

Matteo Ruggero Ronchi



## **COCO and Places Visual Recognition Challenges Workshop**

**Sunday, October 29th, Venice, Italy**



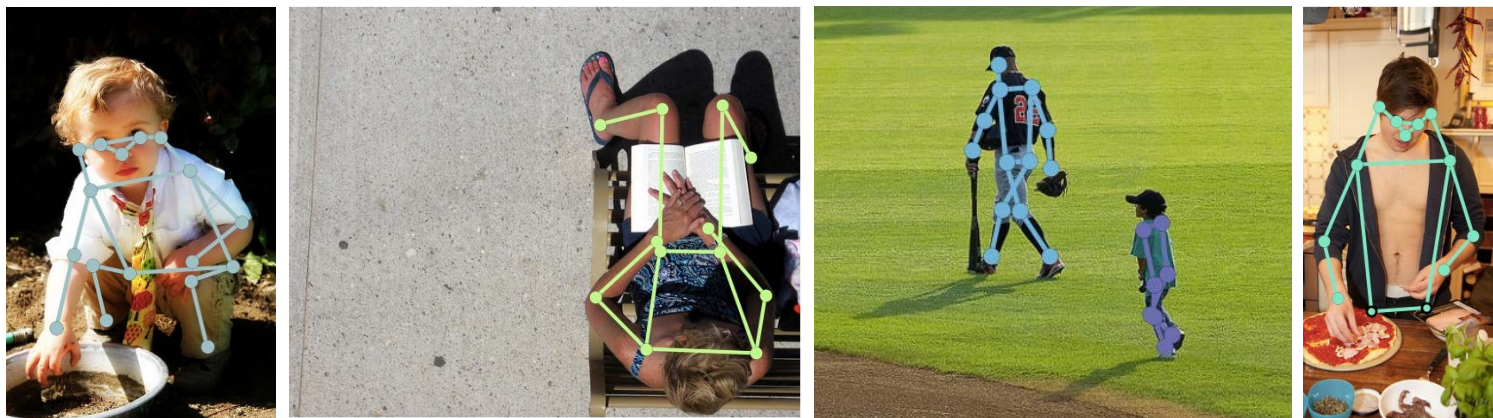
# COCO Keypoints Dataset





# COCO Keypoints Dataset (I)

## Multiple Perspectives, Instances, Sizes, Occlusions:



## Overall Statistics (train/val):

- 17 types of keypoints.
- 58,945 images.
- 156,165 annotated people.
- 1,710,498 total keypoints.



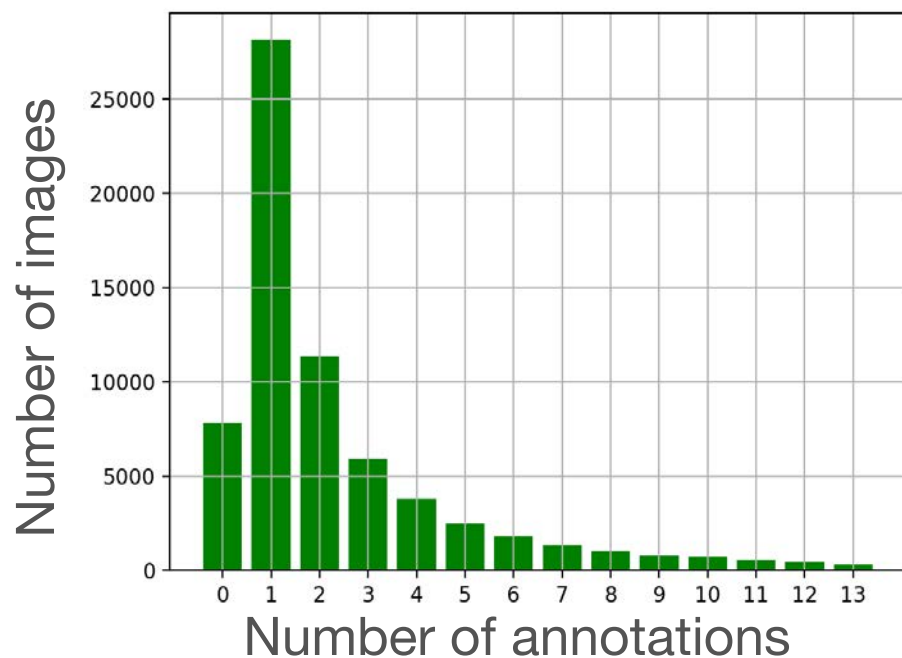




# COCO Keypoints Dataset (II)

## Multi-Instance Dataset:

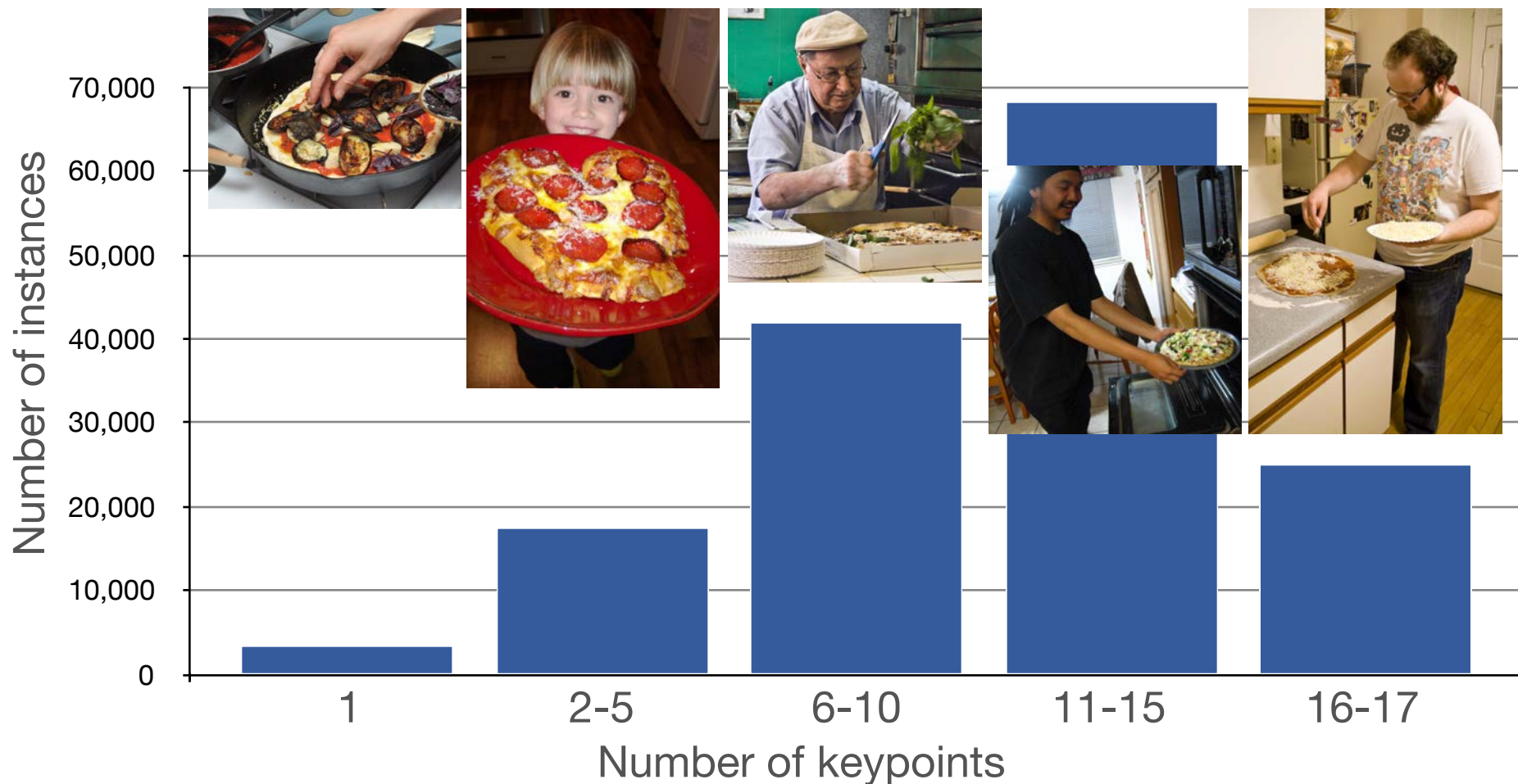
- Avg of ~2 annotated people per image.
- Up to 13 annotated people per image.





