



# Learning to Refine Human Pose Estimation

Mihai Fieraru

Anna Khoreva

Leonid Pishchulin

Bernt Schiele

Max Planck Institute for Informatics, Saarbrücken, Germany

## Motivation

Keypoint localization still fails in challenging cases of various human pose estimation tasks.



Single-Person Pose Estimation

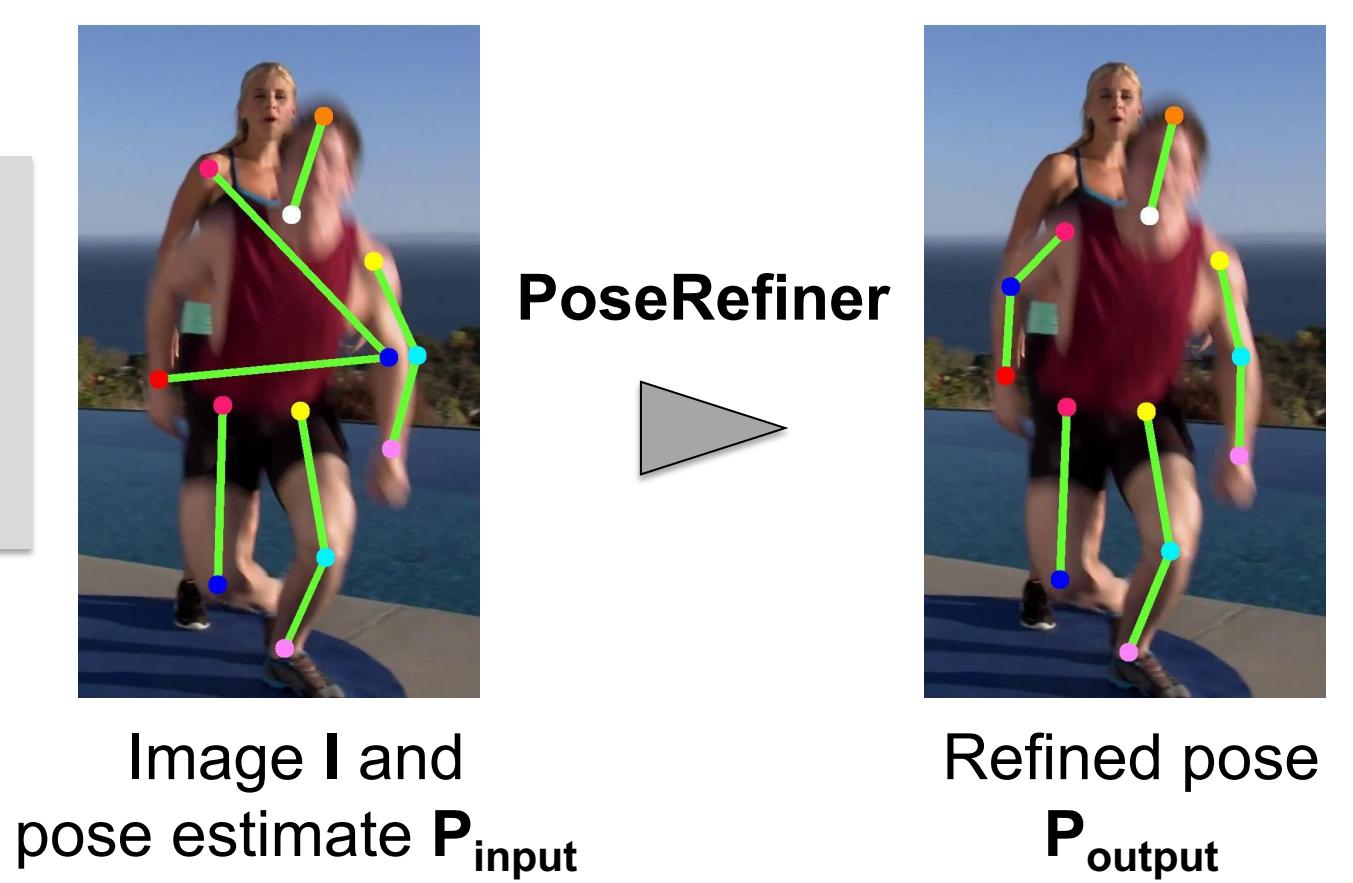


Multi-Person Pose Estimation

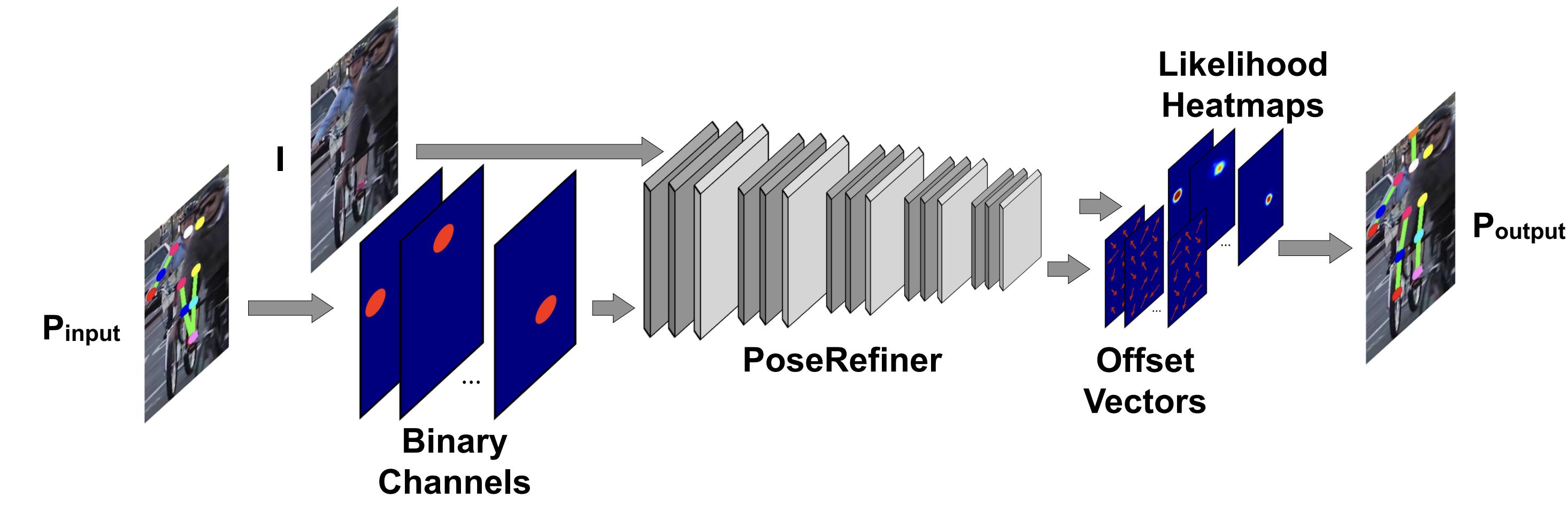


Multi-Person Articulated Tracking

**Goal:** Learn a post-processing step **PoseRefiner** that corrects localization mistakes of any estimated body pose.

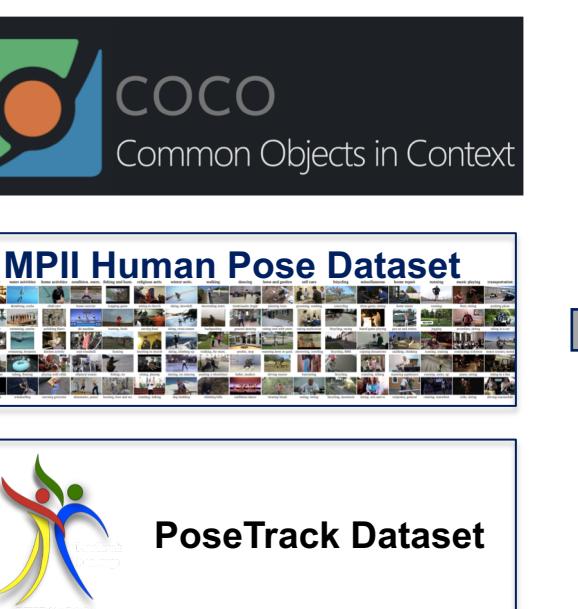


