MSCOCO Instance Segmentation Challenges 2018

Megvii (Face++) Team
lizeming@megvii.com

I. COCO’18 Instance Seg

Zeming LI

Yueqing ZHUANG

Xiangyu ZHANG

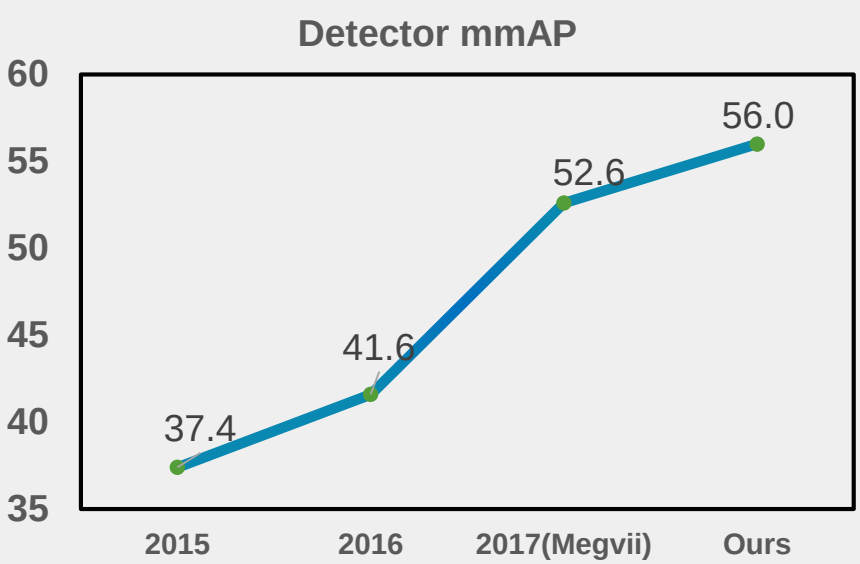
Gang YU

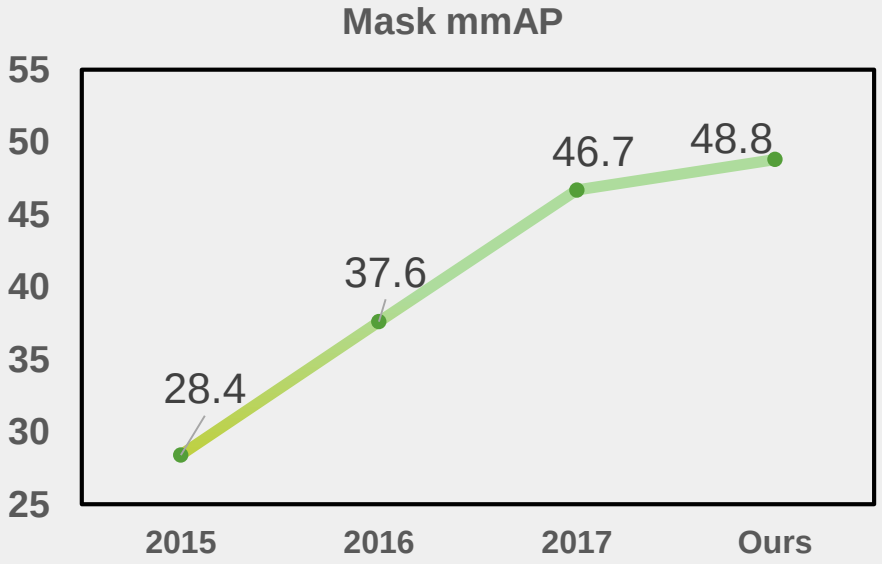
Jian SUN

Overview

Improvements

The results is obtained on test-dev





Object Detector

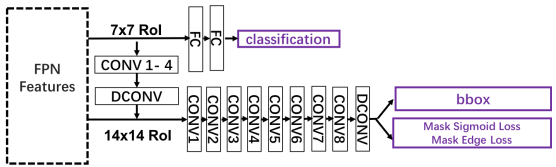
3.4% improvement

Instance Segmentation

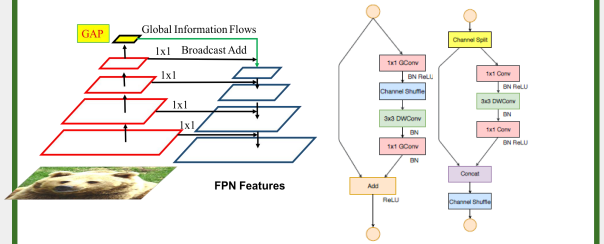
2.1% improvement

Outline

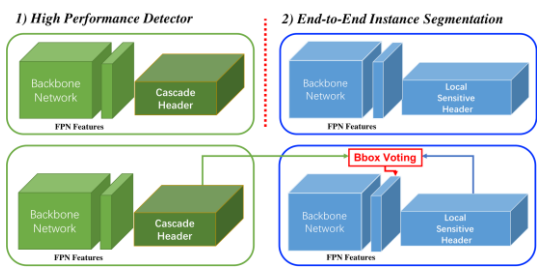
Face++ 旷视



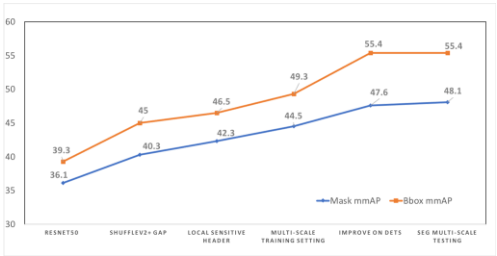
1) Location Sensitive Header



2) Backbone Improvement

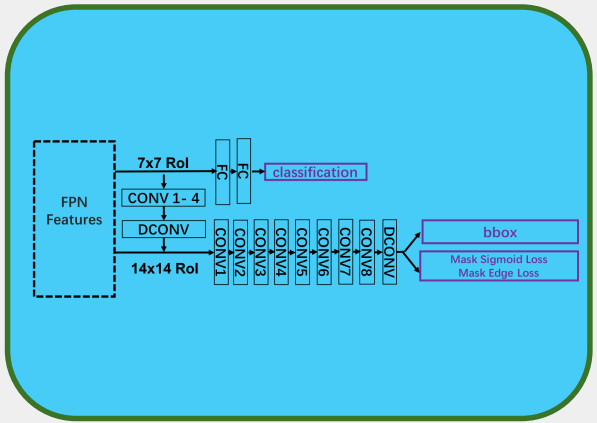


3) Two-Pass Pipeline

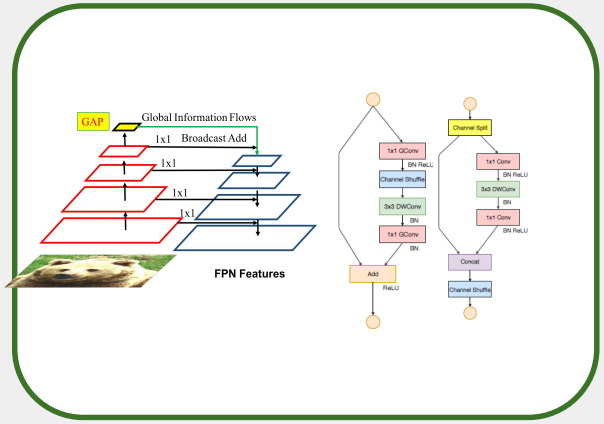


4) Results

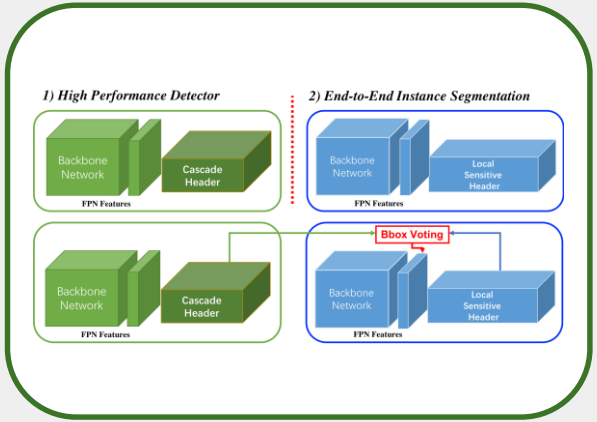
Outline



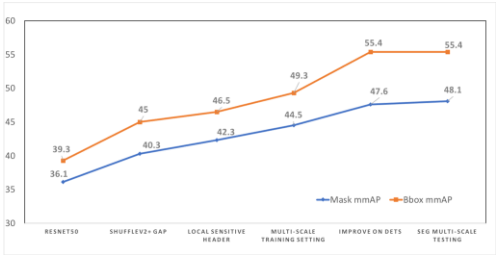
1) Location Sensitive Header



2) Backbone Improvement

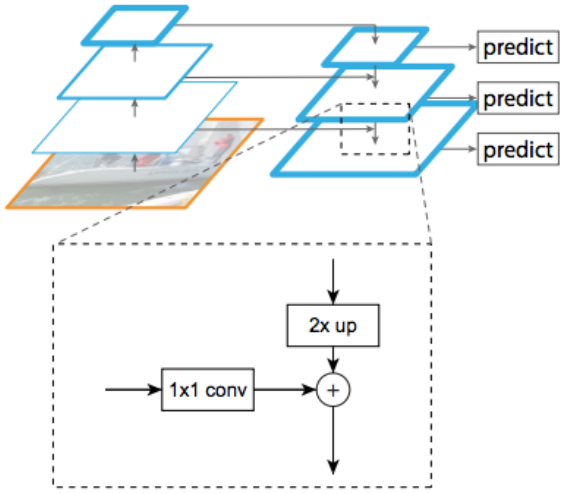


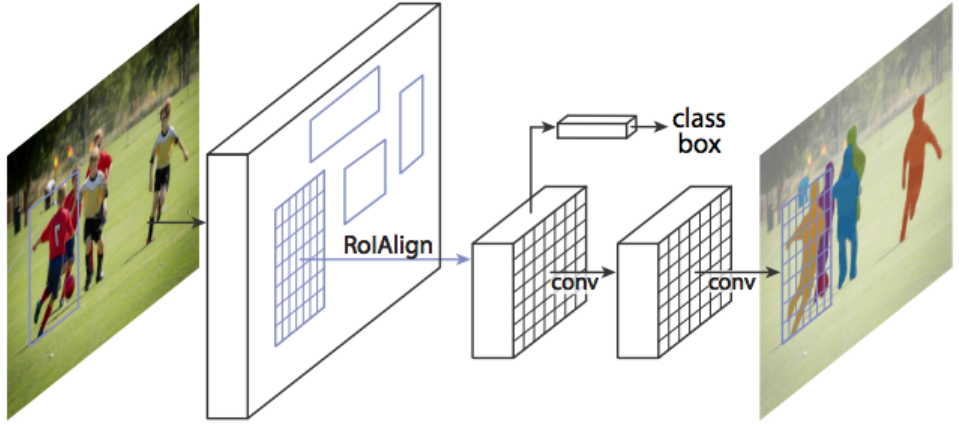
3) Two-Pass Pipeline



4)Results

Mask RCNN Baseline





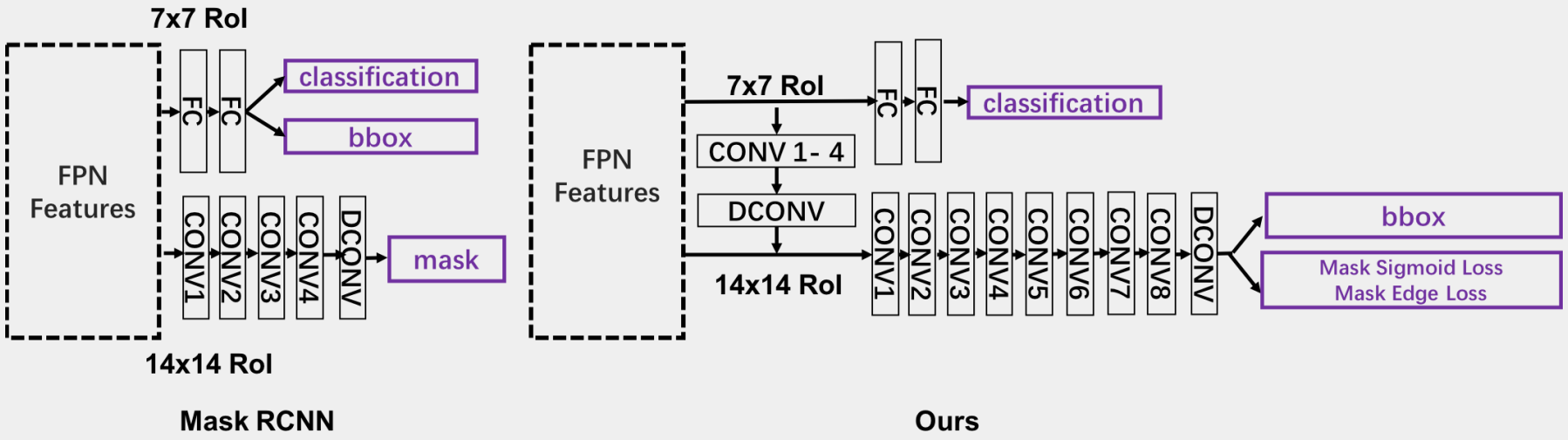
FPN

Original Mask Head

	Instance Seg mmAP	Det mmAP
Original Paper(detectron 1x)	33.6	-
Our Re-implement	34.4	37.0