# COCO Keypoints Dataset (III)

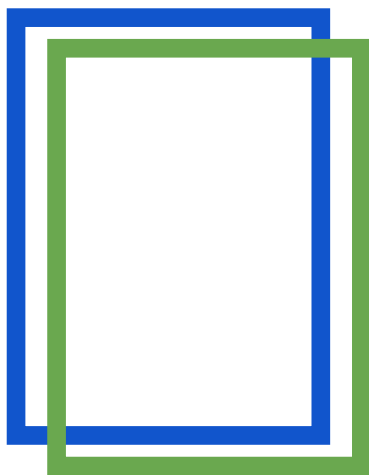
## Distribution of the number of keypoints:



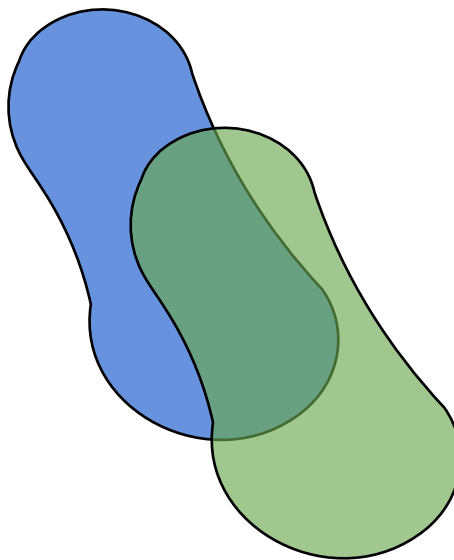


# Evaluating Keypoint Predictions

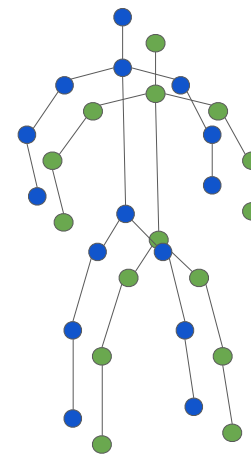
How to measure localization accuracy:



Bounding Box IoU



Mask IoU



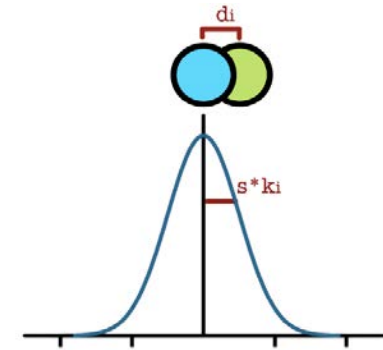
Object  
Keypoint  
Similarity



# Keypoints Evaluation Metric

## Object Keypoint Similarity (OKS):

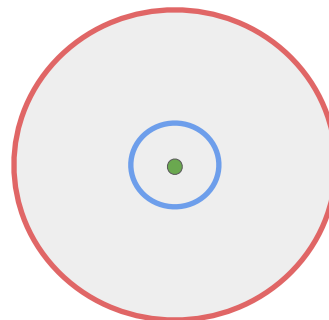
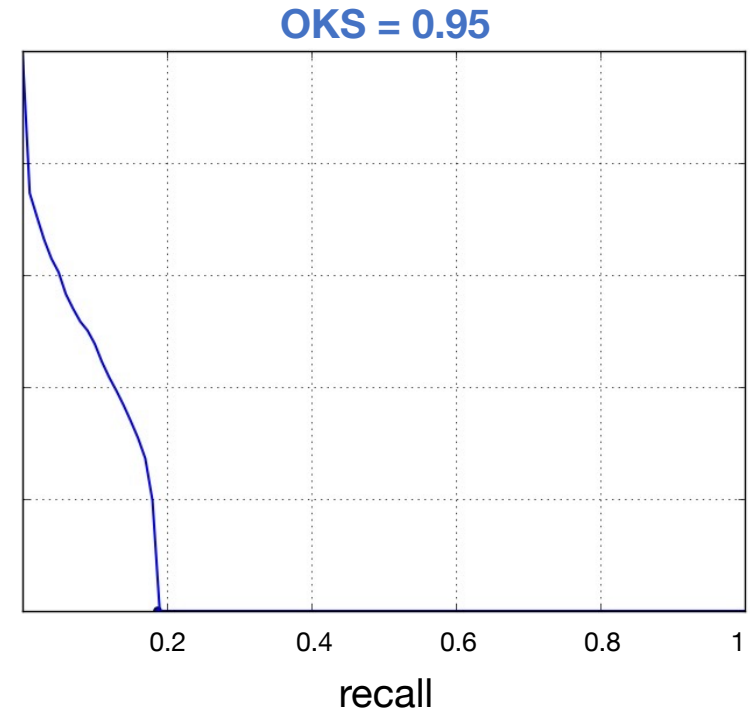
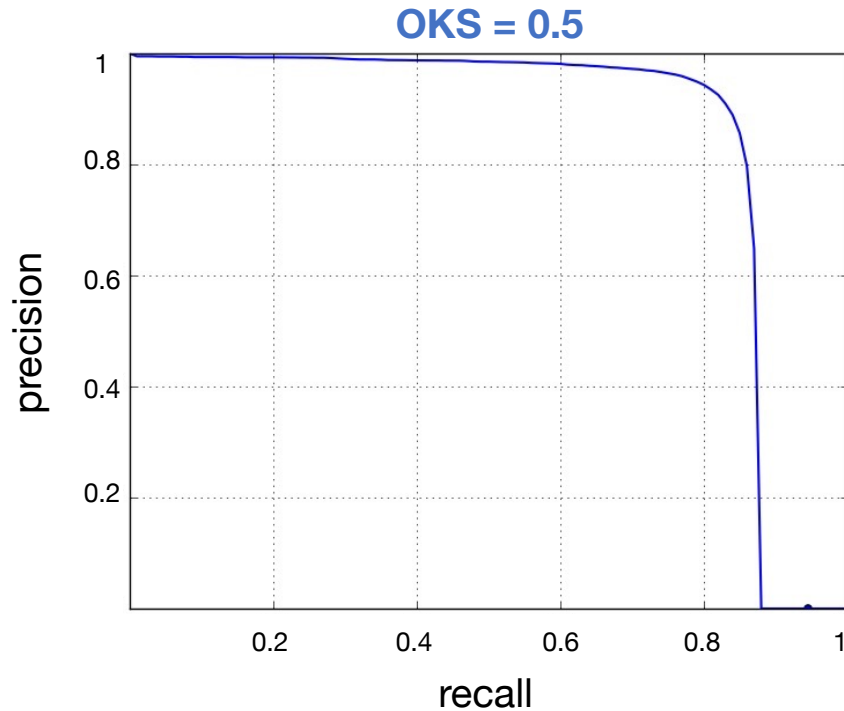
$$\begin{cases} ks(\hat{\theta}_i^{(p)}, \theta_i^{(p)}) &= e^{-\frac{\|\hat{\theta}_i^{(p)} - \theta_i^{(p)}\|_2^2}{2s^2 k_i^2}} \\ OKS(\hat{\theta}^{(p)}, \theta^{(p)}) &= \frac{\sum_i ks(\hat{\theta}_i^{(p)}, \theta_i^{(p)}) \delta(v_i > 0)}{\sum_i \delta(v_i > 0)} \end{cases}$$





