



3rd Detection / Segmentation Challenge

Tsung-Yi Lin



COCO and Places Visual Recognition Challenges Workshop

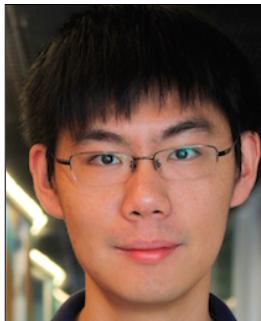
Sunday, October 28th, ICCV 2017



Workshop Organizers



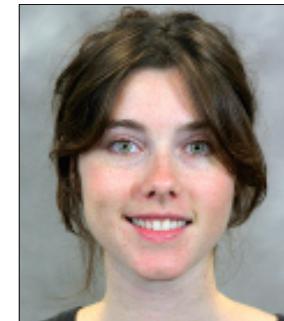
Tsung-Yi Lin
Google Brain



Yin Cui
Cornell Tech



**Matteo Ruggero
Ronchi**
Caltech



**Genevieve
Patterson**
Microsoft Research



**Holger
Caesar**
University of Edinburgh

Workshop Advisors:

Michael Maire
Serge Belongie
Lubomir Bourdev
Ross Girshick
James Hays
Pietro Perona
Larry Zitnick
Piotr Dollár

Award Committee:

Deva Ramanan
Pietro Perona
James Hays
Lubomir Bourdev
Serge Belongie
Michael Maire
Matteo Ruggero Ronchi
Genevieve Patterson
Yin Cui

Vittorio Ferrari
Jasper Uijlings
Holger Caesar



COCO Dataset



- 80 object categories
- 160k images
- 1M instances (350k people)
- Every instance segmented
- 106k people with keypoints
- **65k images with stuff segm**



Available for download at
cocodataset.org





COCO Website on Github

[This organization](#)

[Search](#)

[Pull requests](#)

[Issues](#)

[Marketplace](#)

[Explore](#)

cocodataset

[Repositories 2](#)

[People 2](#)

[Teams 0](#)

[Projects 0](#)

[Settings](#)

[Type: All ▾](#)

[Language: All ▾](#)

[cocoapi](#)

COCO API - Dataset @ <http://cocodataset.org/>

[Jupyter Notebook](#)

[★ 691](#)

[366](#)

[BSD-2-Clause](#)

[Updated 12 days ago](#)

[cocodataset.github.io](#)

[HTML](#)

[★ 10](#)

[1](#)

[Updated 29 days ago](#)

4 /24



COCO Train/val Sets

Train2017 = Train2014 + Val2014 - Val2017

115k images

Val2017 = 5k subset of Val2014

5k images

No image added, only reorganization.



COCO Test Sets

The 2017 COCO Test set consists of ~40k test images. Test splits stayed the same.

Test-dev (publications)

Used to score entries for the Public Leaderboard

Test-challenge (competitions)

Used to score workshop competition.

Test-standard and Test-reserve

Dropped



Challenges at ICCV 2017



coco

Common Objects in Context



Detection



Keypoints



Stuff





Evaluation Metrics

Average Precision (AP) :

AP	% AP at IoU=.50::.05::.95 (determines challenge winner)
AP _{IoU=.50}	% AP at IoU=.50 (PASCAL VOC metric)
AP _{IoU=.75}	% AP at IoU=.75 (strict metric)

AP Across Scales:

AP _{small}	% AP for small objects: area < 32 ²
AP _{medium}	% AP for medium objects: 32 ² < area < 96 ²
AP _{large}	% AP for large objects: area > 96 ²

Average Recall (AR) :

AR _{max=1}	% AR given 1 detection per image
AR _{max=10}	% AR given 10 detections per image
AR _{max=100}	% AR given 100 detections per image

AR Across Scales:

AR _{small}	% AR for small objects: area < 32 ²
AR _{medium}	% AR for medium objects: 32 ² < area < 96 ²
AR _{large}	% AR for large objects: area > 96 ²



Evaluation Metrics

Average Precision (AP) :

AP

AP_{IoU=.50}

AP_{IoU=.75}

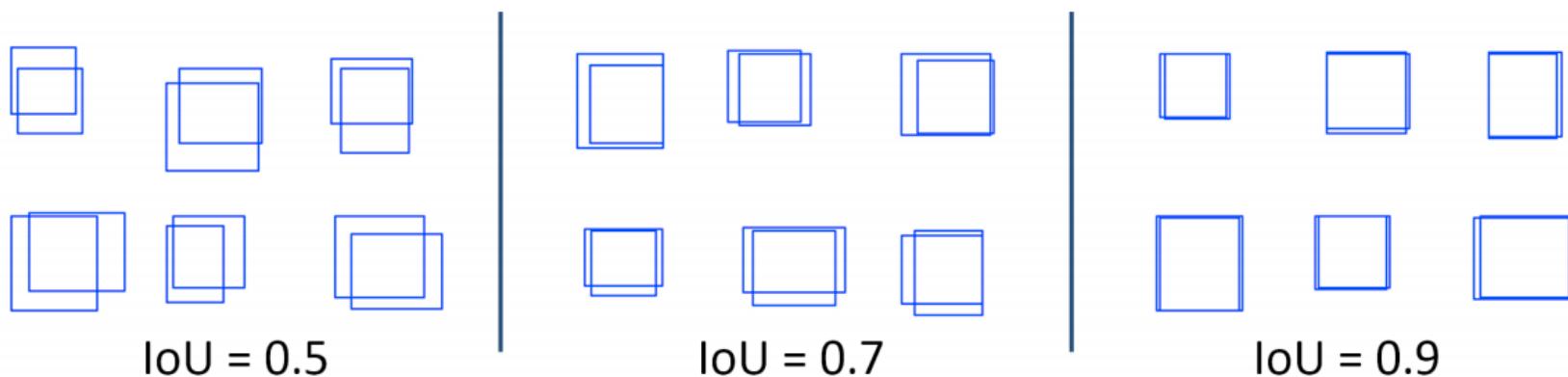
% AP at IoU=.50:.05:.95 (**determines challenge winner**)

% AP at IoU=.50 (PASCAL VOC metric)

% AP at IoU=.75 (strict metric)

Challenges Score: AP

- AP is averaged over multiple IoU values between 0.5 and 0.95.
- More comprehensive metric than the traditional AP at a fixed IoU value (0.5 for PASCAL).