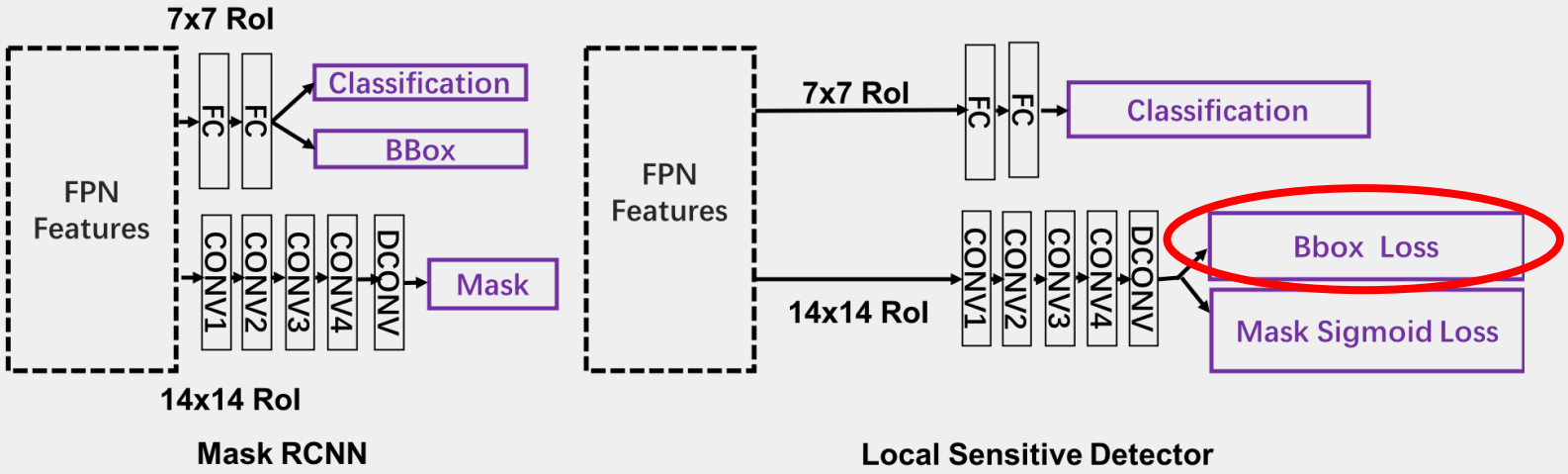
Location Sensitive Header

Overall Architecture Comparison



Location Sensitive Header

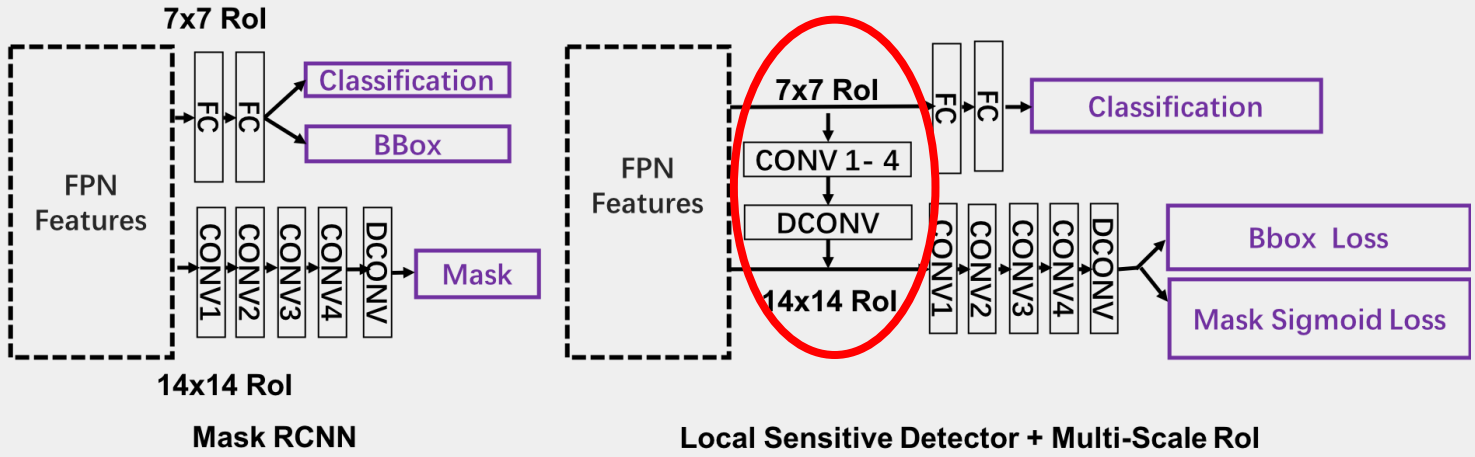
1) Location Sensitive Detector



name	Mask AP	Bbox AP	Improvement
Baseline	34.4	37.0	-
+ Local Sensitive Detector	35.4	38.7	+ 1.0 / +1.7

Location Sensitive Header

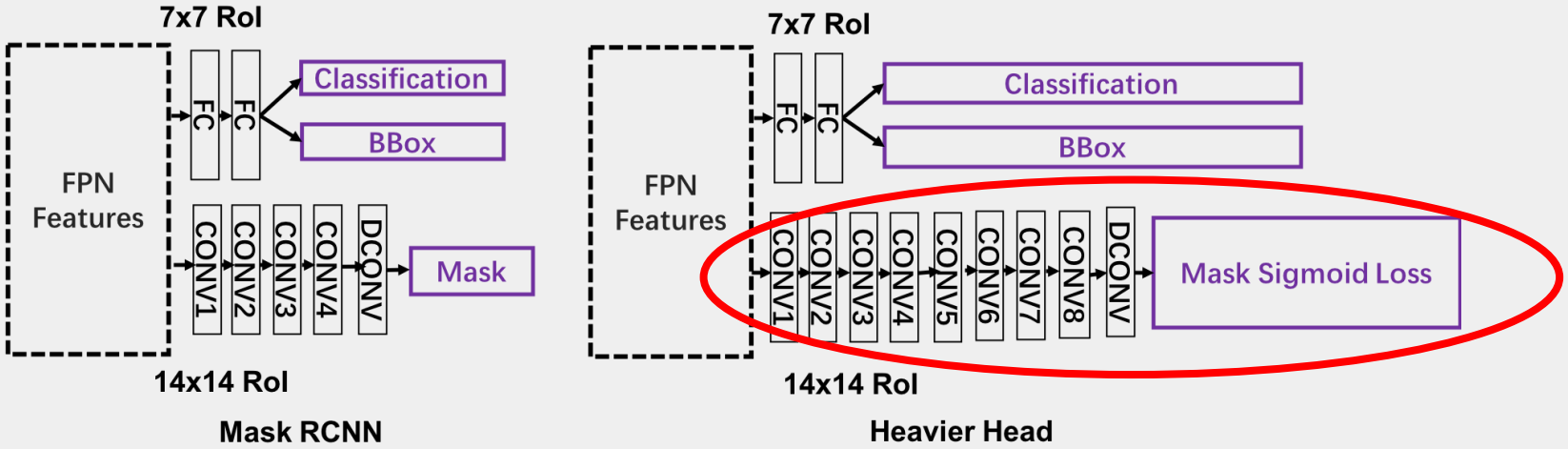
2) Multi-Scale RoI



name	Mask AP	Bbox AP	Improvement
Baseline	34.4	37.0	-
+ Local Sensitive Detector	35.6	38.7	+ 1.0 / +1.7
+ Multi-Scale RoI	35.8	38.9	+ 0.2 / +0.2

Location Sensitive Header

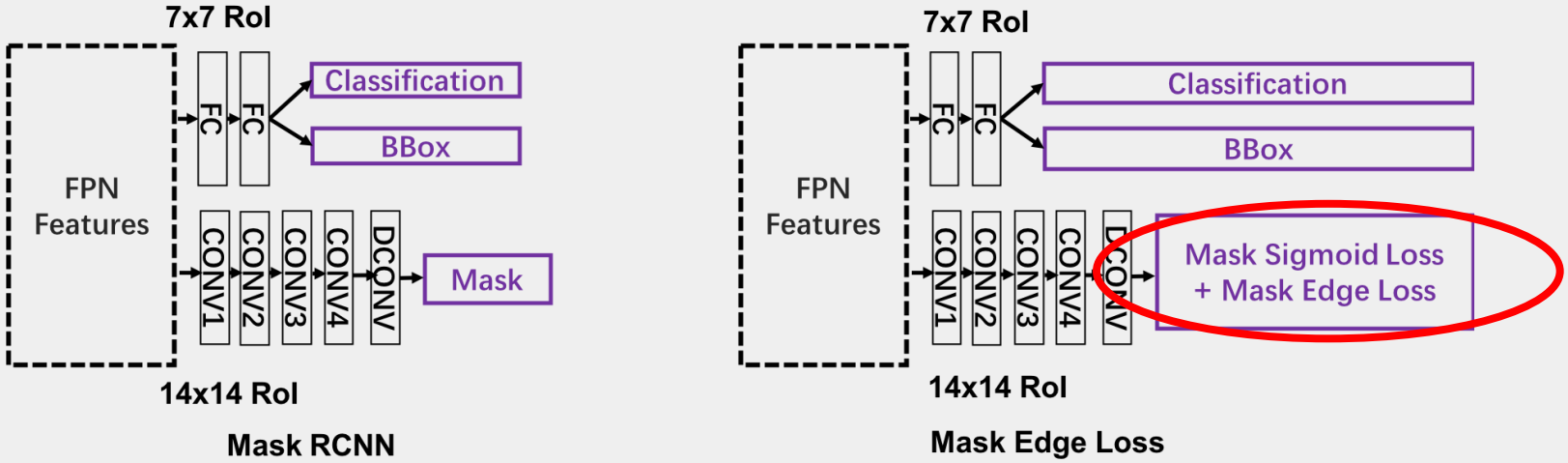
3) Heavier Header



name	Mask AP	Bbox AP	Improvement
Baseline	34.4	37.0	-
Heavier Header	35.3	36.8	+ 0.9 / -0.2

Location Sensitive Header

4) Mask Edge Loss



name	Mask AP	Bbox AP	Improvement
Baseline	34.4	37.0	-
Mask Edge Loss	35.0	37.0	+ 0.6 / +0.0