# COCO Keypoints Task

## Simultaneous detection and keypoint estimation:



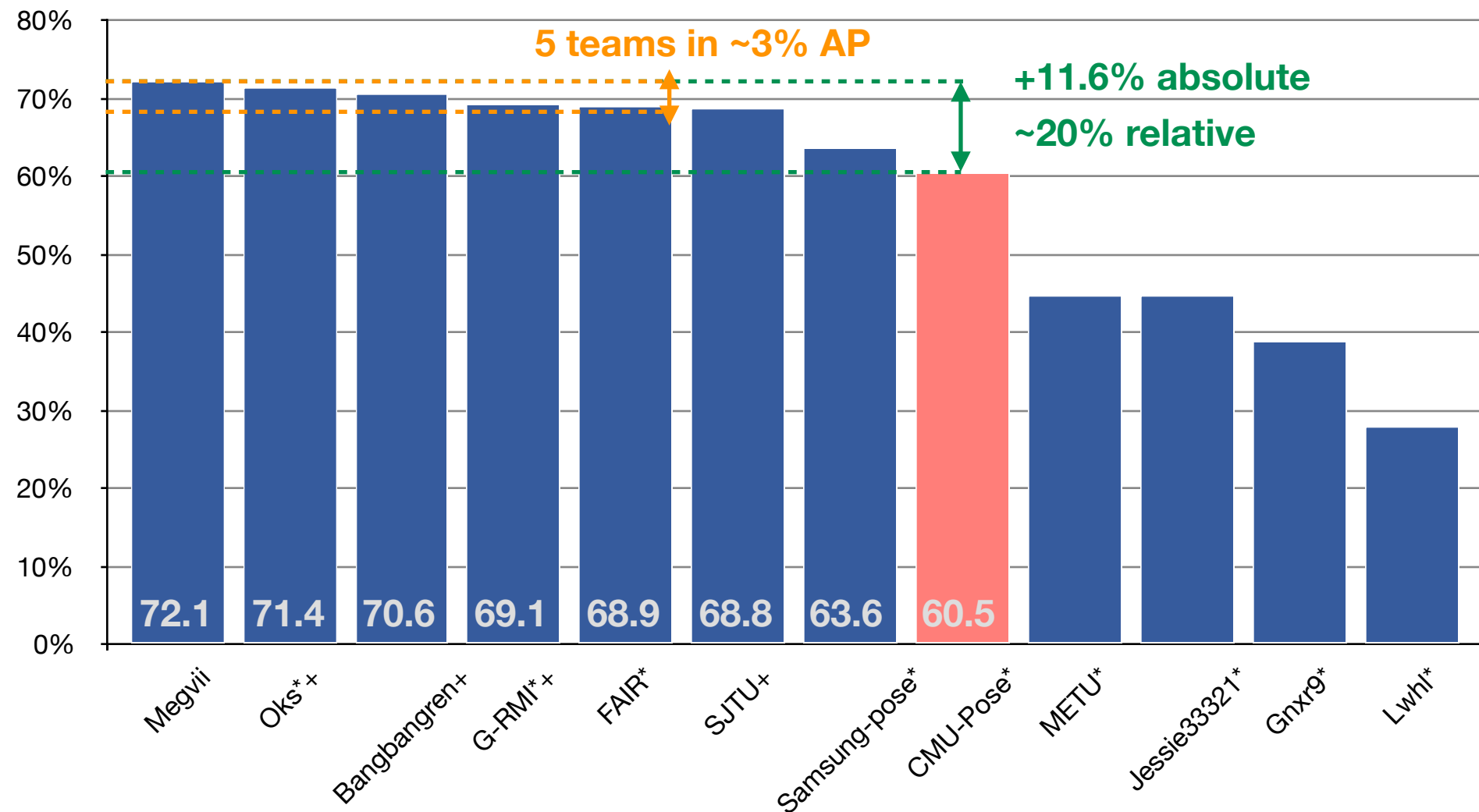
- OKS .5
- OKS .95
- Ground Truth





# 2017 Keypoints Challenge Leaderboard (I)

## COCO AP (average over all OKS)



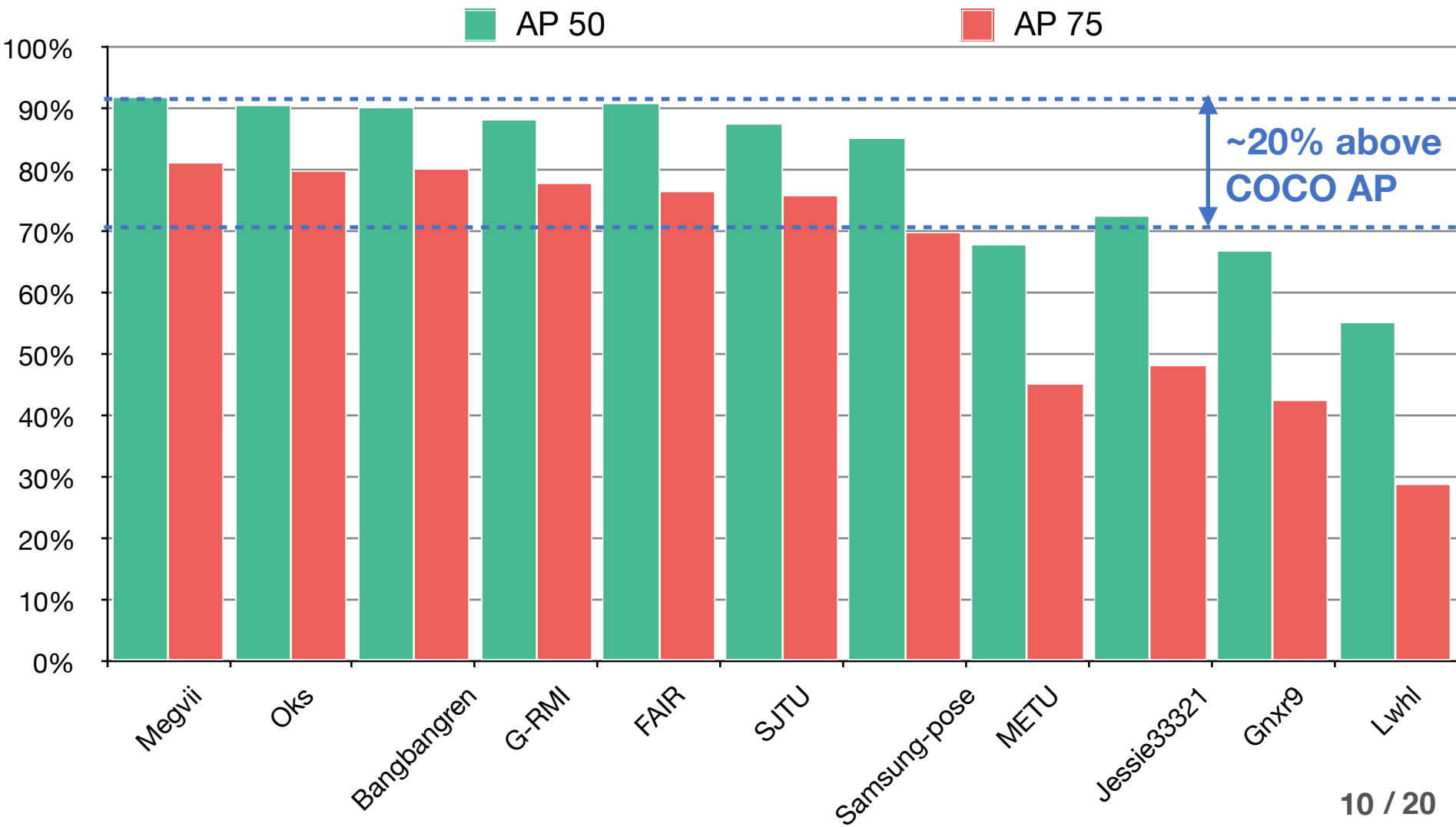