## Approach

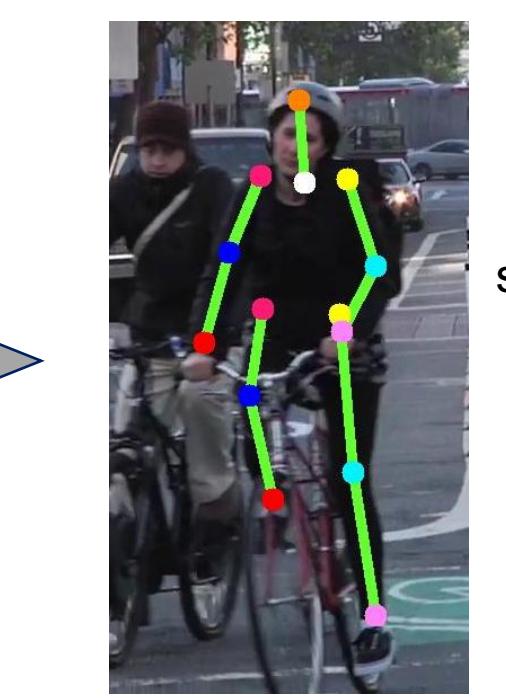


## Training Data

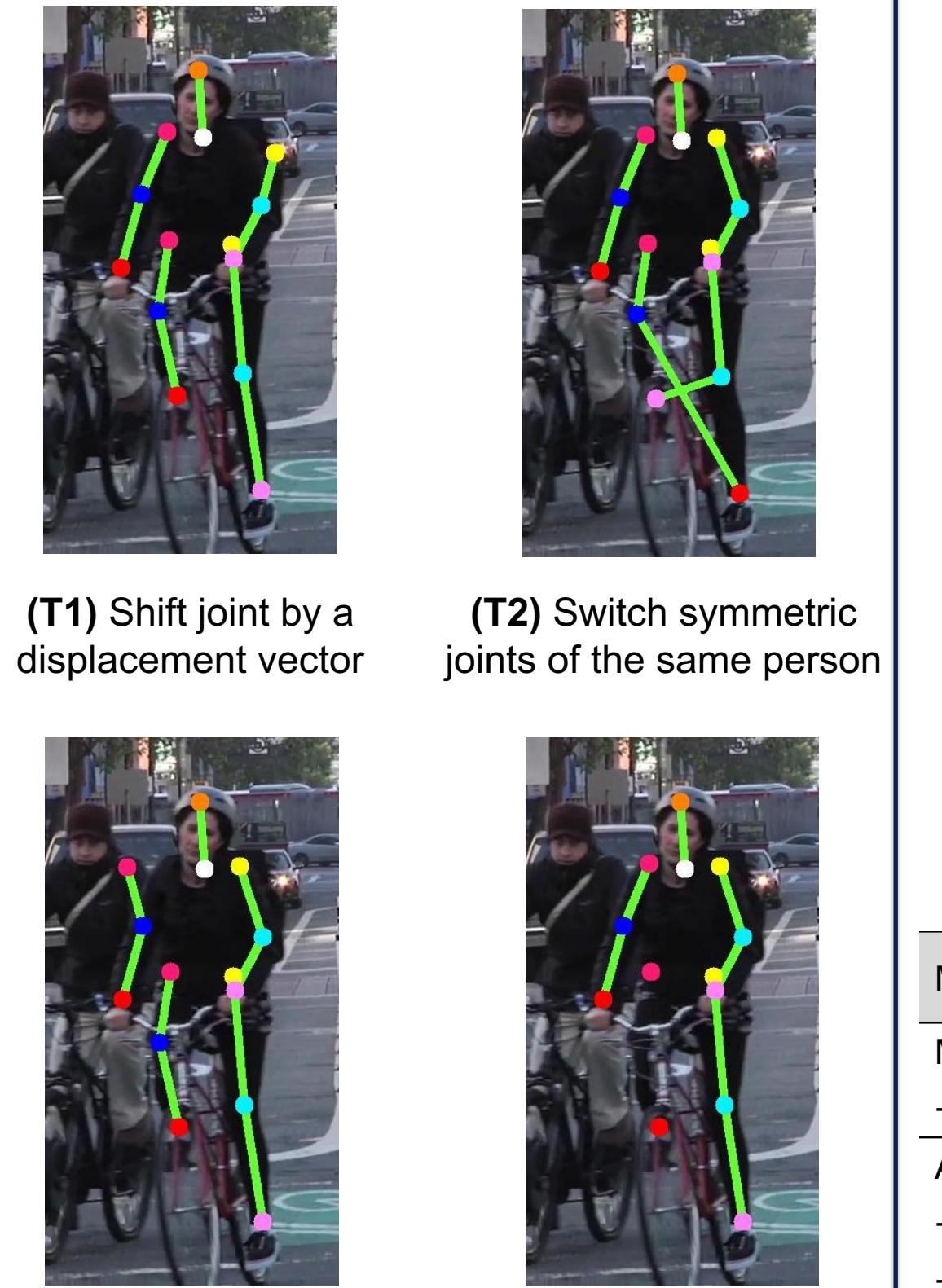
### Pose Estimation Datasets



### $I$ and $P_{\text{output}}$



### Examples of synthetic $P_{\text{input}}$



We synthesize  $P_{\text{input}}$  by applying several transformations to the ground truth  $P_{\text{output}}$ .