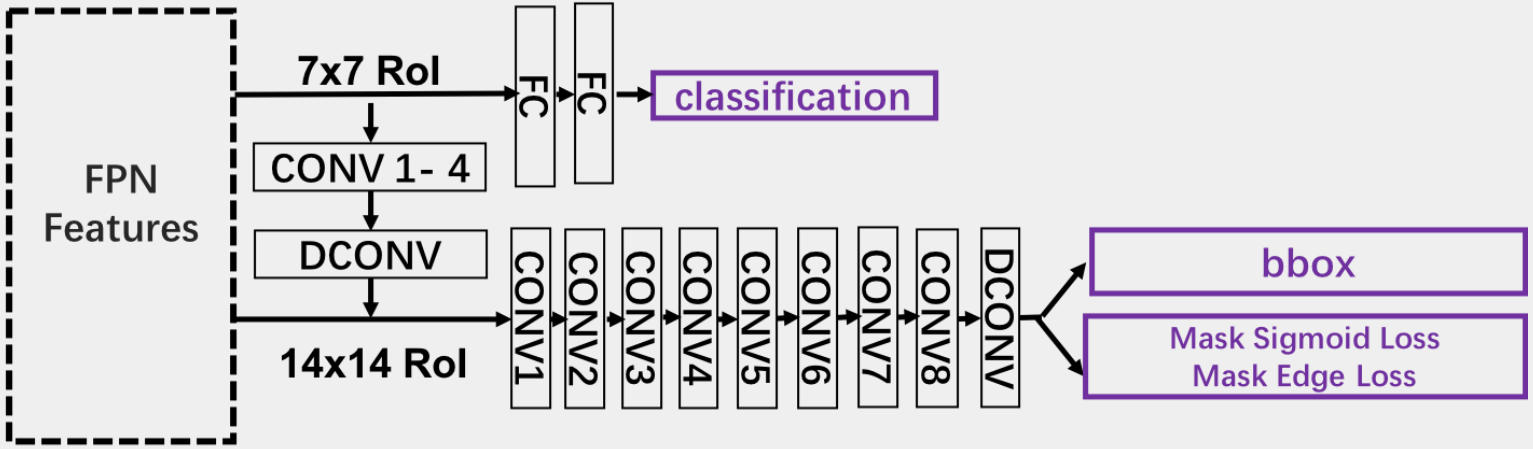
Location Sensitive Header

4) Mask Edge Loss

Sigmoid Cross Entropy

Location Sensitive Header

Review of overall Architecture



Location Sensitive Header:

1) **Location Sensitive Detector**
2) **Multi-Scale RoI**
3) **Heavier Header**
4) **Mask Edge Loss**

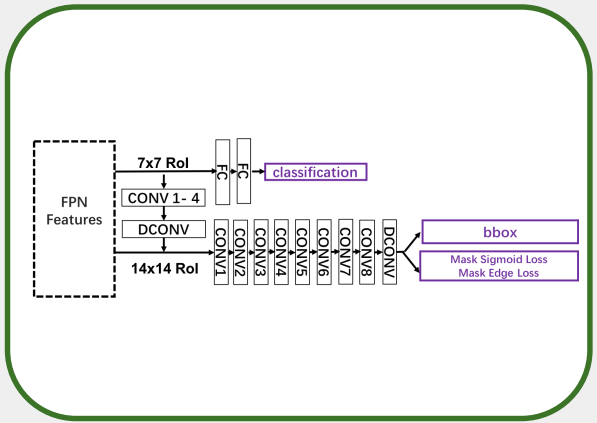
Location Sensitive Header

Overall Performance in Small and Large Model

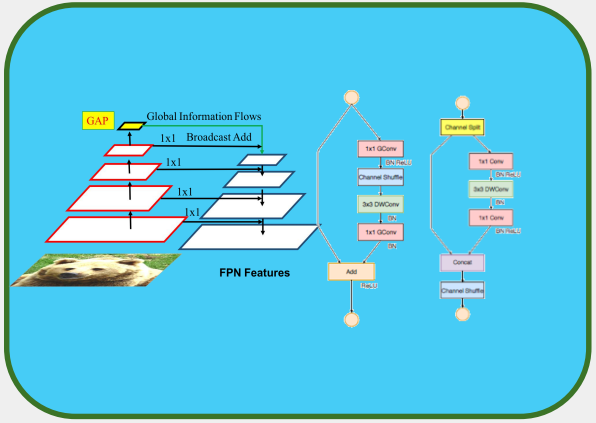
BackBone	Header	Mask AP	Bbox AP	Improvement
ResNet50	Baseline	34.4	37.0	**-**
	Location Sensitive Header	37.0	39.3	+ 2.6 / + 2.0
ShuffleV2-GAP	Baseline	40.3	45.0	-
	Location Sensitive Header	42.3	46.5	+2.0/+1.5