\* Single model method

+ Used external keypoints training dataset



# 2017 Keypoints Challenge Leaderboard (II)

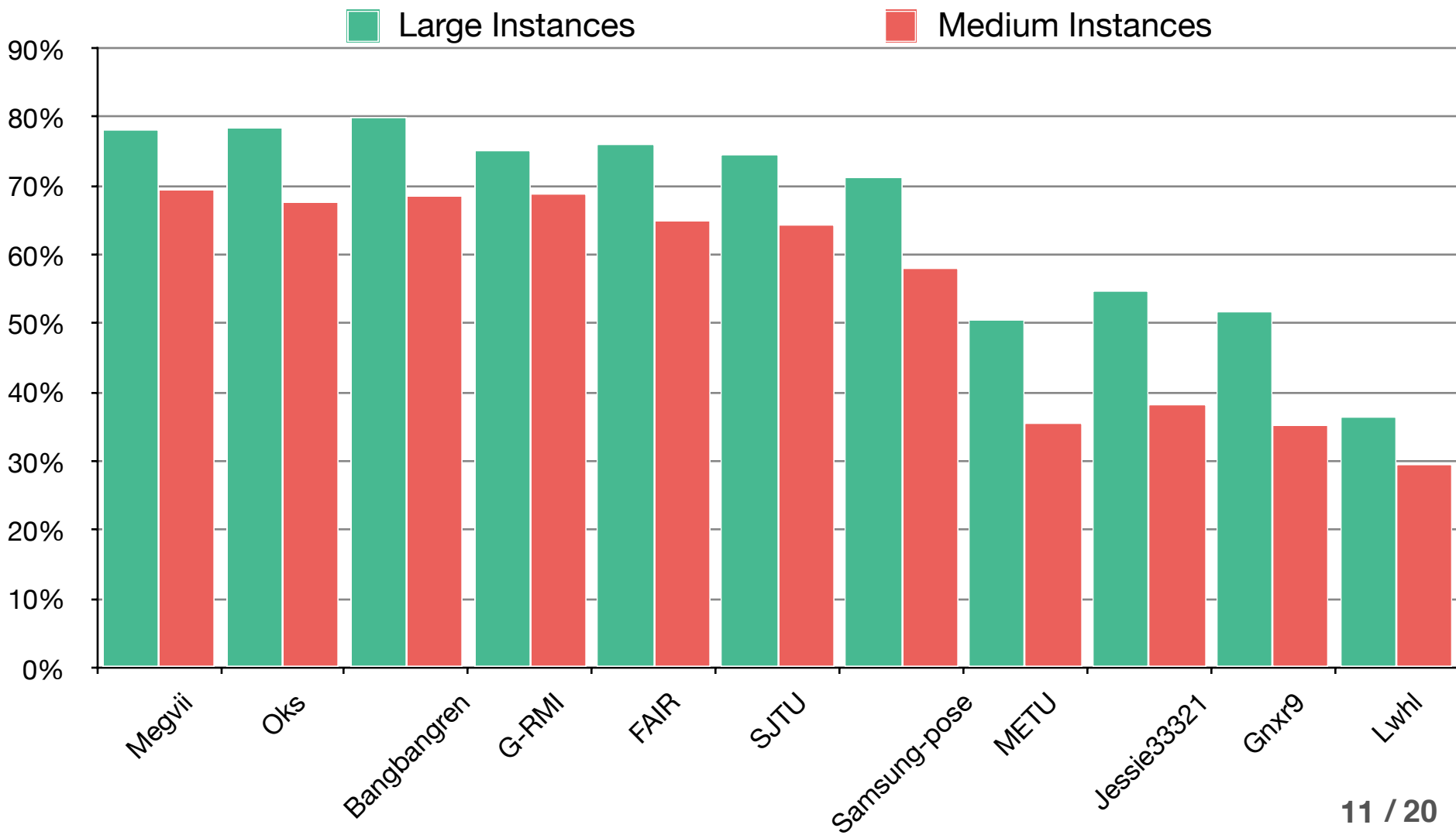
Better performance at looser localization thresholds:





