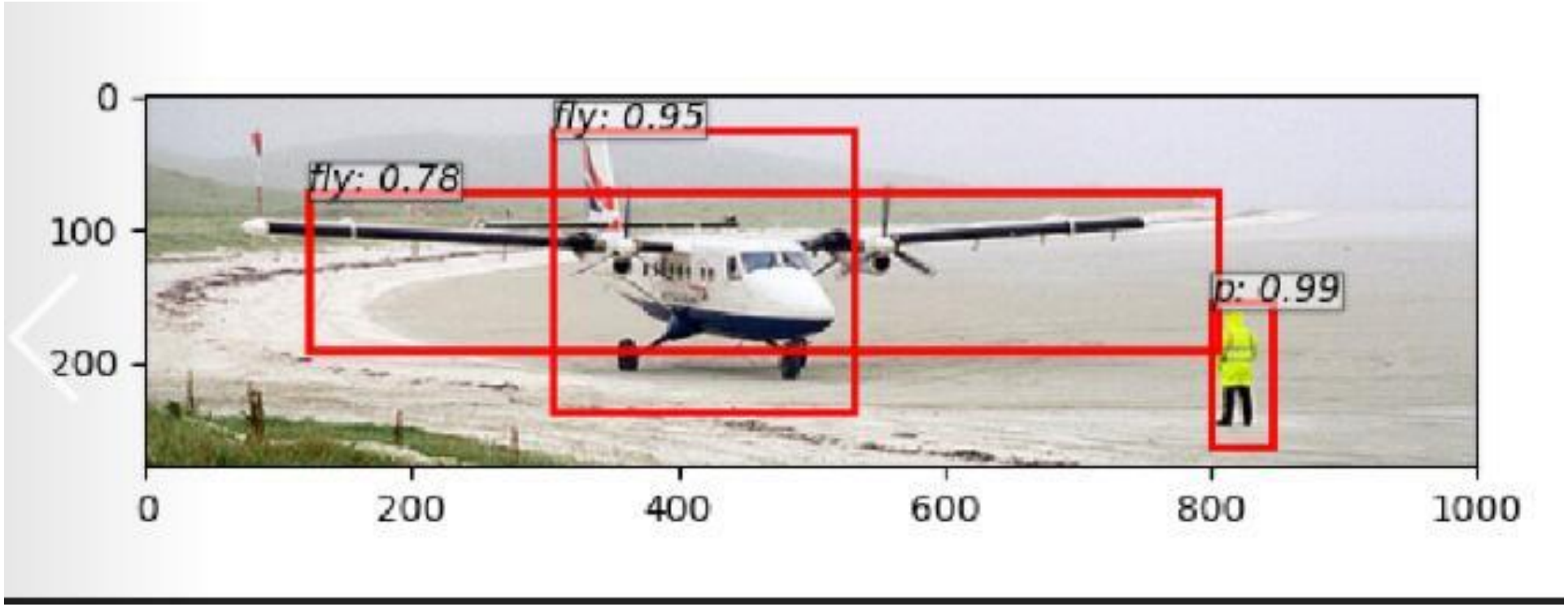
Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