We will introduce backbone in next slides

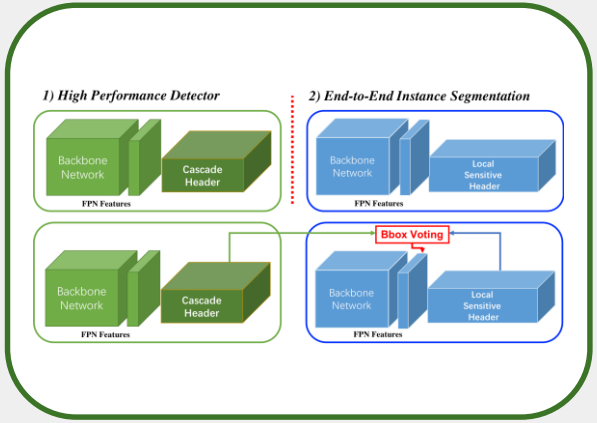
Outline



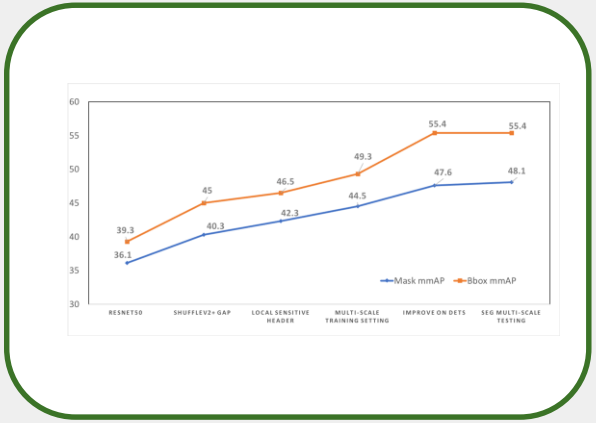
1) Location Sensitive Header



2) Backbone Improvement



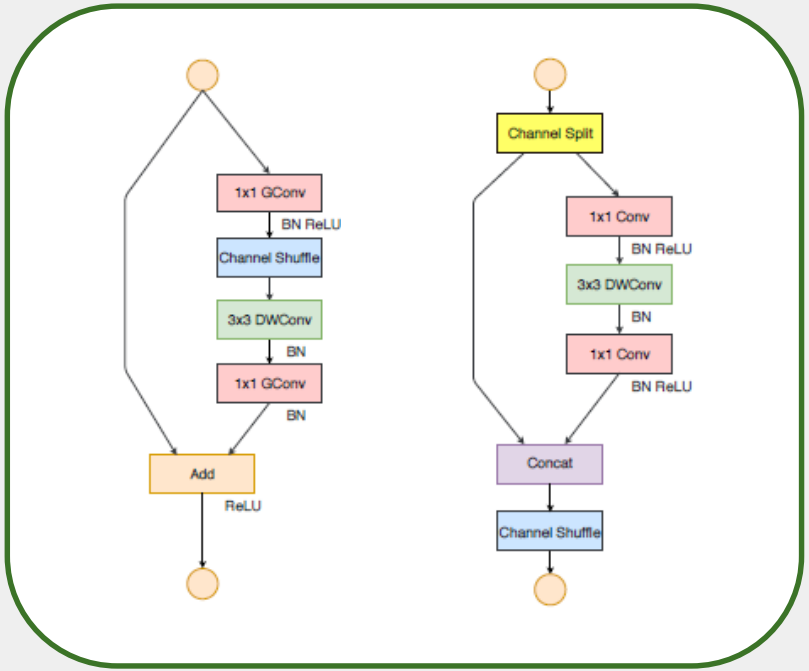
3) Two-Pass Pipeline



4)Results

Backbone Improvement

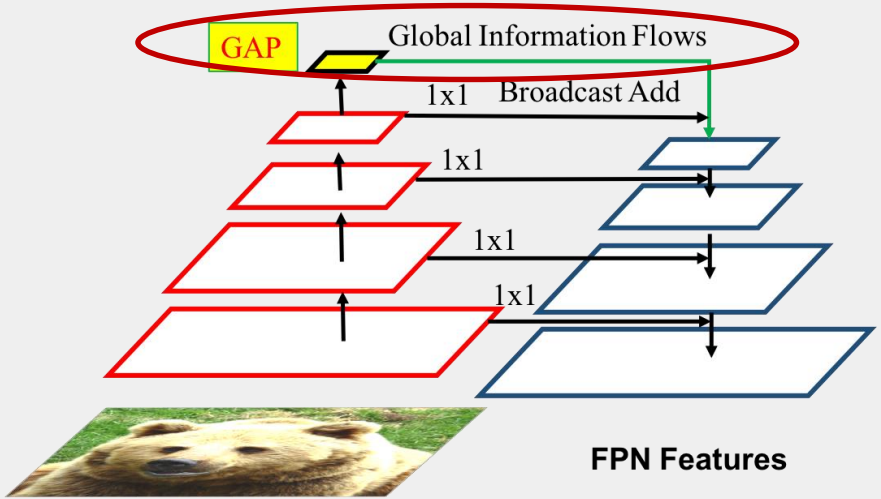
1. Channel Information Flow



Ma N, Zhang X, Zheng H T, et al. ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design[J]. 2018.