# 2017 Keypoints Challenge Leaderboard (III)

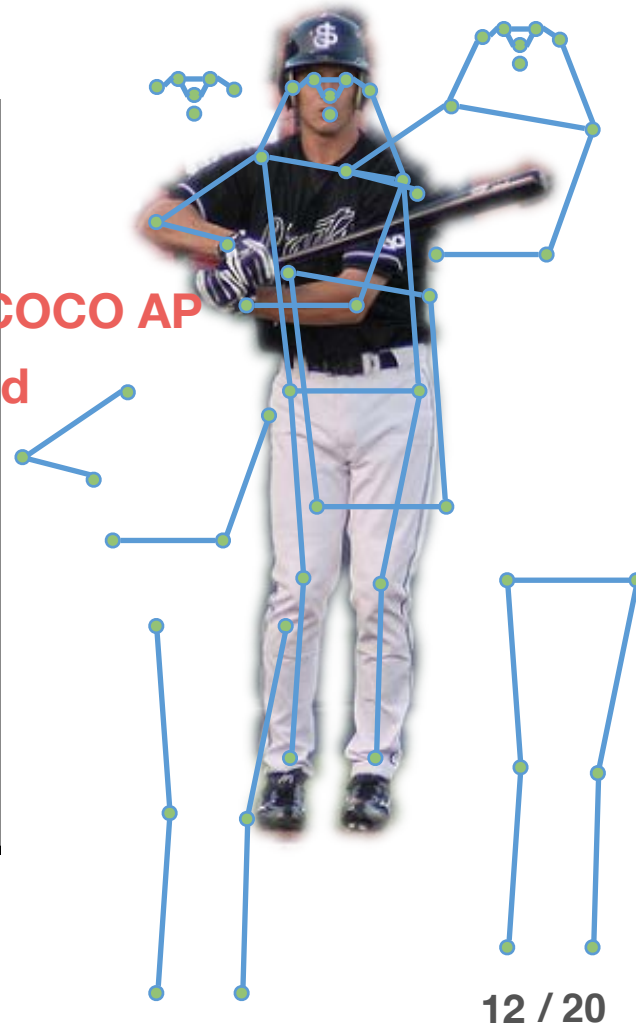
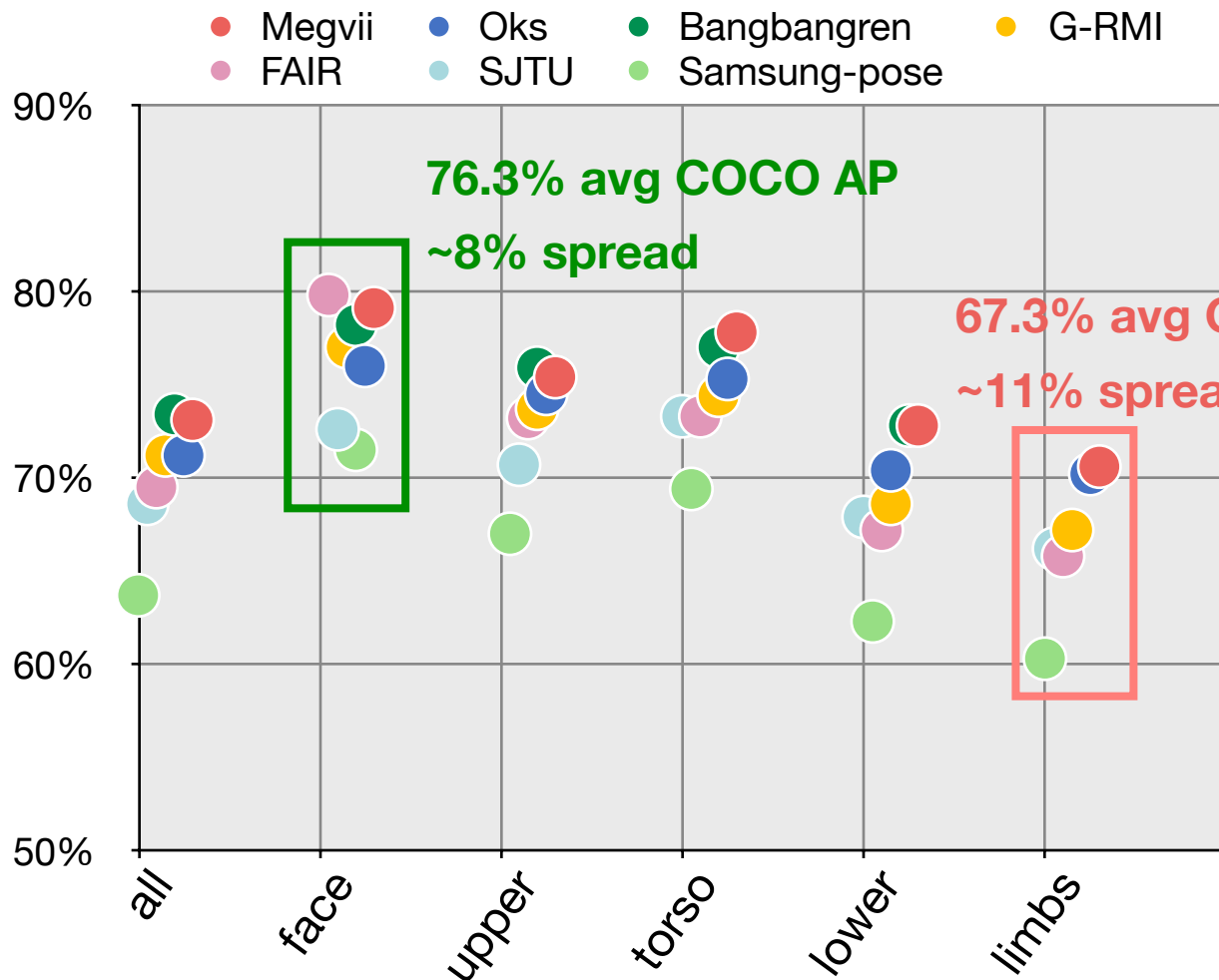
Instance scale is an important factor:





# Performance Breakdown over Keypoints

## COCO AP varies across keypoints







# A Closer Look at Errors

matteorr / coco-analyze

Unwatch 3 Star 16 Fork 6

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

A wrapper of the COCOeval class for extended keypoint error estimation analysis. [Add topics](#) [Edit](#)

31 commits 1 branch 0 releases 2 contributors MIT

Branch: release New pull request Create new file Upload files Find file Clone or download

matteorr Delete masknn\_performance\_report.pdf Latest commit e7deb0e 24 days ago

analysisAPI	updated repo to work with new coco_url	a month ago
annotations	updated repo to work with new coco_url	a month ago
detections	added fakekeypoints100 annotations	4 months ago
latex	Finalized version of coco-analyze repository.	4 months ago
pycocotools	Finalized version of coco-analyze repository.	4 months ago
.gitignore	added fakekeypoints100 annotations	4 months ago
COCOanalyze_demo.ipynb	updated repo to work with new coco_url	a month ago
LICENSE	Initial commit	8 months ago
README.md	updated README	24 days ago
run_analysis.py	updated repo to work with new coco_url	a month ago





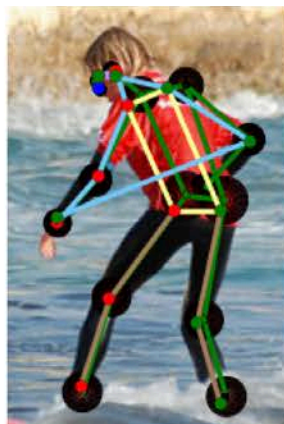
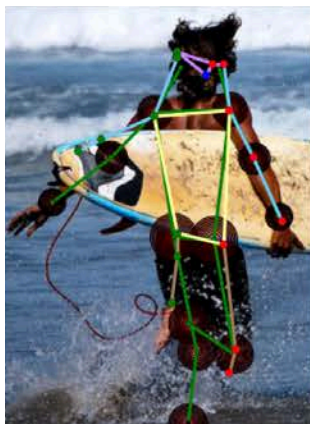
# A Closer Look at Errors (I)

## Taxonomy of Errors for Multi-Instance Pose Estimation:

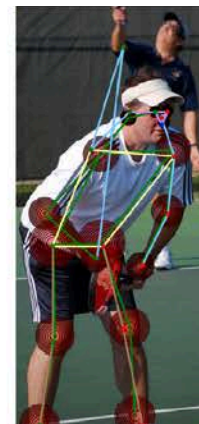
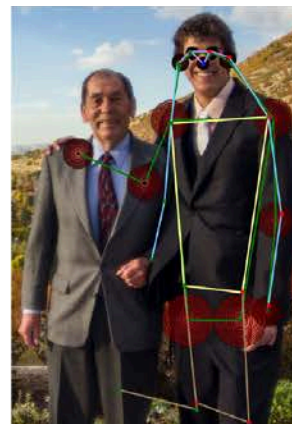
**JITTER**



**INVERSION**



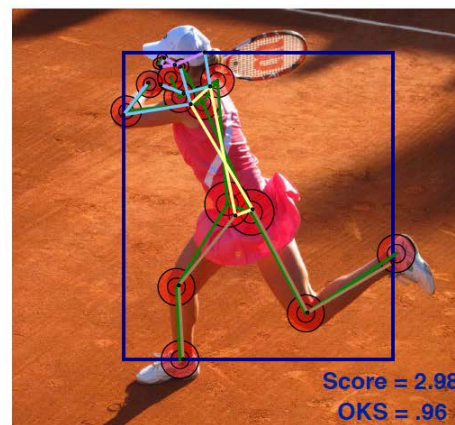
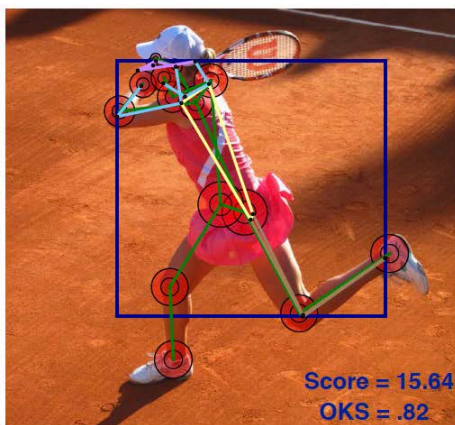
**SWAP**