Evaluation Metrics

AP Across Scales:

AP_{small}

% AP for small objects: area < 32^2

AP_{medium}

% AP for medium objects: $32^2 < \text{area} < 96^2$

AP_{large}

% AP for large objects: area > 96^2

Other Scores: Size AP

- AP is averaged over instance size:
 - small ($A < 32 \times 32$)
 - medium ($32 \times 32 < A < 96 \times 96$)
 - large ($A > 96 \times 96$)

$A < 32 \times 32$



$A > 96 \times 96$



$32 \times 32 < A < 96 \times 96$





Evaluation Metrics

Average Recall (AR) :

$AR_{max=1}$

% AR given 1 detection per image

$AR_{max=10}$

% AR given 10 detections per image

$AR_{max=100}$

% AR given 100 detections per image

AR Across Scales:

AR_{small}

% AR for small objects: $area < 32^2$

AR_{medium}

% AR for medium objects: $32^2 < area < 96^2$

AR_{large}

% AR for large objects: $area > 96^2$

Other Scores: AR

- Measures the maximum recall over a fixed number of detections allowed in the image of 1, 10, 100.
- AR is averaged over small ($A < 32 \times 32$), medium ($32 \times 32 < A < 96 \times 96$) and large ($A > 96 \times 96$) instances of objects.



coco Challenges Results



coco

Common Objects in Context



Detection



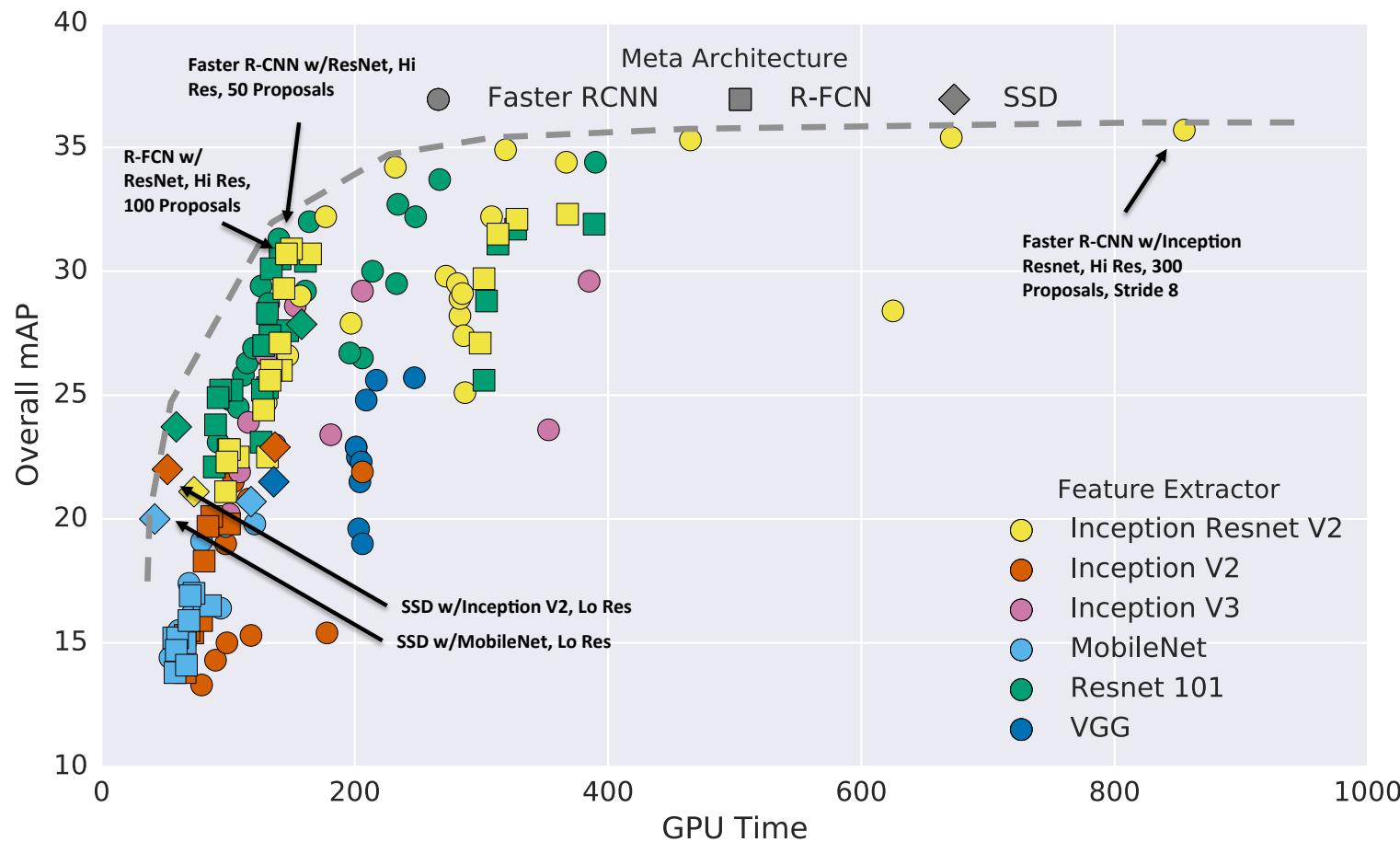
Segmentation





Shout-out to previous algorithms!