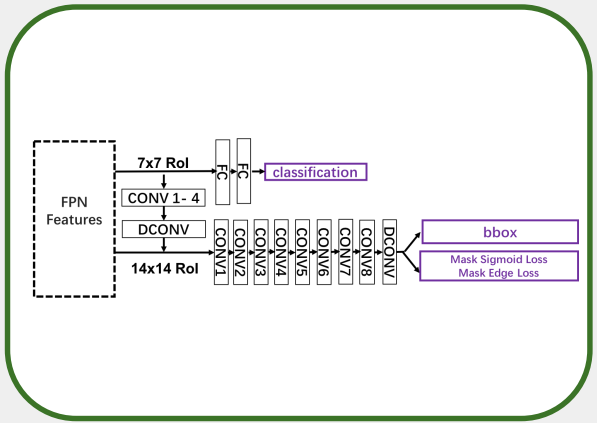
Backbone Improvement

2. Add Global Information

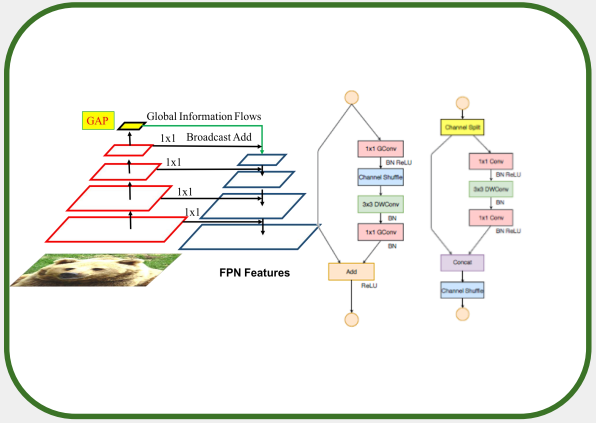


name	Mask AP	Bbox AP	Improvement
Baseline	34.4	37.0	-
+GAP	35.1	37.7	+0.7/+ 0.7

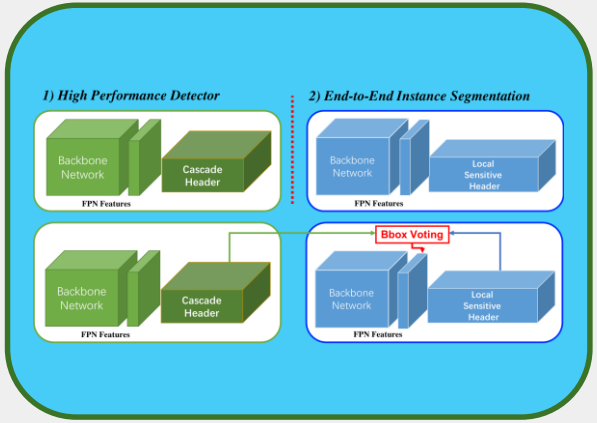
Outline



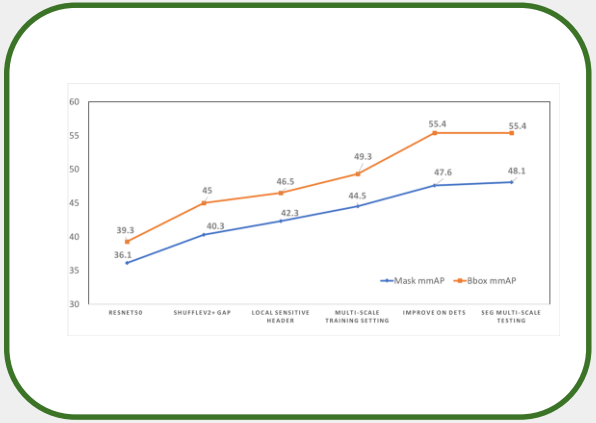
1) Location Sensitive Header



2) Backbone Improvement

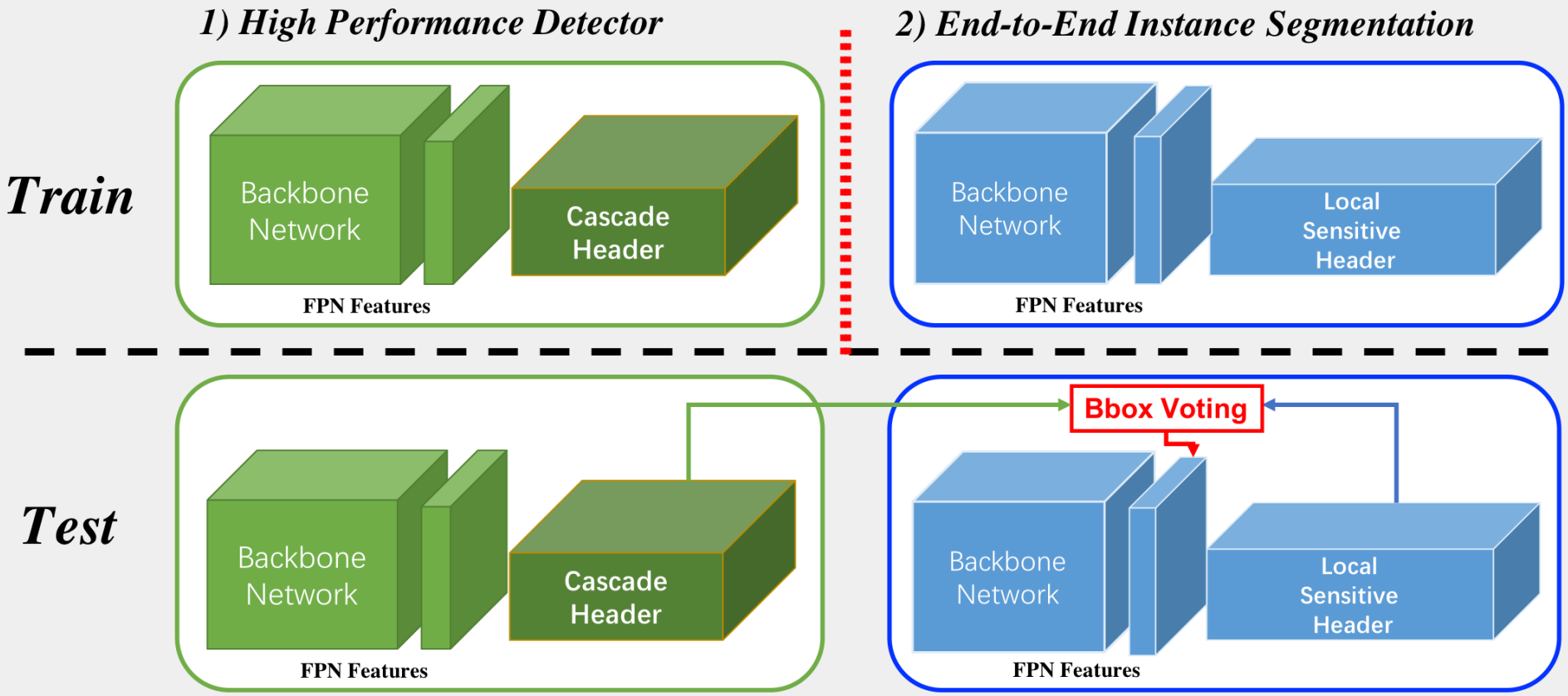


3) Two-Pass Pipeline

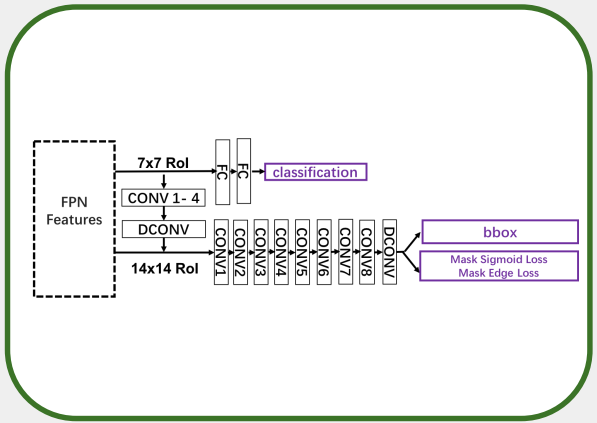


4)Results

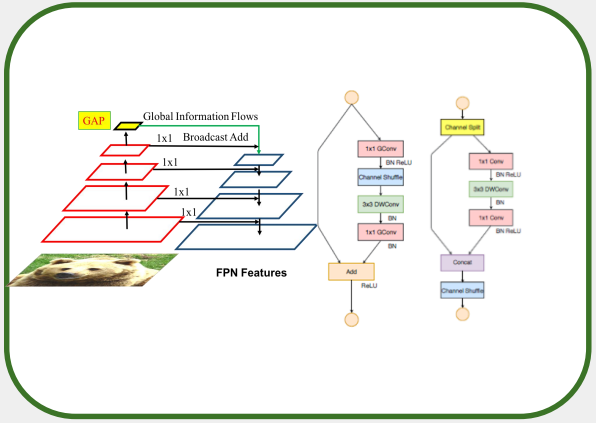
Two-Pass Pipeline



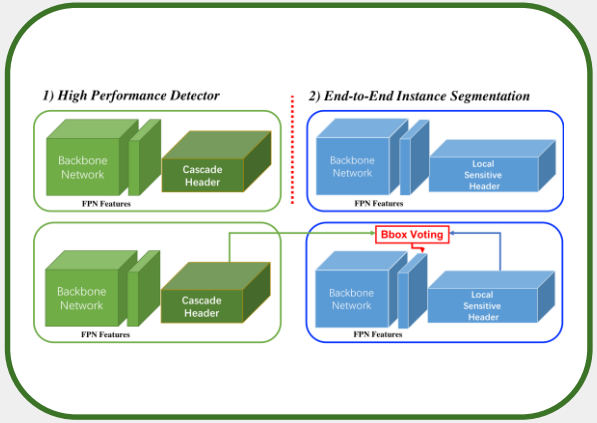
Outline



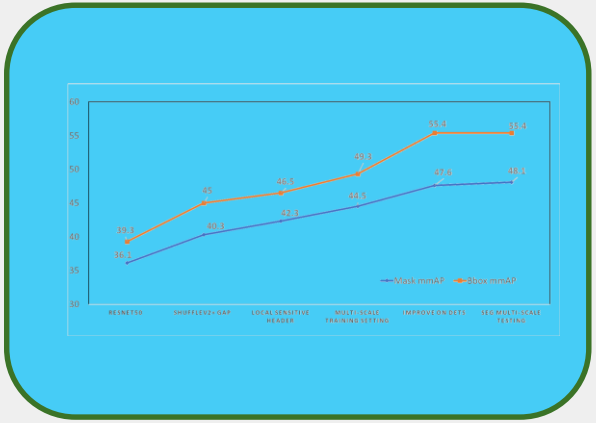
1) Location Sensitive Header



2) Backbone Improvement



3) Two-Pass Pipeline



4)Results

Results

Trained On Megvii's Megbrain

name	Mask AP(val)	Bbox AP(val)	Improvement
ResNet50 (2x-2batch-setting)	36.1	39.3	-
ShuffleV2 (1batch)	40.3	45.0	+3.8/+5.7

2x Means 2x training setting used in Detectron

Results

Trained On Megvii's Megbrain

name	Mask AP(val)	Bbox AP(val)	Improvement
ResNet50 (2x-2batch-setting)	36.1	39.3	-
ShuffleV2 (1batch)	40.3	45.0	+3.8/+5.7
+ Location Sensitive Header	42.3	46.5	+2.0 /+1.5

Results

Trained On Megvii's Megbrain

name	Mask AP(val)	Bbox AP(val)	Improvement
ResNet50 (2x-2batch-setting)	36.1	39.3	-