**MISS**



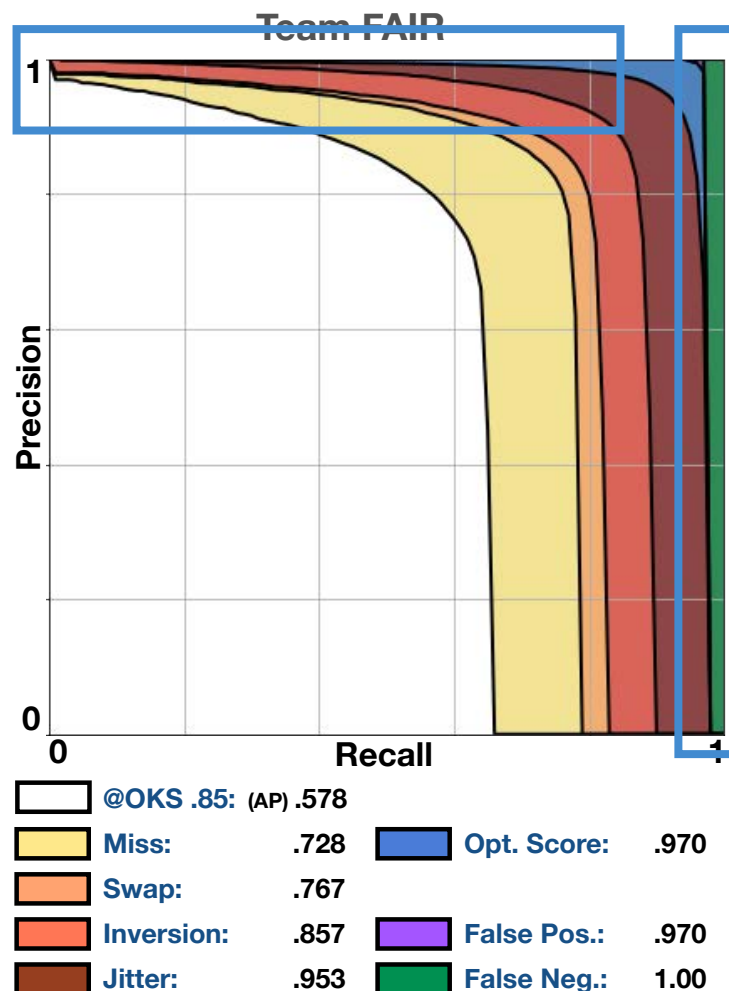
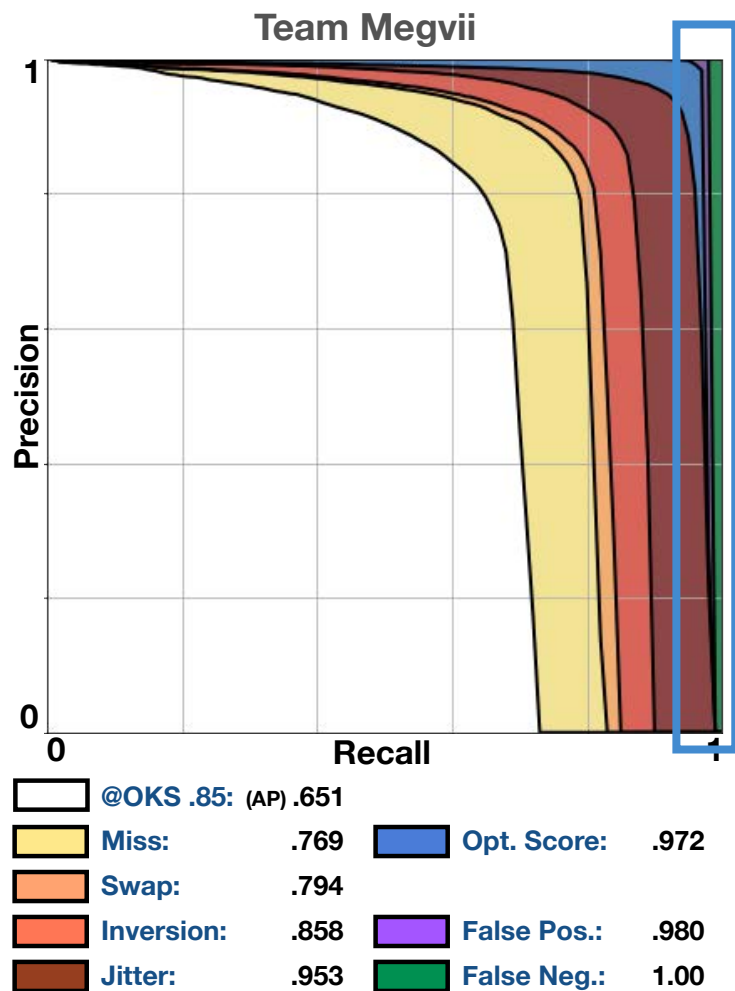
**SCORING**