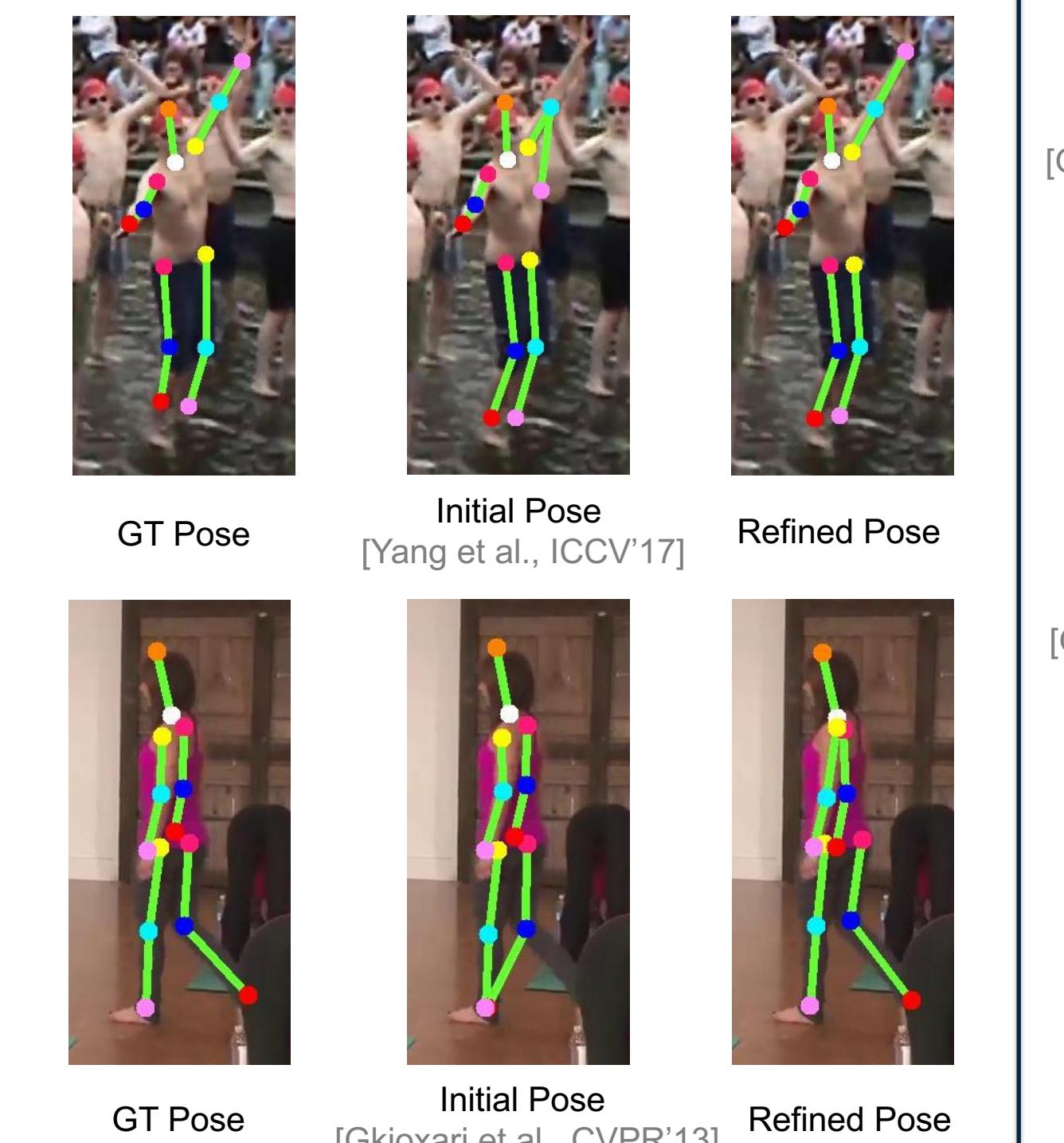
## Experimental Results

### MPII Single-Person Pose Estimation

Method	mPCK <sub>h</sub> @0.5	AUC	$\Delta$
Pyramid Residual Module [Yang et al., ICCV'17]	<b>92.0</b>	64.2	-
+ Refinement *	<b>92.0</b>	64.7	+0.3
Adversarial PoseNet [Chen et al., CoRR'17]	91.9	61.6	-
+ Refinement *	<b>92.1</b>	63.6	+1.1
DeeperCut [Insafutdinov et al., ECCV'16]	88.5	60.8	-
+ Refinement	<b>89.1</b>	62.3	+1.0
Chained Predictions [Gkioxari et al., CVPR'13]	86.1	57.3	-
+ Refinement	<b>88.0</b>	61.2	+2.9
Iterative Error Feedback [Carreira et al., CVPR'16]	81.3	49.1	-
+ Refinement	<b>85.6</b>	58.4	+6.8

\*Using a refinement model trained with only (T1) transformations.

Effect of PoseRefiner on MPII Single-Person [Andriluka et al., CVPR'14]



### MPII Multi-Person Pose Estimation

Method	mAP	$\Delta mAP$
Associative Embedding [Newell et al., NIPS'17]	77.5	-
+ Refinement	<b>78.0</b>	+0.5
Part Affinity Fields [Cao et al., CVPR'17]	75.6	-
+ Refinement	<b>76.9</b>	+1.3
ArtTrack [Insafutdinov et al., CVPR'17]	74.2	-
+ Refinement	<b>75.1</b>	+0.9
[Varadarajan et al., arXiv'17]	72.2	-
+ Refinement	<b>75.1</b>	+2.9

Effect of PoseRefiner on MPII Multi-Person [Andriluka et al., CVPR'14]



GT Pose      Initial Pose      Refined Pose

[Varadarajan et al., arXiv'17]    [Cao et al., CVPR'17]