ShuffleV2 (1batch)	40.3	45.0	+3.8/+5.7
+ Local Sensitive Header	42.3	46.5	+2.0 /+1.5
+ 2 Batch Per GPU + Multi Scale Training + BN training	44.5	49.3	+2.2/ 2.8

Results

Trained On Megvii's Megbrain

name	Mask AP(val)	Bbox AP(val)	Improvement
ResNet50 (2x-2batch-setting)	36.1	39.3	-
ShuffleV2 (1batch)	40.3	45.0	+3.8/+5.7
+ Local Sensitive Header	42.3	46.5	+2.0 /+1.5
+ 2 Batch Per GPU + Multi Scale Training + BN training	44.5	49.3	+2.2/ 2.8
+ Improve on Dets	47.6	55.4	+3.1/ 6.1

Results

Trained On Megvii's Megbrain

name	Mask AP(val)	Bbox AP(val)	Improvement
ResNet50 (2x-2batch-setting)	36.1	39.3	-
ShuffleV2 (1batch)	40.3	45.0	+3.8/+5.7
+ Local Sensitive Header	42.3	46.5	+2.0 /+1.5
+ 2 Batch Per GPU + Multi Scale Training + BN training	44.5	49.3	+2.2/ 2.8
+ Improve on Dets	47.6	55.4	+3.1/ 6.1
+ Seg Multi-scale Testing	48.1	55.4	+0.5/0.0

Results

Trained On Megvii's Megbrain

name	Mask AP(val)	Bbox AP(val)	Improvement
ResNet50 (2x-2batch-setting)	36.1	39.3	-
ShuffleV2 (1batch)	40.3	45.0	+3.8/+5.7
+ Local Sensitive Header	42.3	46.5	+2.0 /+1.5
+ 2 Batch Per GPU + Multi Scale Training + BN training	44.5	49.3	+2.2/ 2.8
+ Improve on Dets	47.6	55.4	+3.1/ 6.1
+ Seg Multi-scale Testing	48.1/ 48.8(dev)	55.4/ 56.0(dev)	+0.5/0.0