# A Closer Look at Errors (II)

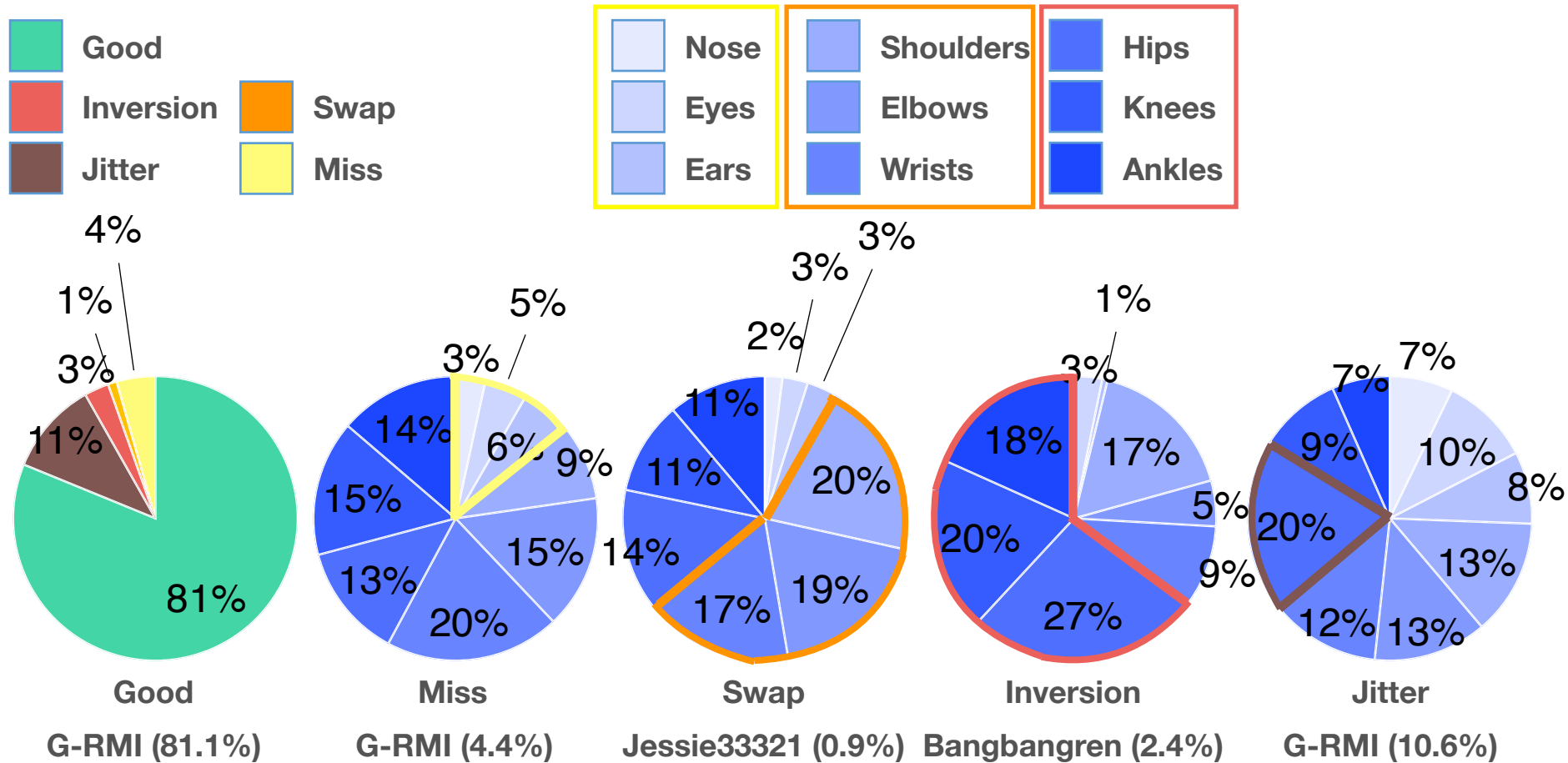
## Fine-grained Precision-Recall Curves





# Localization Errors

## Best performance for each type of Localization Error



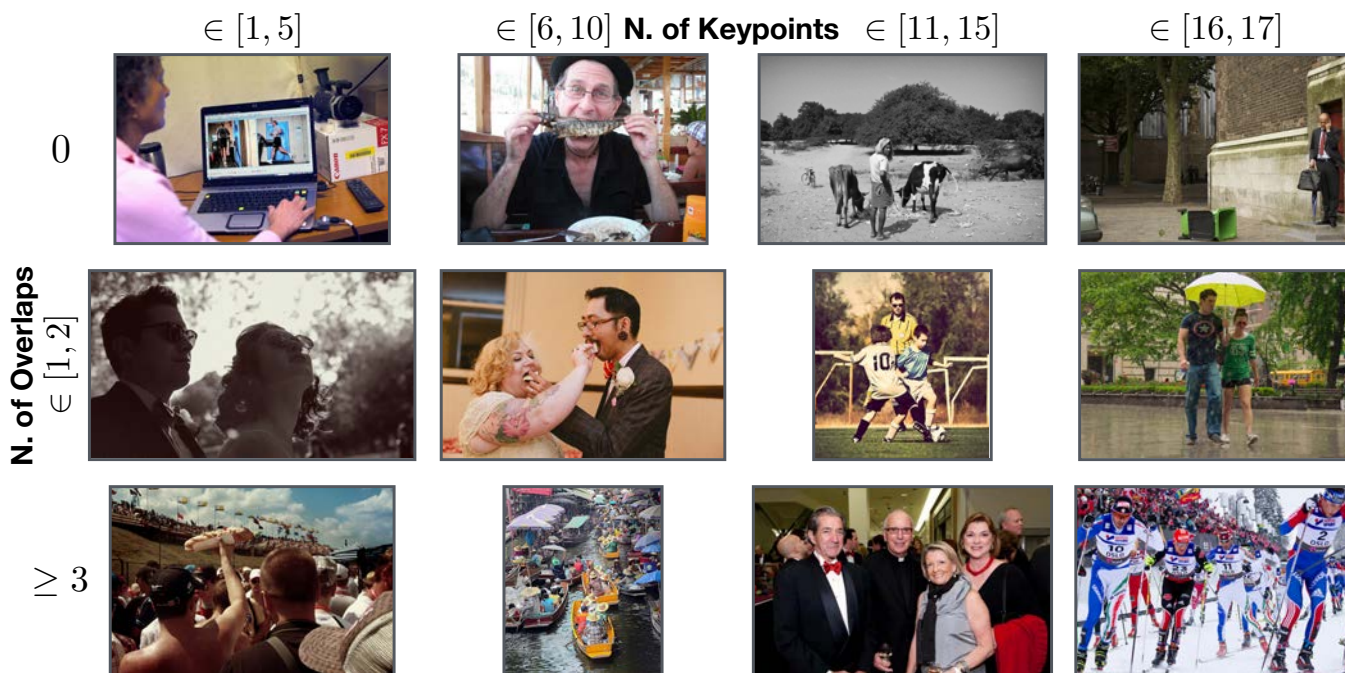




# Occlusion and Crowding Benchmarks (I)

## COCO Benchmarks of image complexity:

- **Occlusion:** number of visible keypoints
- **Crowding:** number of overlapping instances ( $\text{IoU} > 0.1$ )