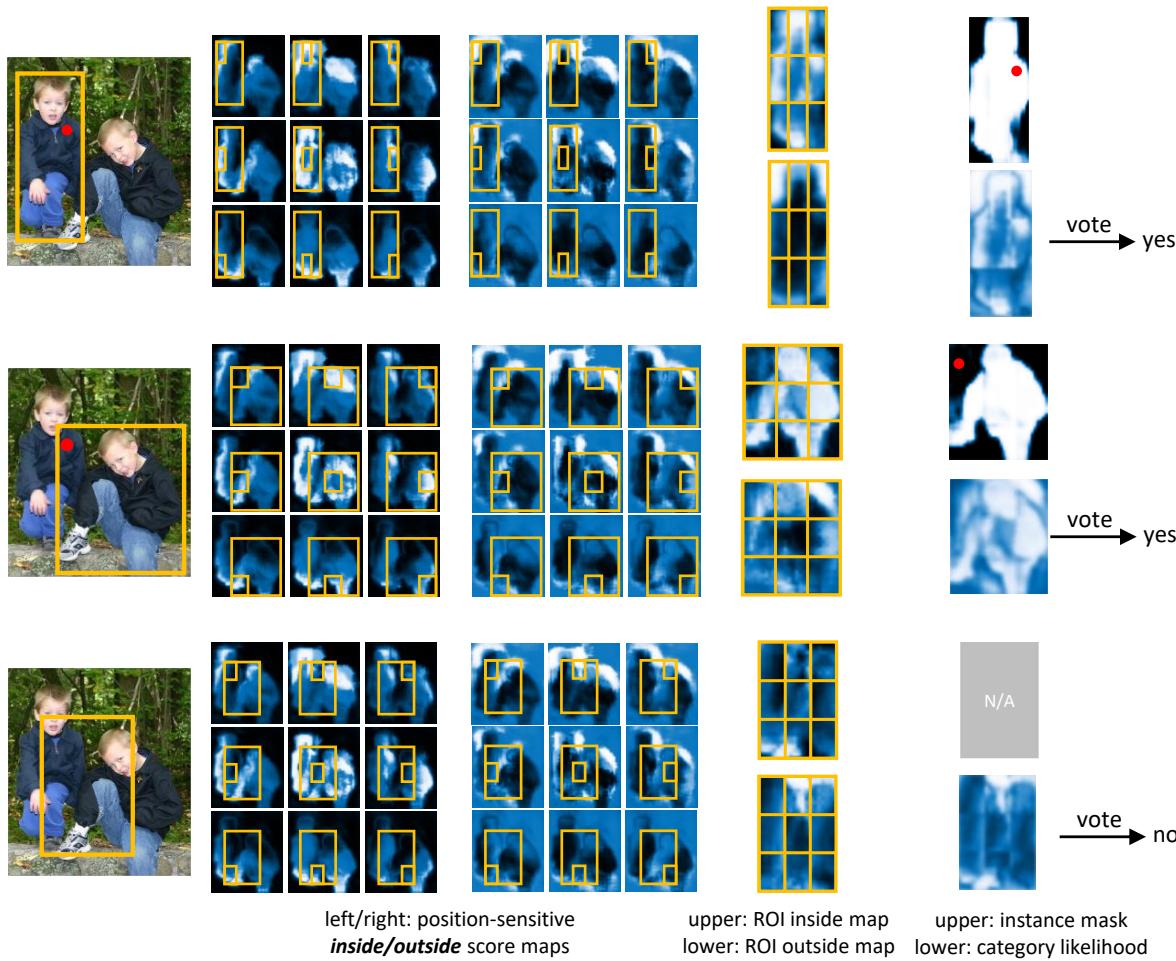
Google Research





Shout-out to previous algorithms!

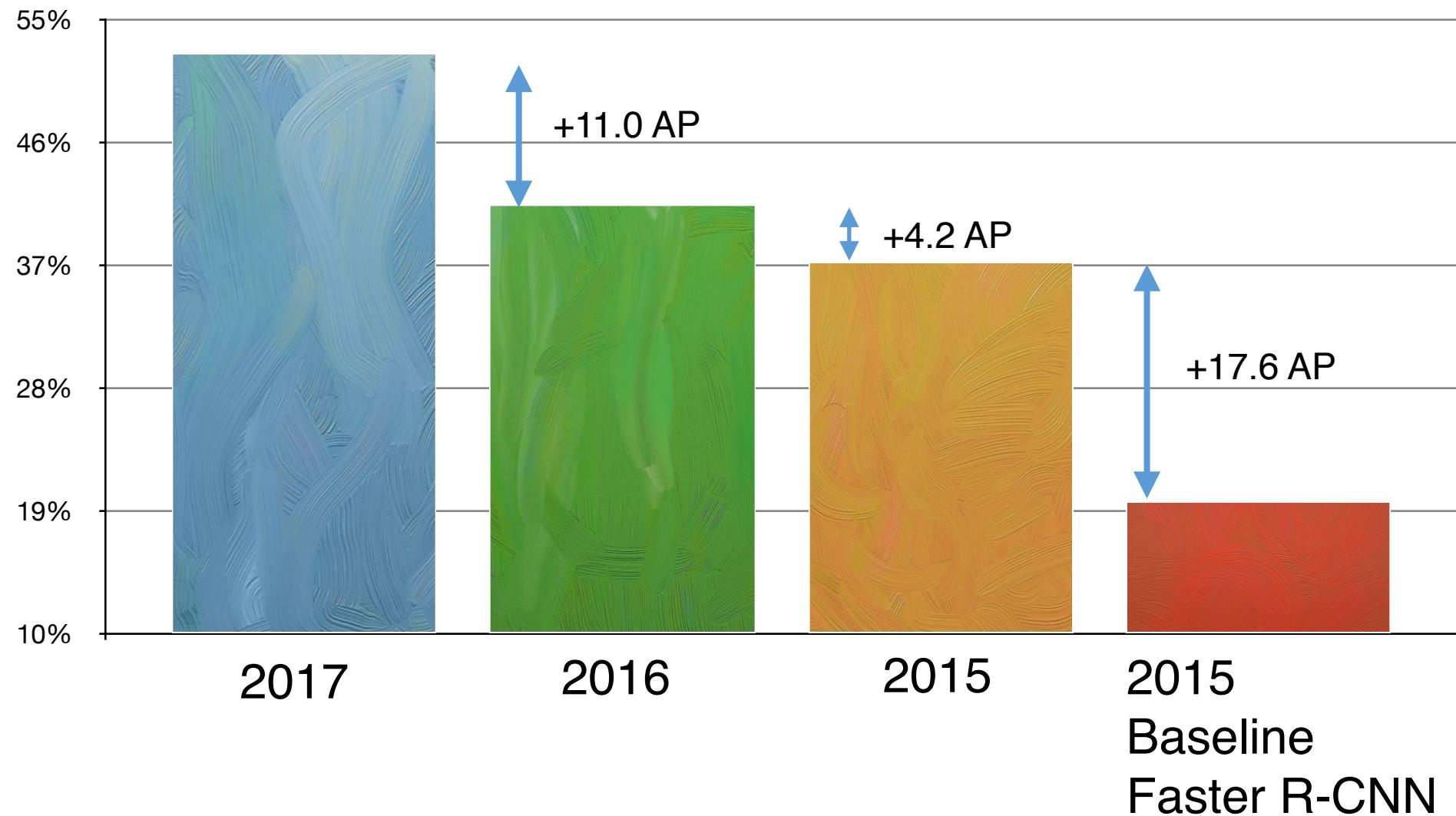
Microsoft Research Asia





Bounding Boxes Leaderboard (I)