Multiple RPNs

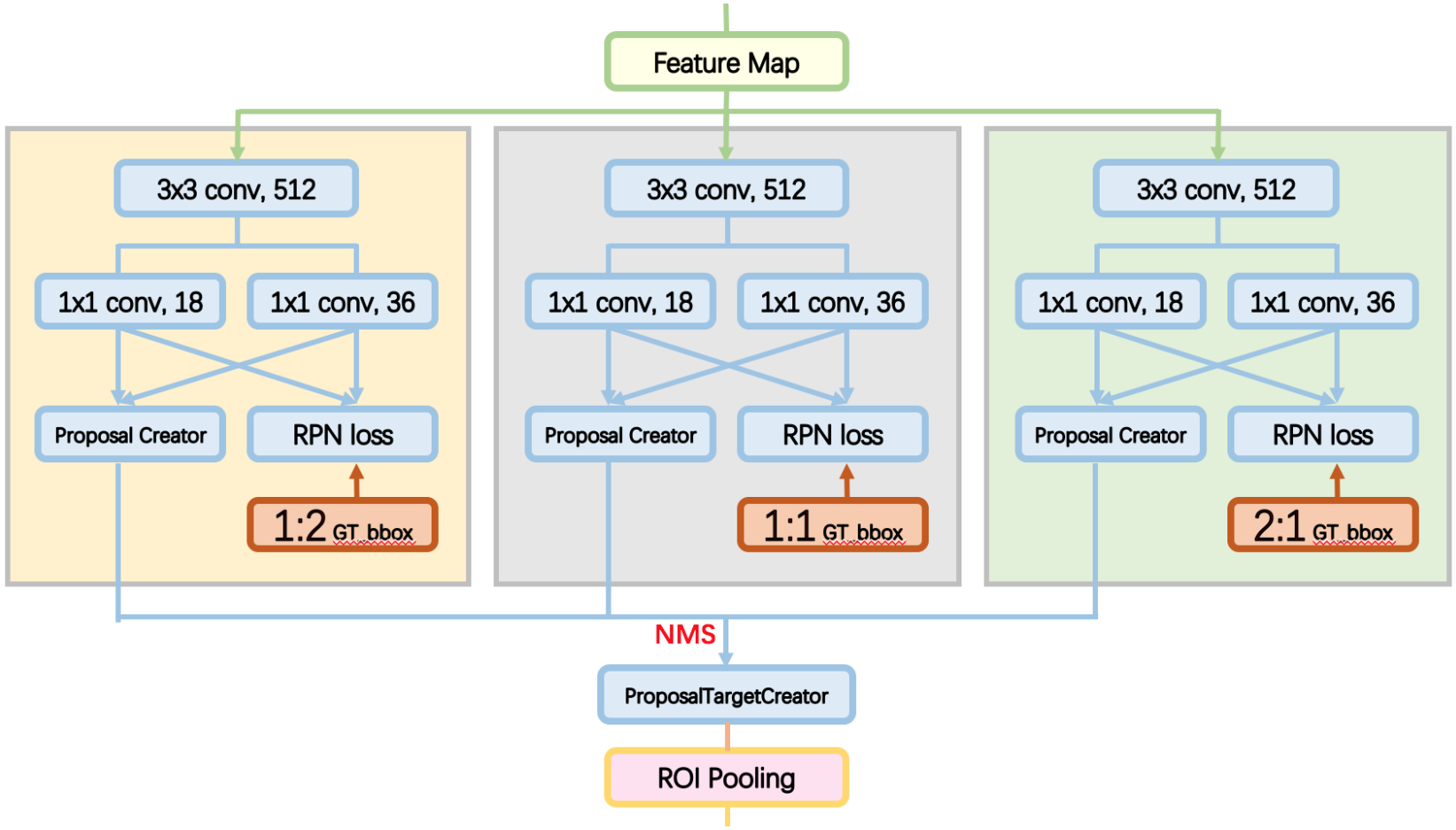


mAP remains unchanged

More False Positive



Multiple RPNs

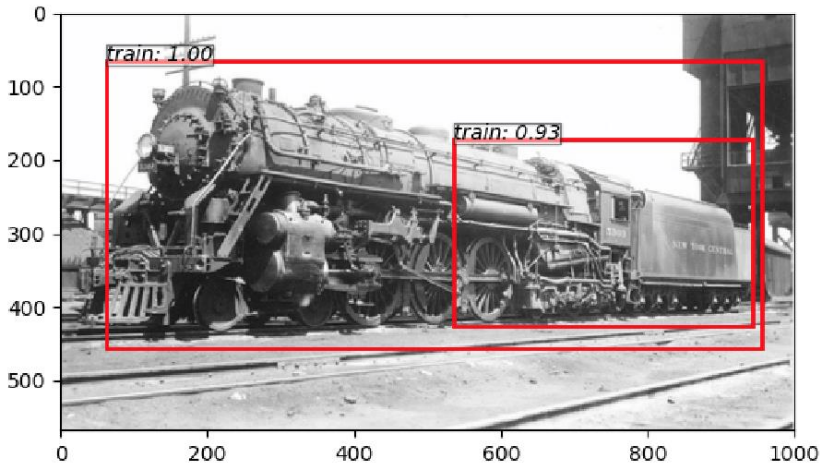