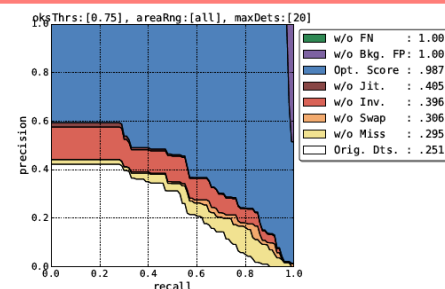
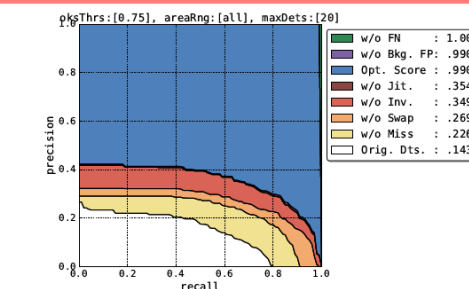
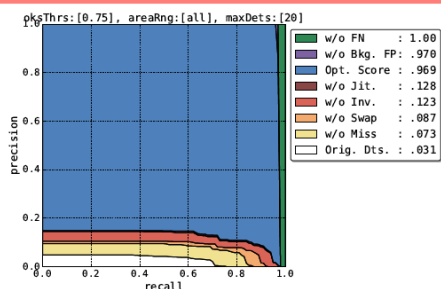
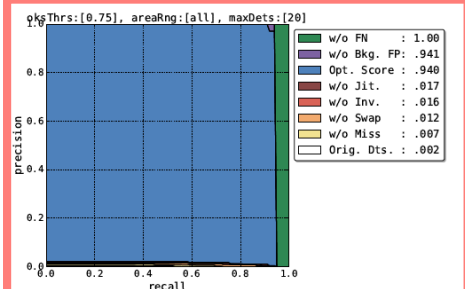
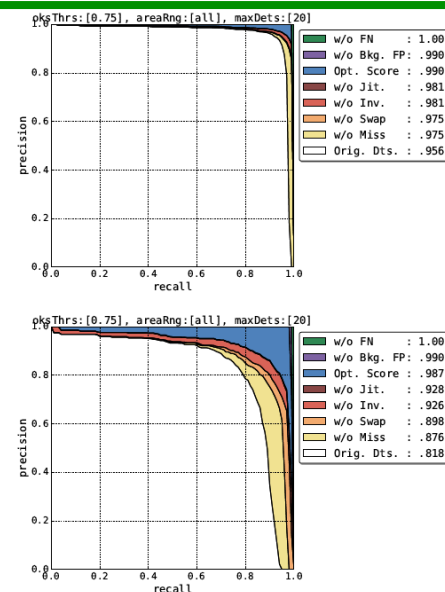
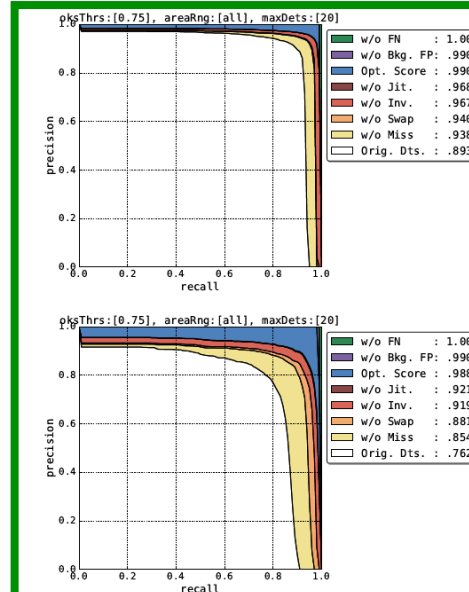
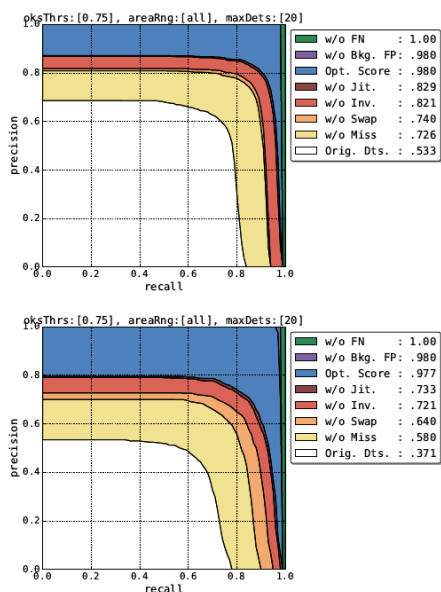
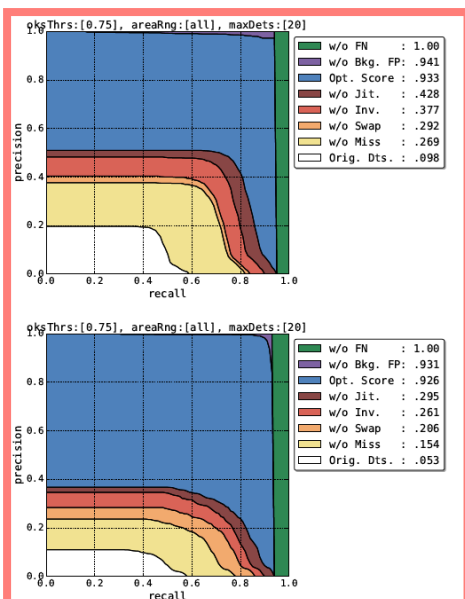
		N. of Keypoints			
		$\in [1, 5]$	$\in [6, 10]$	$\in [11, 15]$	$\in [16, 17]$
N. of Overlaps	0	9098	15968	29165	12246
	$\in [1, 2]$	4876	11735	16384	4636
	$\geq 3$	243	644	780	193



# Occlusion and Crowding Benchmarks (II)

Overall challenge performance is saturated by the easiest benchmarks:

Team Megvii





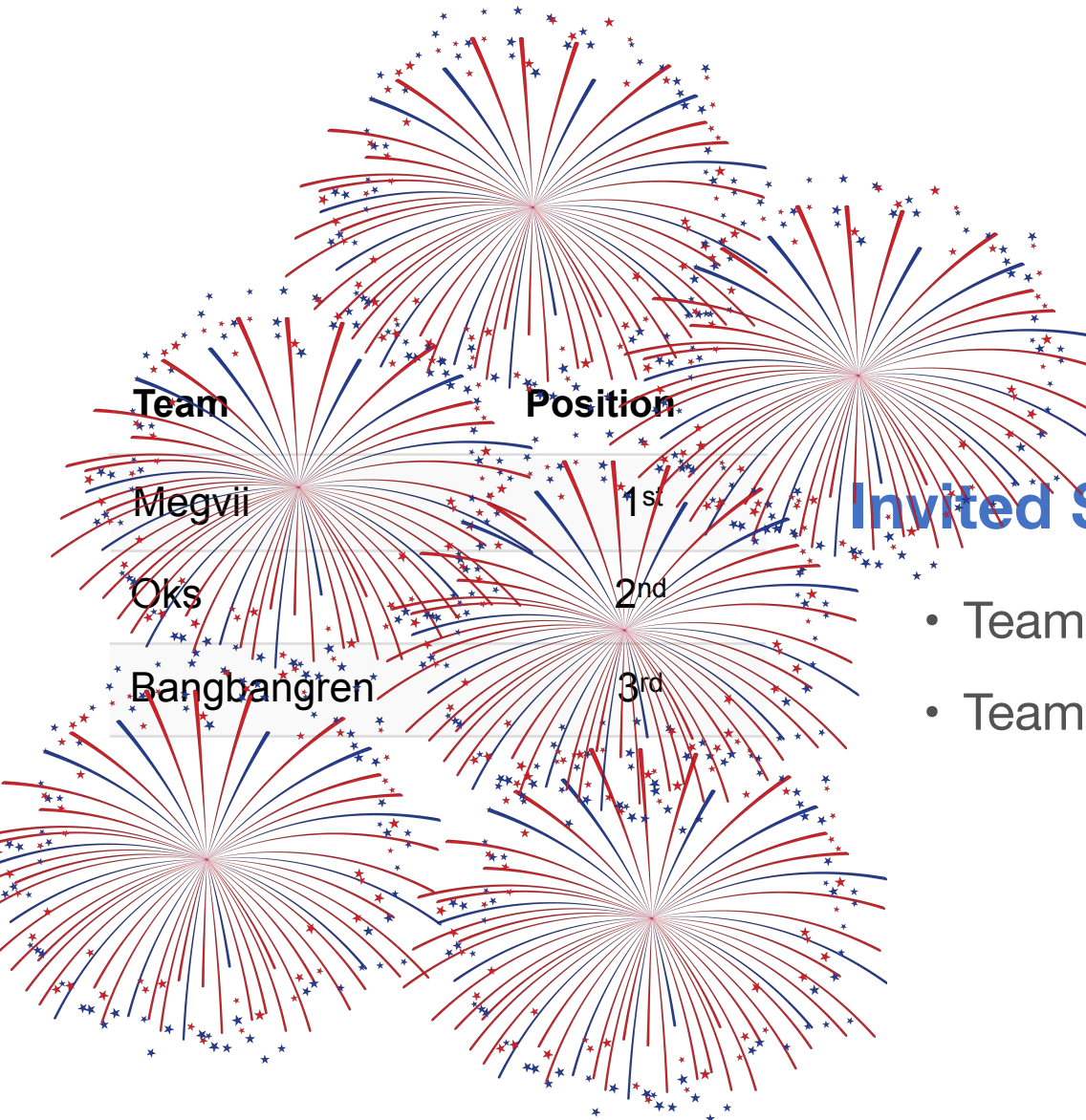
# Summary of Findings

## 2017 Keypoint Challenge Take-aways:

- About 20% relative AP improvement over last year's challenge.
- Very small performance gap between top entries.
- Single model performance is on par with ensembles.
- Single performance metrics do not capture the complex causes of diverse errors.
- We need to broaden current benchmarks with challenging images (high occlusion / low number of keypoints).



# 2017 COCO Keypoints Challenge



**Team**

**Position**

Megvii

1st

Oks

2nd

Bangbangren

3rd

## Invited Speakers:

- Team Megvii / (10:50am - 11:05am)
- Team Oks / (11:05am - 11:20am)