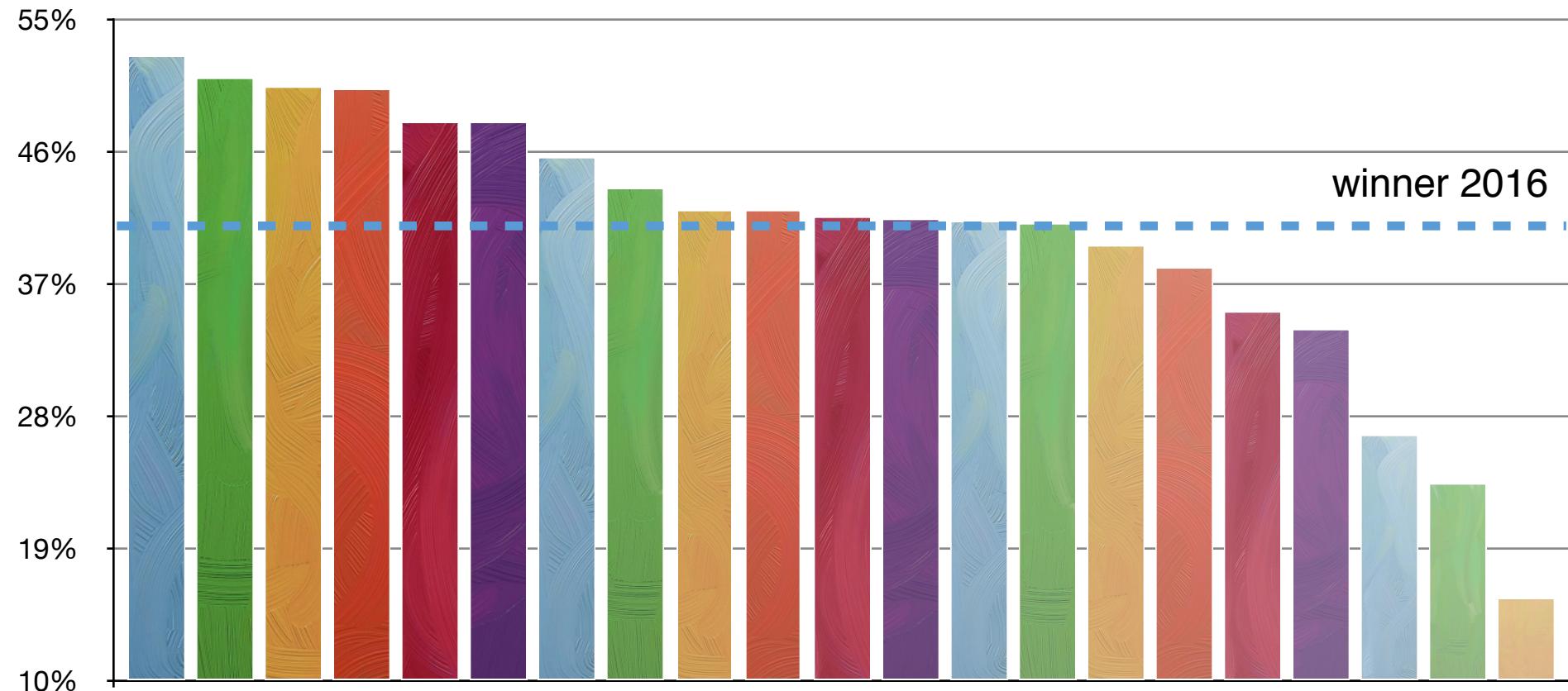
COCO AP (over all IoU)





Bounding Boxes Leaderboard (II)

COCO AP (over all IoU)