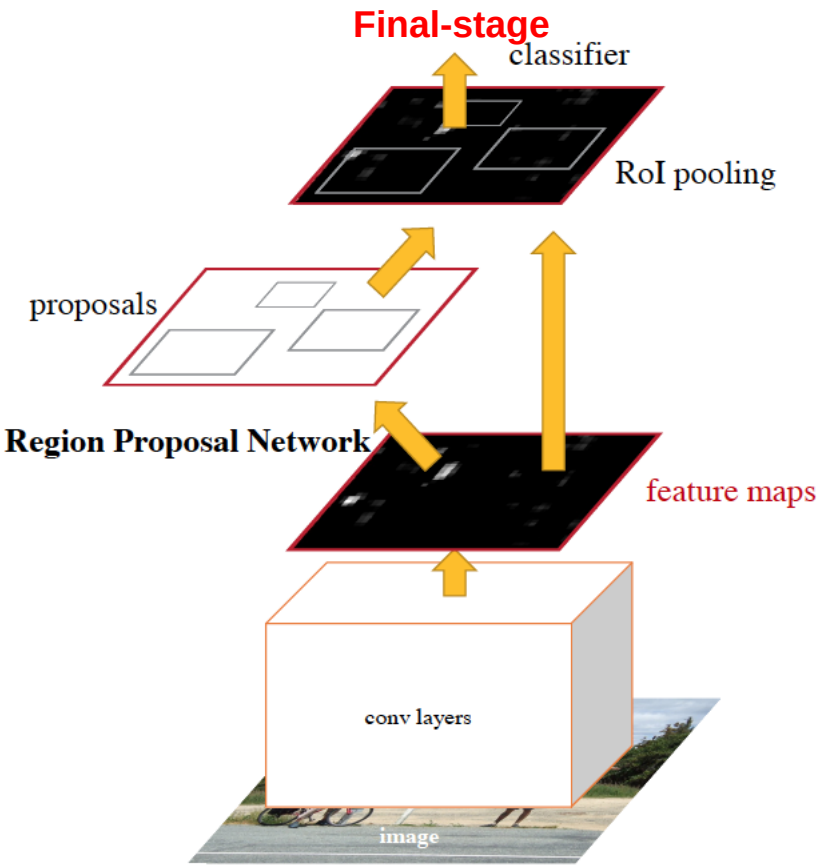
NMS again after RPN?

- Keep the proposal number same (900->300)
- That’s doesn’t work (at this stage)
  - mAP decreased by 1% (RPNs are isolated)