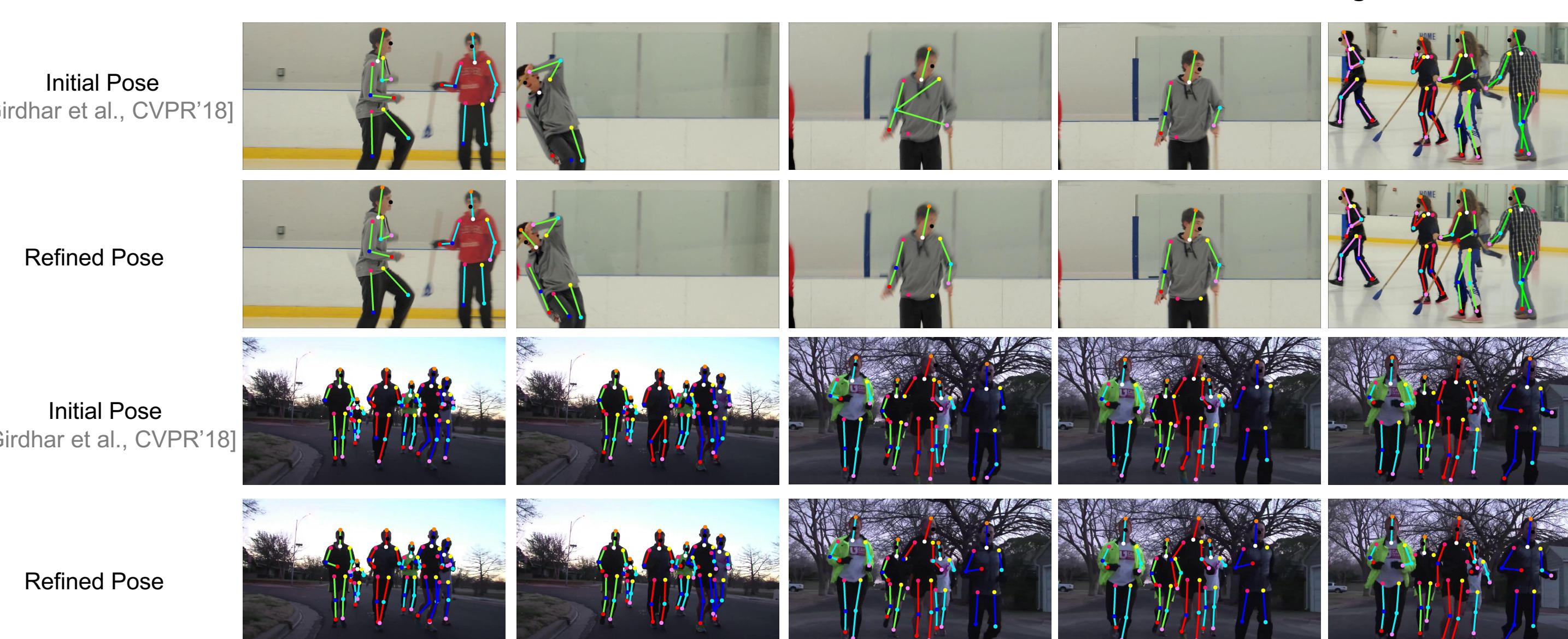
### PoseTrack Multi-Person Pose Estimation and Articulated Tracking

Method	mAP	$\Delta mAP$
ML_Lab [Zhu et al., ICCVw'17]	71.9	-
+ Refinement	<b>73.8</b>	+1.9
ArtTrack [Insafutdinov et al., CVPR'17] (best mAP)	68.6	-
+ Refinement (w/o noise)	<b>70.0</b>	+1.4
+ Refinement (with noise)	69.7	+1.1
BUTD [Jin et al., ICCVw'17] (best mAP)	67.8	-
+ Refinement	<b>70.9</b>	+3.1
Detect-and-Track [Girdhar et al., CVPR'18]	60.4	-
+ Refinement	<b>65.7</b>	+5.3

Effect of PoseRefiner on PoseTrack [Andriluka et al., CVPR'18]

Method	mAP	mMOTA	$\Delta mMOTA$
BUTD [Jin et al., ICCVw'17] (best mMOTA)	62.5	56.0	-
+ Refinement	<b>64.3</b>	<b>58.4</b>	+2.4
Detect-and-Track [Girdhar et al., CVPR'18]	60.4	55.1	-
+ Refinement	64.1	<b>57.3</b>	+2.2
ArtTrack [Insafutdinov et al., CVPR'17] (best mMOTA)	66.7	50.2	-
+ Refinement (w/o noise)	66.5	53.3	+3.1
+ Refinement (with noise)	67.0	<b>54.1</b>	+3.9
ML_Lab [Zhu et al., ICCVw'17]	71.9	48.6	-
+ Refinement	70.1	<b>53.5</b>	+4.9

Effect of PoseRefiner on PoseTrack [Andriluka et al., CVPR'18]



Systematic improvement over all methods (incl. state-of-the-art) across different datasets and tasks.