21 teams joined the competition

12 teams achieved better performance than last year's winner

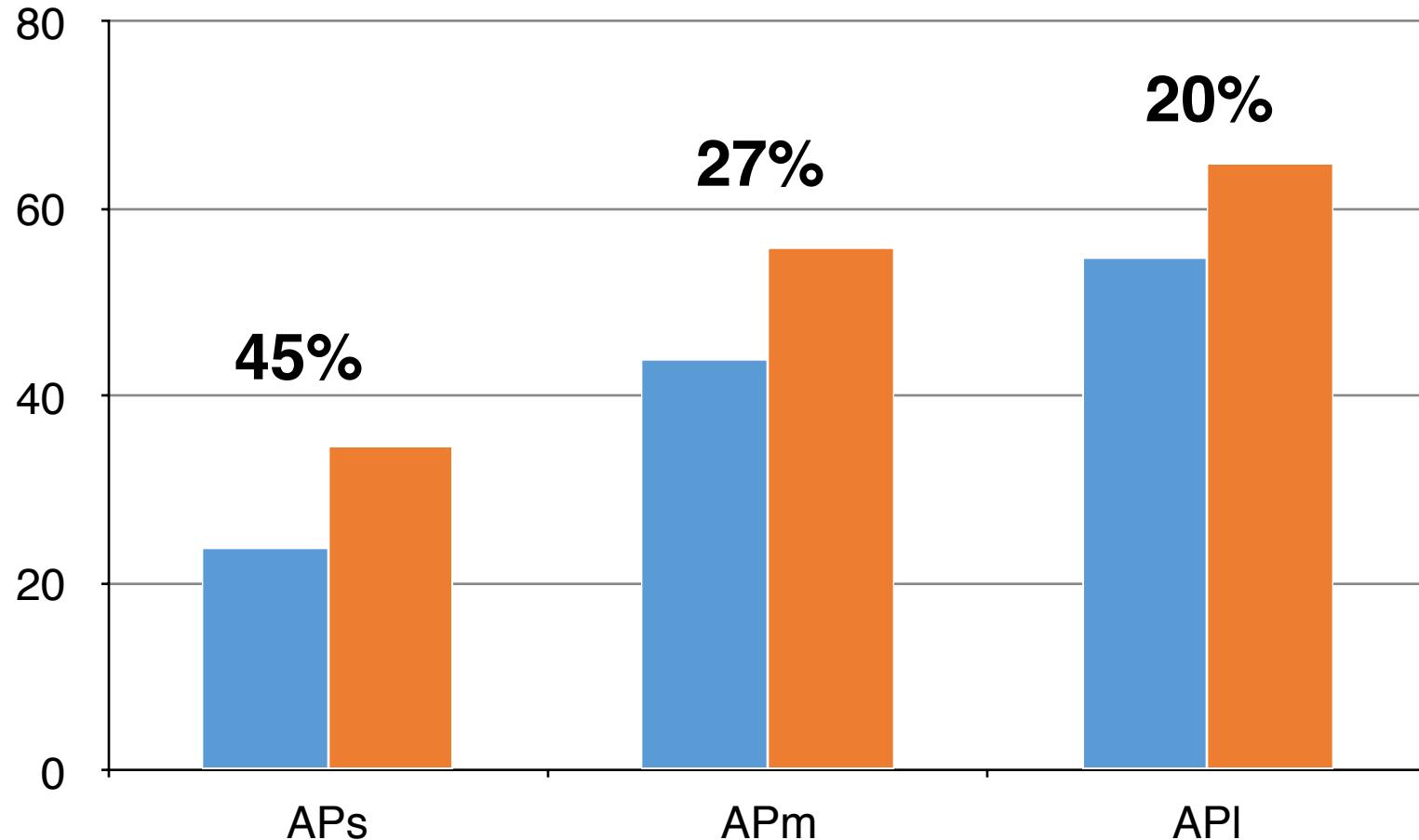
4 teams > 50 AP



Bounding Boxes Leaderboard (III)

Winner 2016

Winner 2017

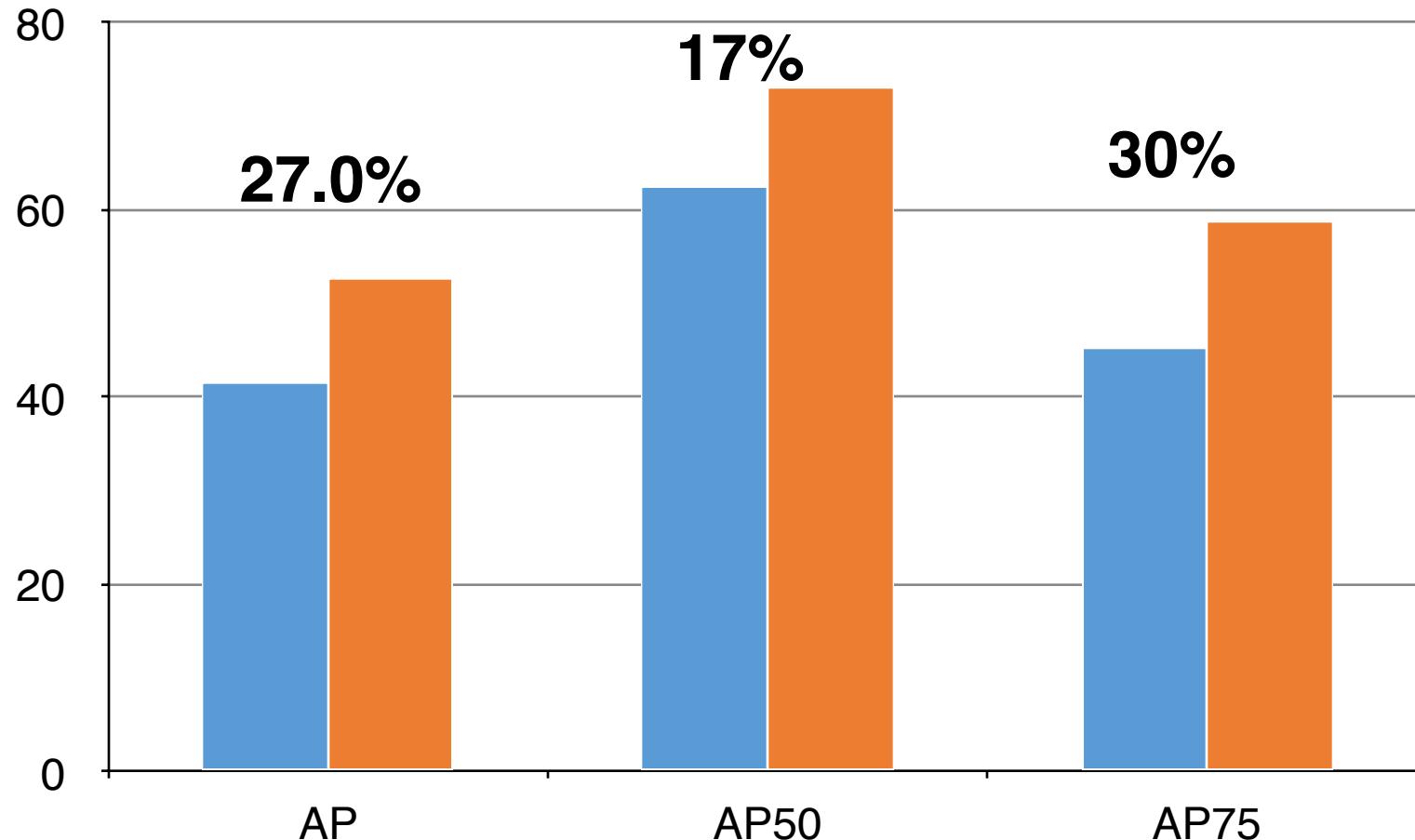




Bounding Boxes Leaderboard (III)