Final-stage Suppression

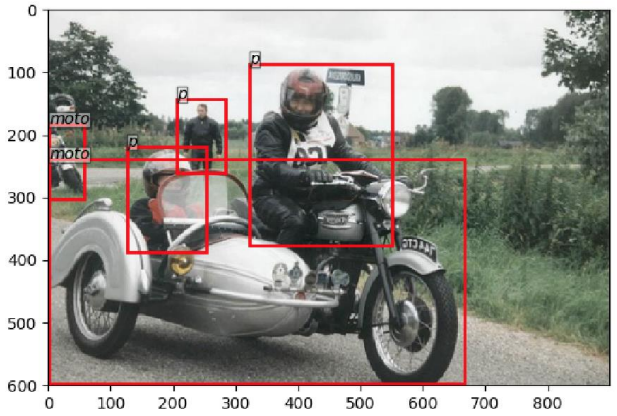
- Overlapping bbox

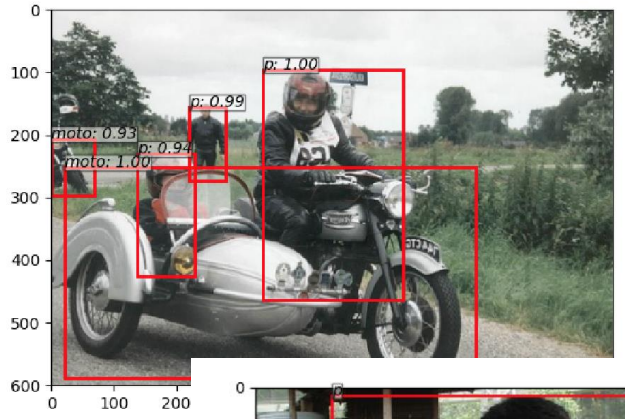


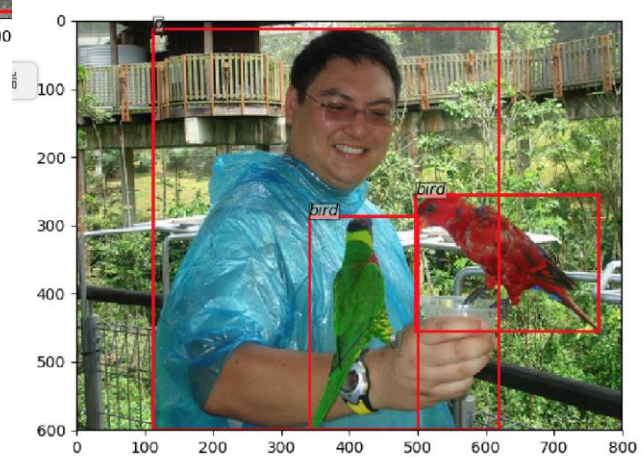


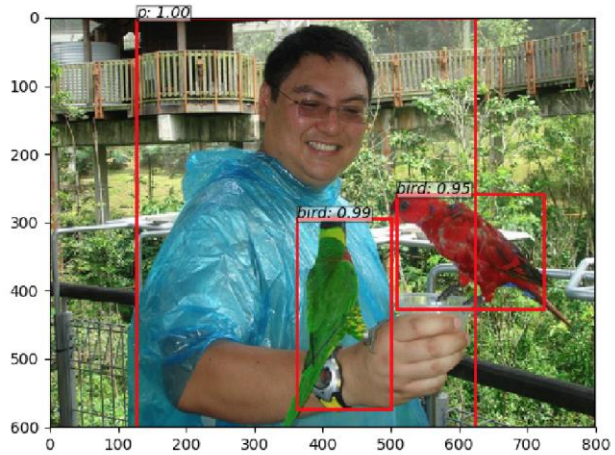
Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. 2015.

Some Results









Contents

Final-stage Suppression

ARSNet, no suppression, 900 proposals	71.7
ARSNet, with suppression, 900 proposals	73.3