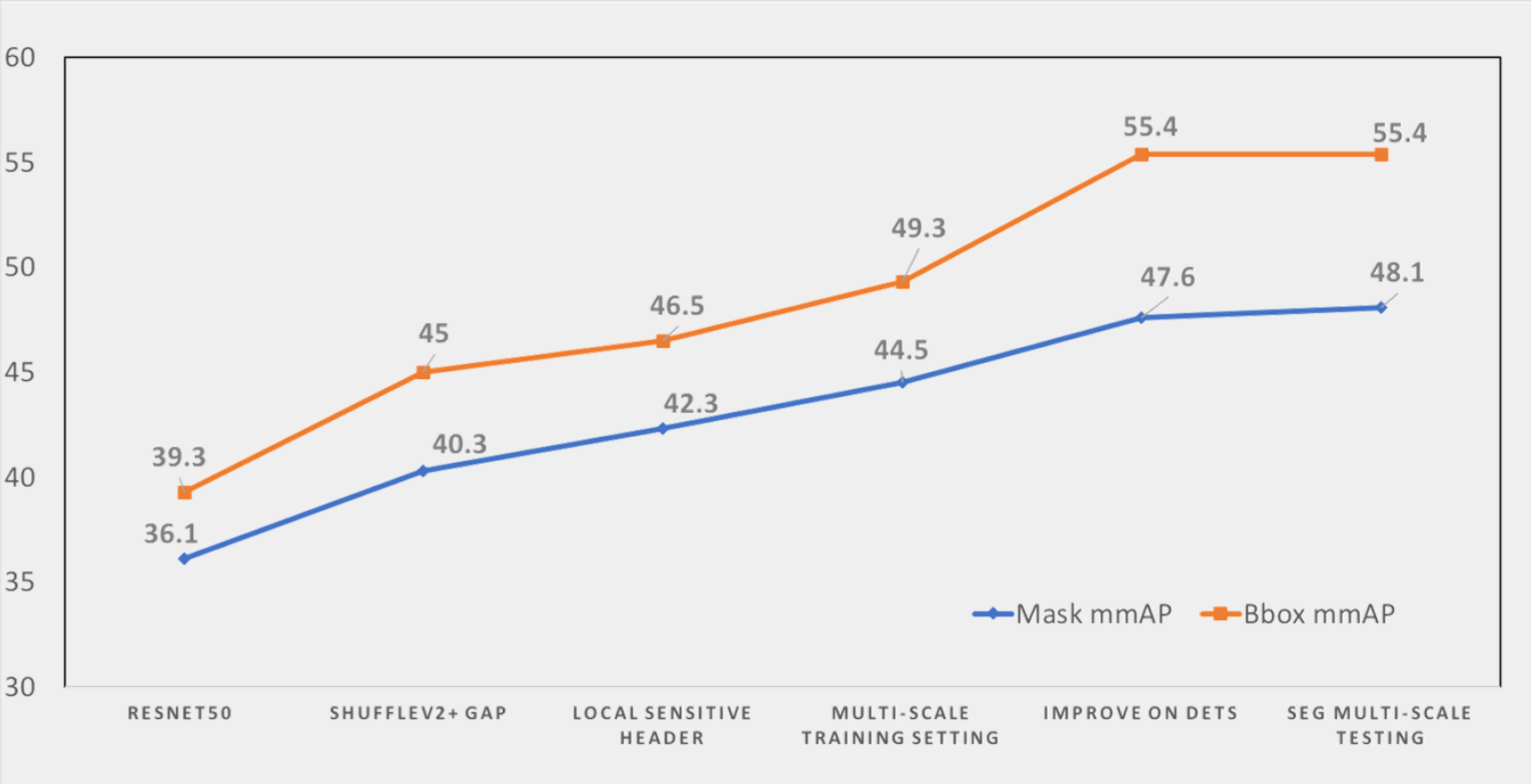
Instance Segmentation is obtained by single instance segmentation model

Results

Trained On Megvii's Megbrain

name	Bbox AP(val)	Improvement
Baseline	49.3	-
+Soft-Nms	49.8	+0.5
+Multi-scale Testing	51.6	+1.8
+Ensemble	53.6	+2.0
add an additional model for ensemble: +with cascade R-CNN +external COCO++ 11W data	55.4	+1.8

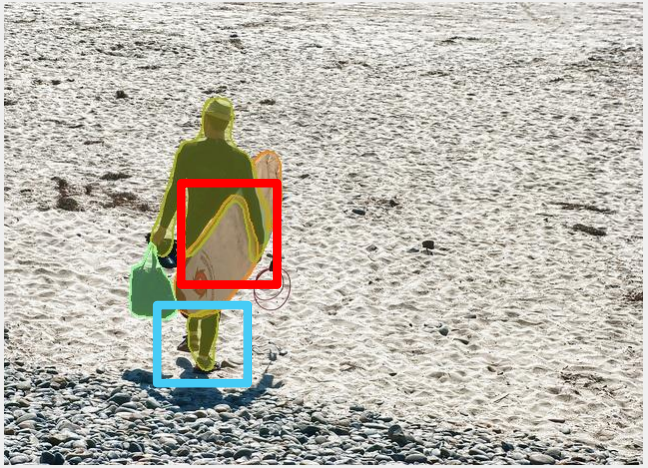
Results


60
55
50
45
40
35
30
39.3
45
46.5
49.3
55.4
55.4
36.1
40.3
42.3
44.5
47.6
48.1
Mask mmAP
Bbox mmAP
RESNET50
SHUFFLEV2+ GAP
LOCAL SENSITIVE HEADER
MULTI-SCALE TRAINING SETTING
IMPROVE ON DETS
SEG MULTI-SCALE TESTING