Winner 2016

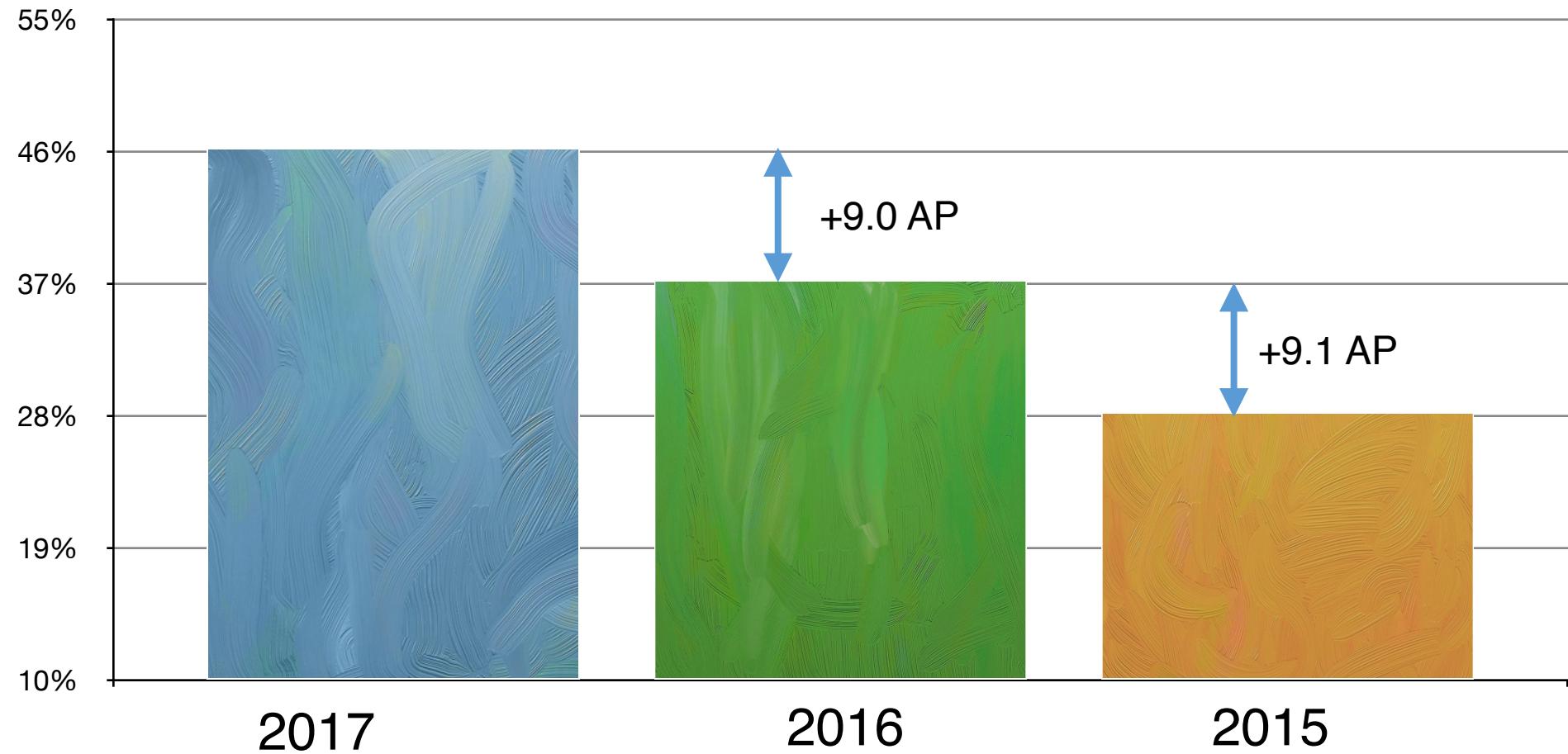
Winner 2017





Segmentation Leaderboard (I)

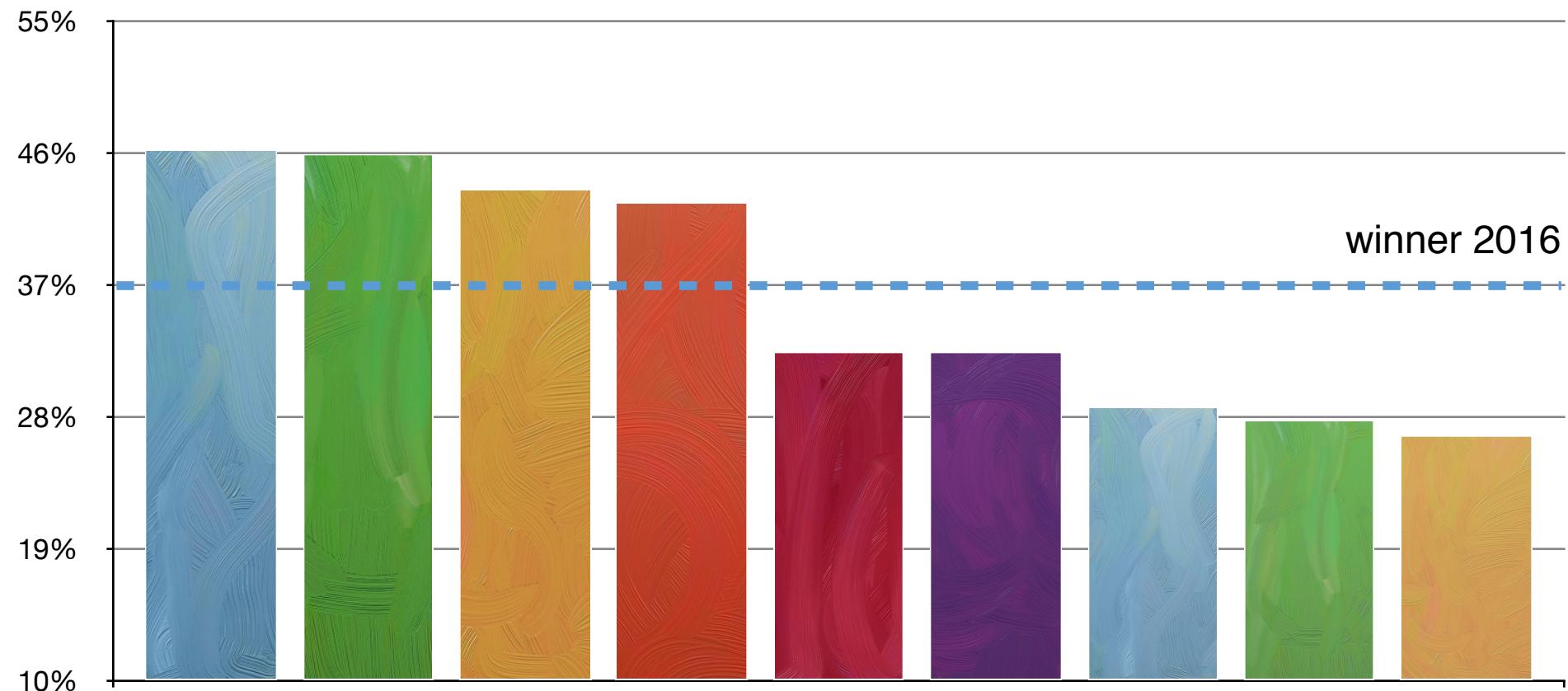
COCO AP (over all IoU)