Ours vs Original Fast-RCNN

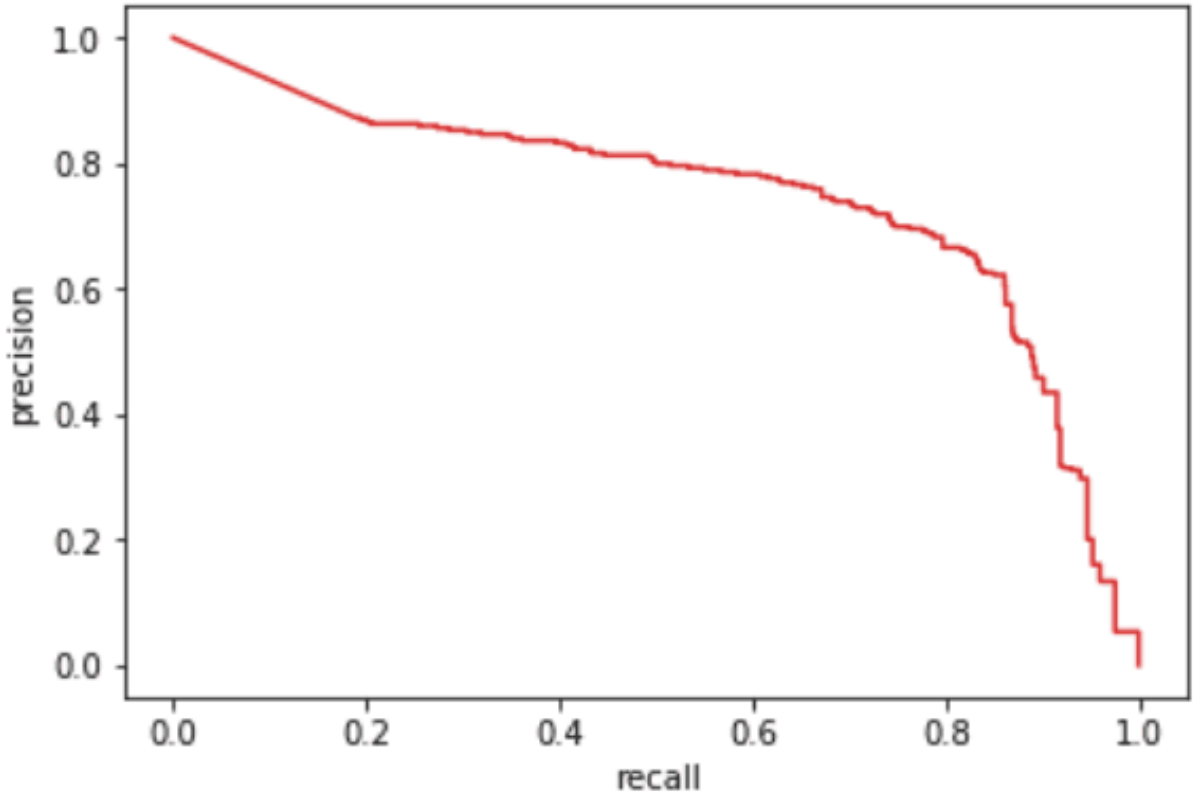
More precious on certain region

method	proposals	mAP	bottle	bus	cow	sofa	plant	sheep
Faster RCNN	300	71.8	52.3	80.7	76.8	65.9	44.7	72.5
Our approach	300	73.0	56.1	83.4	82.9	70.0	40.7	69.7
Our approach	900	73.3	58.0	85.2	81.4	70.7	42.2	69.5

method	proposals	data	mAP	areo	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv
Faster RCNN(Github)	300	VOC07	71.8	73.5	81.5	68.5	53.7	52.3	80.7	85.3	84.3	52.5	76.8	71.5	81.3	84.9	75.1	79.6	44.7	72.5	65.9	79.8	72.3
Our approach	300	VOC07	73.0	75.2	81.3	70.9	55.4	56.1	83.4	83.3	87.4	53.6	82.9	71.5	83.7	87.6	74.6	77.3	40.7	69.7	70.0	81.8	73.9
Our approach	900	VOC07	73.3	75.0	80.6	71.7	56.5	58.0	85.2	83.6	87.9	53.1	81.4	71.9	83.6	87.5	74.2	77.3	42.2	69.5	70.7	82.3	73.4

https://github.com/chenyuntc/simple-faster-rcnn-pytorch

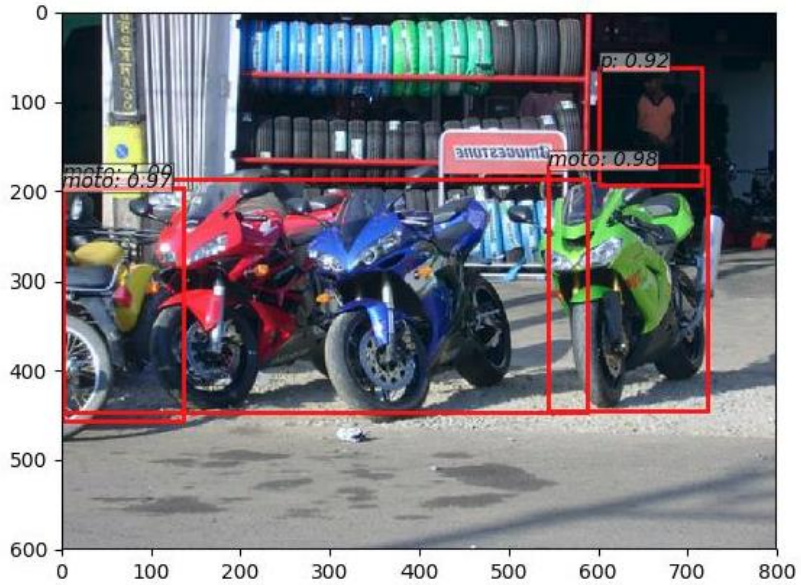
Performance



Contents

Discussion

- Hard negative
  - No obvious improvement on VOC 2007



- Unbalanced size of proposal from each RPN net
- Change ratio (shape) of anchor

Conclusion (Contribution)

- Multiple RPN based on shape (aspect ratio) is simple and useful
  especially for some aspect-ratio sensitive objects.
- More Proposal -> More false positive -> Suppress at the final-stage
- Future Work: Train and Evaluation on coco

Reference

Works mentioned in the presentation only

[1] Redmon, Joseph, et al. "You only look once: Unified, real-time object detection." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

[2] Redmon, Joseph, and Ali Farhadi. "YOLO9000: Better, Faster, Stronger." *Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on*. IEEE, 2017.

[3] M. Everingham, L. Van Gool, C. Williams, J. Winn, A. Zisserman, *The PASCAL Visual Object Classes Challenge 2007 (VOC2007)*.

[4] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems*. 2015.

[5] Singh, Bharat, and Larry S. Davis. "An Analysis of Scale Invariance in Object Detection-SNIP." , 2017.