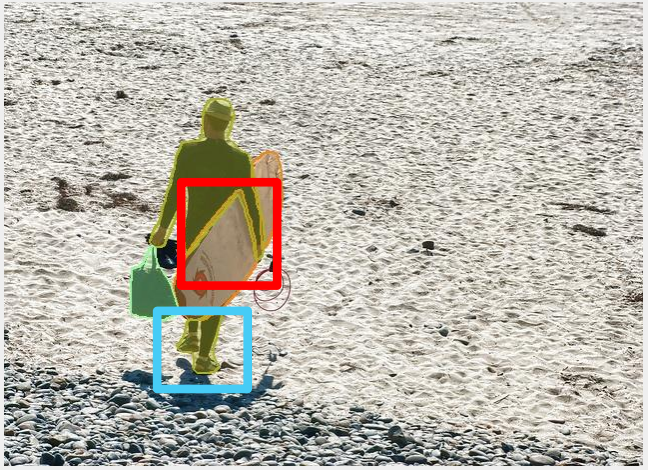
Visualization Comparison

Face++ 旷视

Our baseline

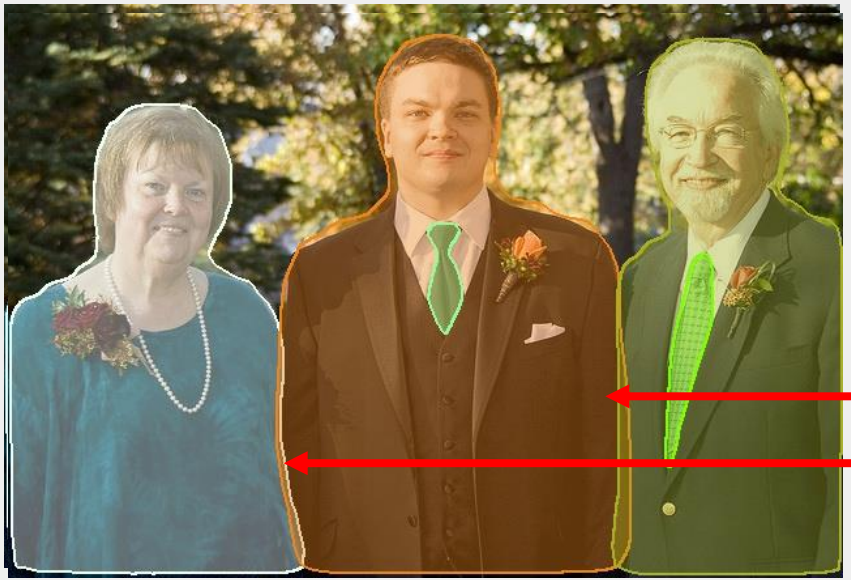


Location Sensitive Header

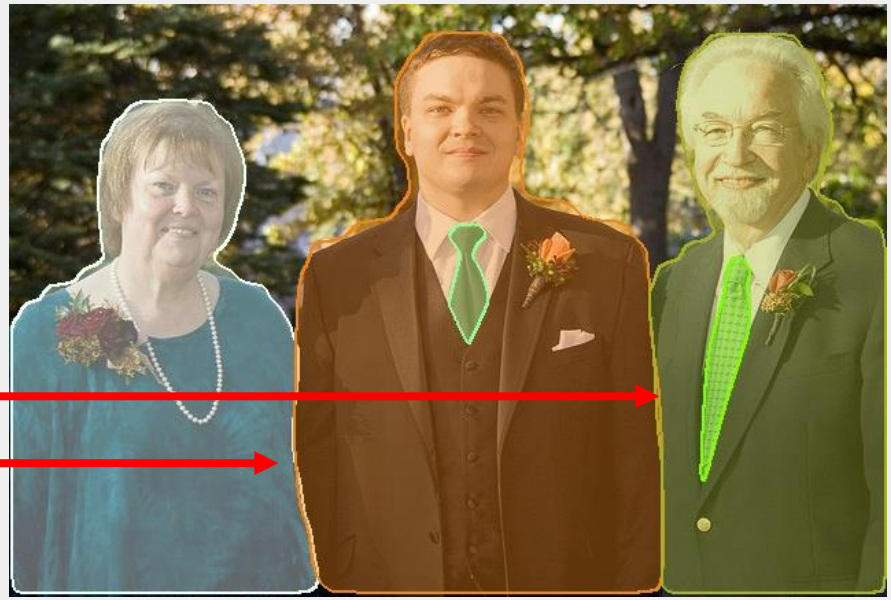


Refine Location Error

Visualization Comparison



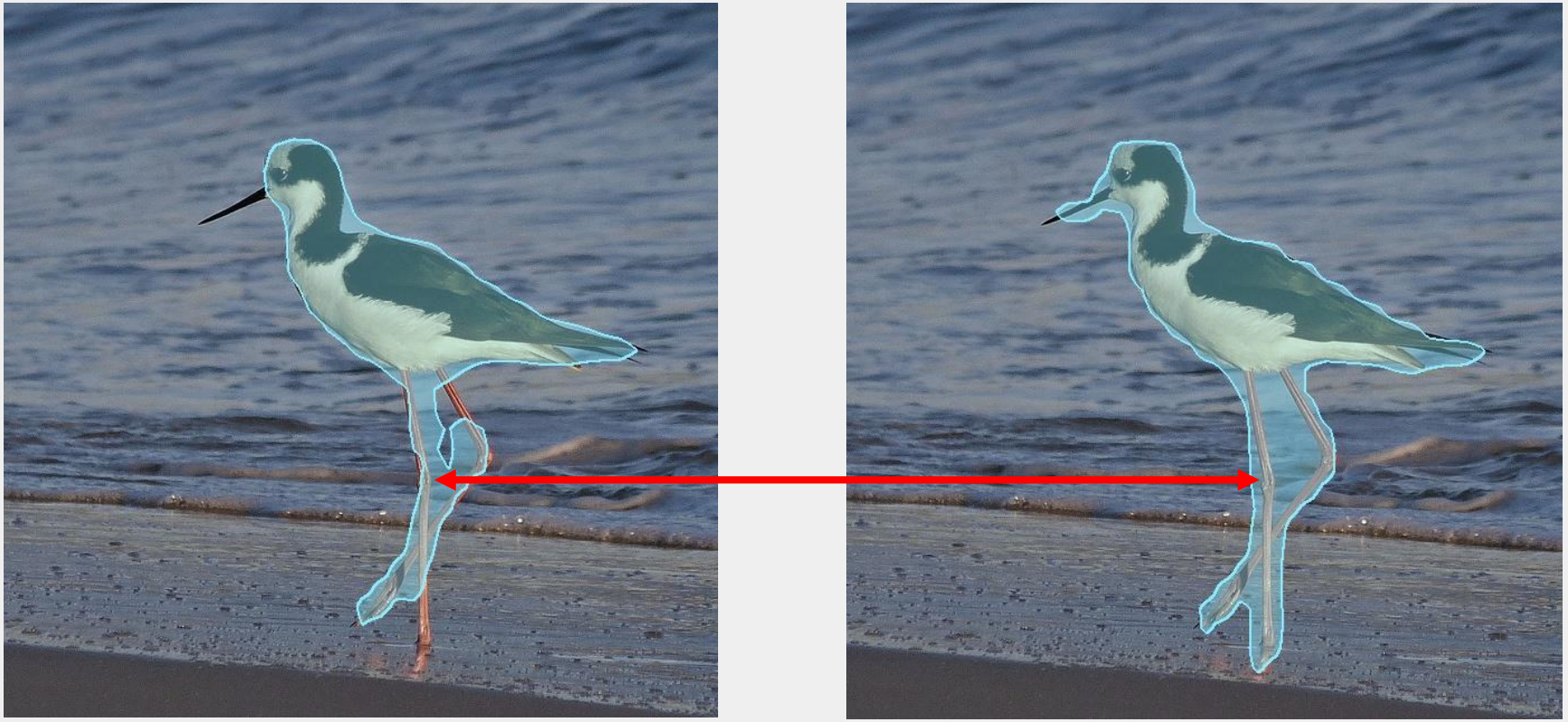
Our Baseline



Location Sensitive Header

Refine Location Error

Visualization Comparison



Our Baseline

Location Sensitive Header

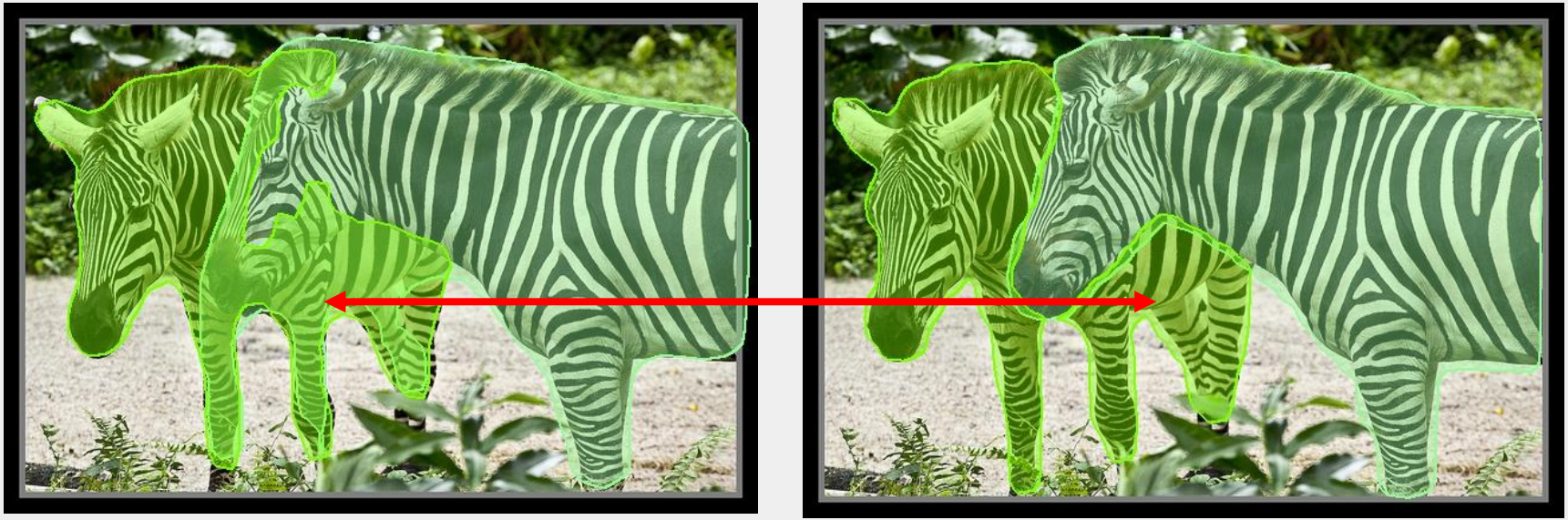
Refine Location Error

Visualization Comparison

Our Baseline

Location Sensitive Header

Visualization Comparison



Our Baseline

Location Sensitive Header

Visualization

Detector Results

Mask Results

Visualization

Detector Results

Mask Results

Summary & thanks

1. Location Sensitive Header
2. Backbone Improvement
3. Pipeline Optimization

Other Improvements:

1. Multi-Scale Training
2. Large Batch (MegDet : [C. Peng, CVPR’ 18])
3. Multi-Scale and Flip Testing
4. Ensemble (only for Detection)

Looking for Interns, Researcher，
Research Engineer

career@megvii.com

yugang@megvii.com