Segmentation Leaderboard (II)

COCO AP (over all IoU)



9 teams joined the competition

4 teams achieved better performance than last year's winner

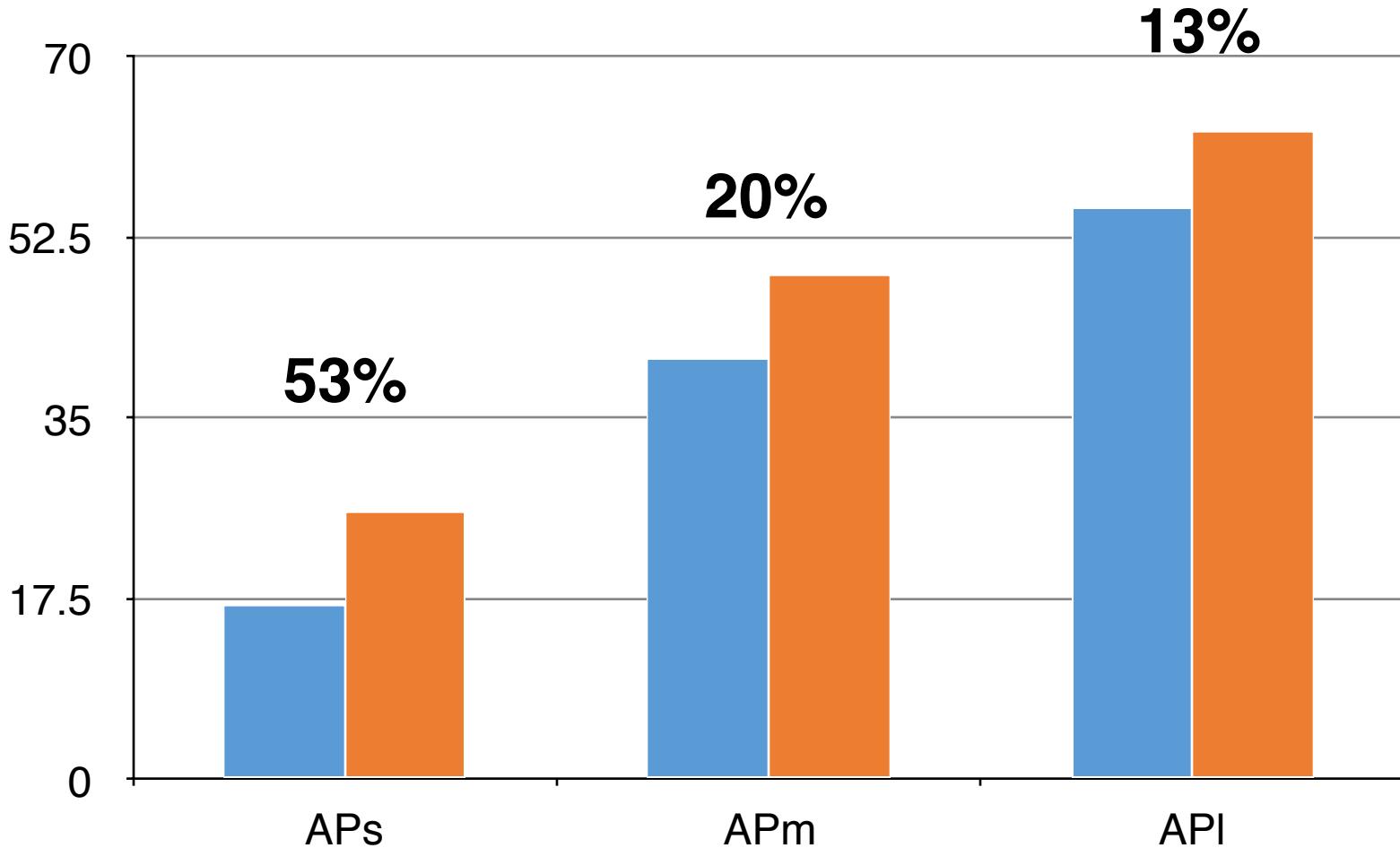
4 teams > 40 AP



Segmentation Leaderboard (III)