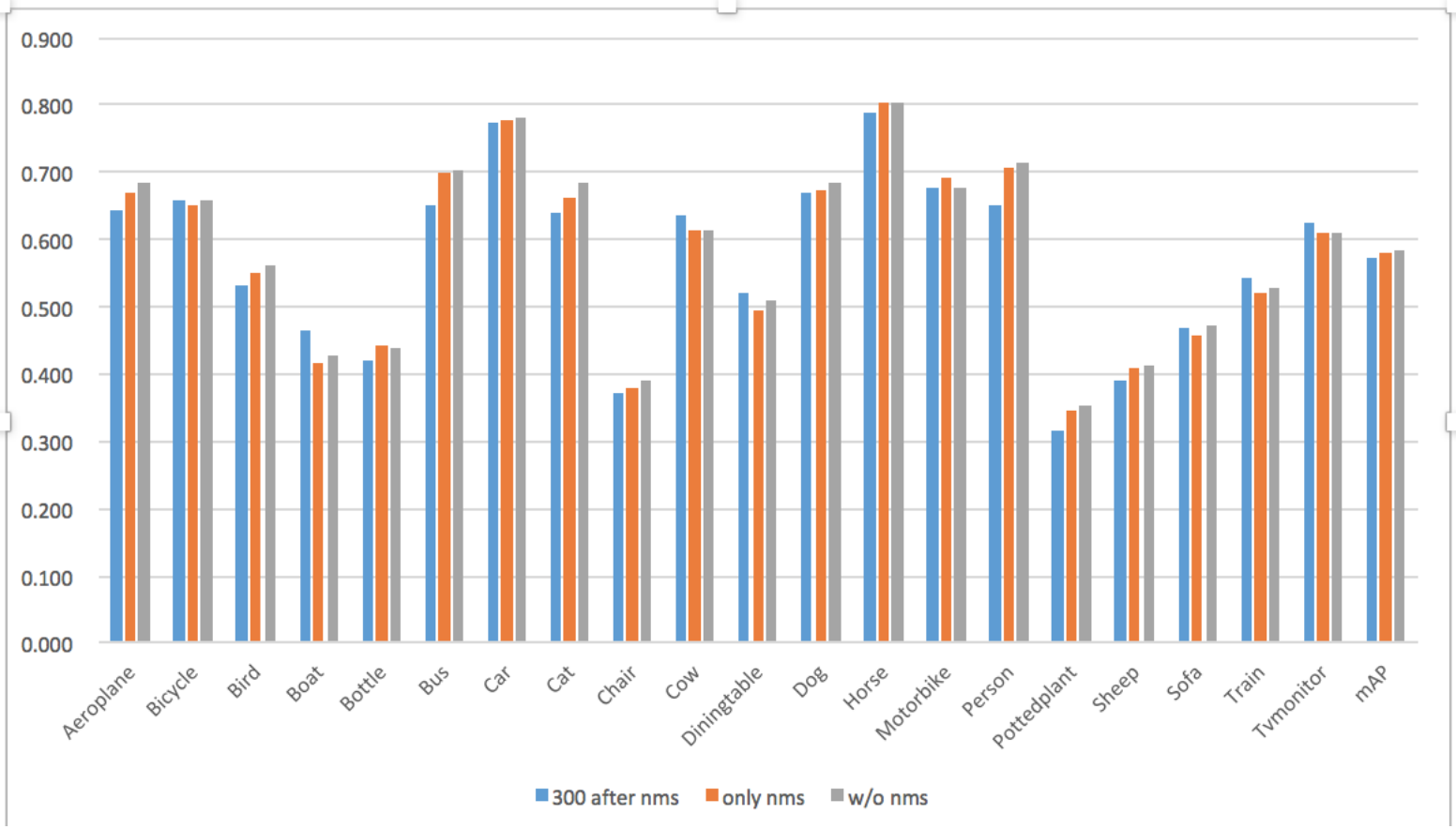
[6] Li, Bo, et al. "Object Detection via Aspect Ratio and Context Aware Region-based Convolutional Networks.", 2016.

[7] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. Technical report, 2014.

Q&A

Appendix

Result of NMS immediately after RPNs



Appendix: Performance

method	proposals	data	mAP
Faster RCNN	300	VOC07	71.8
Our approach	300	VOC07	73.0
Our approach	900	VOC07	73.3

Appendix: number of proposal

method	proposals	data	mAP
Faster RCNN	300	VOC07	71.8
ARSNet, no suppression	900	VOC07	71.7

Appendix: Related Ideas

- Rotated bounding box
  - Less background
  - Ground truth?
  - Split RPN by rotation