Winner 2016

Winner 2017

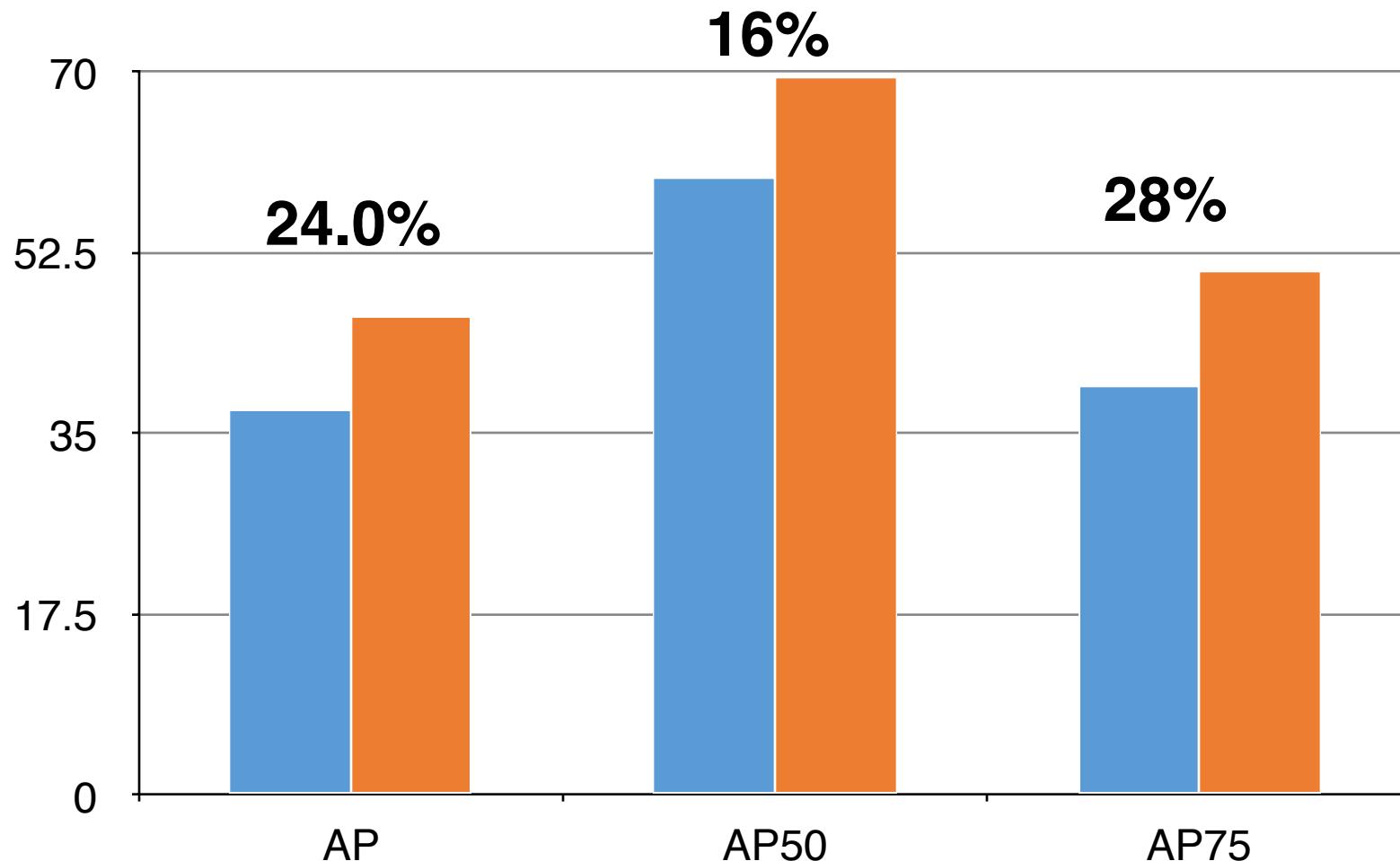




Segmentation Leaderboard (III)

Winner 2016

Winner 2017





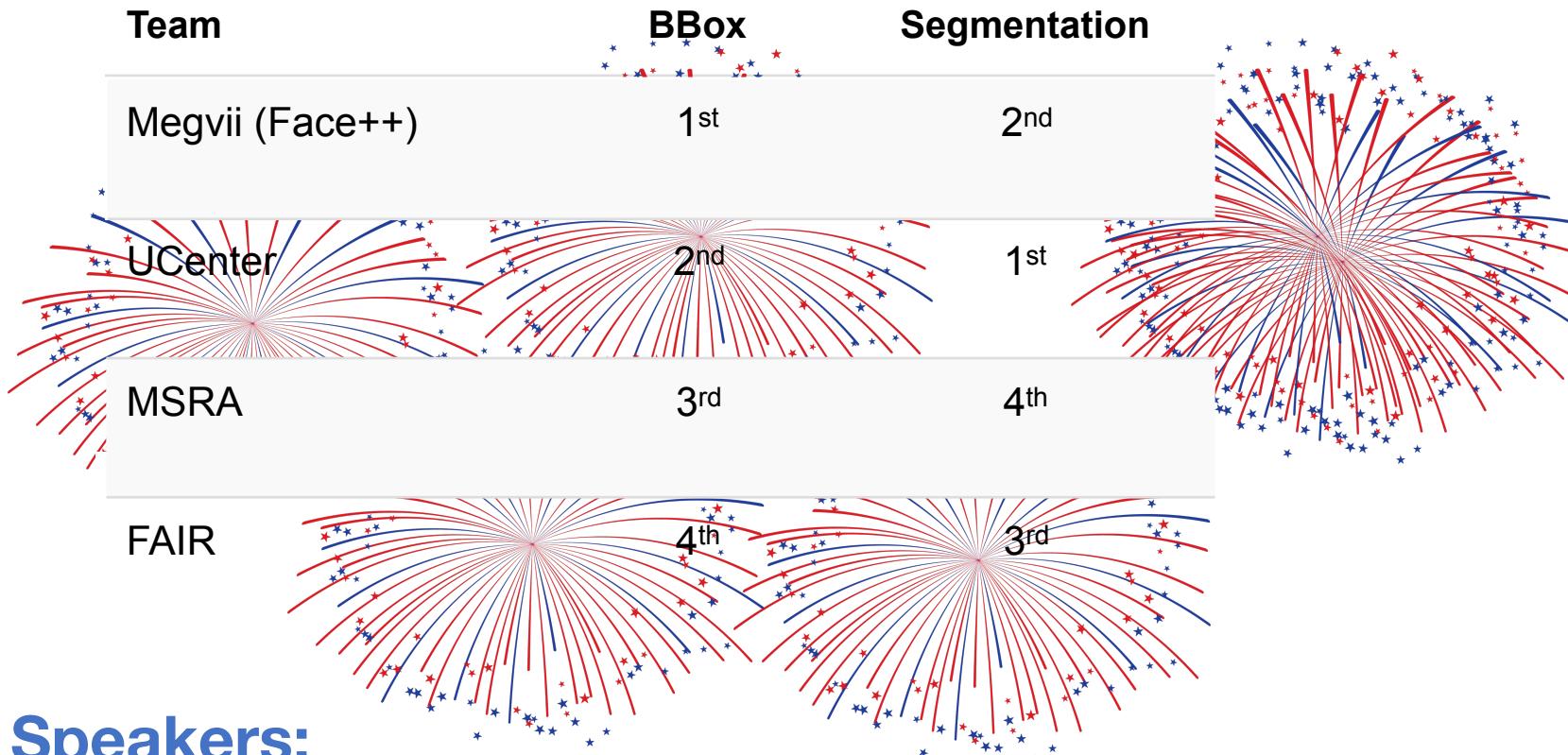
Summary of Findings

2017 Detection Challenge Take-aways

- Exciting breakthrough this year!
- +11 AP for box challenge
- +9 AP for segmentation challenge
- Localization improved greatly in both challenges.
- High relative improvement on small object instances.



Challenges Ranking



Invited Speakers:

- Megvii / 910 - 930a
- UCenter / 930 - 950a
- MSRA / 950 - 1000a
- FAIR / 1000 - 1010a