

#### StrongPose: Bottom-up and Strong Keypoint Heat Map Based Pose Estimation

Author/Presenter: Niaz Ahmad Corresponding author: Jongwon Yoon







## Bottom-up approach for pose estimation



Estimate all keypoints (e.g ears, eyes, nose and joints)







- Bottom-up
  - Estimate all keypoints (e.g ears, eyes, nose and joints)
  - Keypoints are grouped into human instances















StrongPose uses ResNet as a backbone network







- StrongPose uses ResNet as a backbone network
- Pose Estimation Module consists of two parts







- StrongPose uses ResNet as a backbone network
- Pose Estimation Module consist of two parts
  - ✓ Classification
  - ✓ Regression







- SKHM are generated for all hard and soft keypoints
- The role of SKHM is to correctly localize and produce the heat map for each keypoint







- BHM is generated in the same manner as SKHM
- BHM helps to correctly position the human body in the image







- To increase keypoint prediction we define keypoint offset vectore Vk(x)
- Vk(x) compare each predicted 2D keypoint position  $P_i$  with the ground truth position  $G_k$  in the keypoint disk
- The Loss between P<sub>i</sub> and G<sub>k</sub> is panelized by L1 loss







- To increase keypoint prediction we define keypoint offset vectore Vk(x)
- Vk(x) compare each predicted 2D keypoint position  $P_i$  with the ground truth position  $G_k$  in the keypoint disk
- The Loss between P<sub>i</sub> and G<sub>k</sub> is panelized by L1 loss







- To increase keypoint prediction we define keypoint offset vectore Vk(x)
- Vk(x) compare each predicted 2D keypoint position  $P_i$  with the ground truth position  $G_k$  in the keypoint disk
- The Loss between P<sub>i</sub> and G<sub>k</sub> is panelized by L1 loss







 Pose plot module defines associations between keypoints as tuples and group all the associated keypoints into human instances



#### Advantages



Strong Keypoint Heat Map

 High detection rate of adjacent keypoints





#### Advantages





2. Better performance on hidden and occluded keypoints

**Plot Pose** 





#### Advantages





3. Better prediction of keypoints of highly entangled people





## **Visualization results**



#### Single-person pose estimation



Multi-person pose estimation



github.com/niazahamd89/StrongPose



## Evaluation on COCO Val2017 Dataset



Method	Backbone	Input Size	OHKM	AP	AR
Top-down:					
8-stage Hourglass	-	156 x 192	×	0.669	-
8-stage Hourglass	-	156 x 156	×	0.671	-
CPN	ResNet-50	256 x 192	×	0.686	-
CPN	ResNet-50	384 x 288	×	0.706	-
CPN	ResNet-50	256 x 192	$\checkmark$	0.694	-
CPN	ResNet-50	384 x 288	$\checkmark$	0.716	-
HRNet-W48	HRNet-W48	384 x 288	×	0.763	0.812
Bottom-up:					
CMU-Pose	-	-	×	0.618	-
PersonLab(single-Scale)	ResNet-152	-	×	0.665	0.707
PersonLab(multi-scale)	ResNet-152	-	×	0.687	-
StrongPose	ResNet-101	-	×	0.690	0.757
StrongPose	ResNet-152	-	×	0.728	0.800

- 🔏 5.7 % in AP compare to Hourglass
- <sup>1</sup> 2.2 % in AP compare to CPN
- 1.2 % in AP compare to CPN. (OHKM)

- 11.0 % in AP compare to CMU-Pose
- 4.1 % in AP compare to PersonLab

### **Evaluation on COCO Test2017 Dataset**



Method	AP	AP <sup>.50</sup>	AP.75	$AP^M$	$AP^L$
Top-down:					
Mask-RCNN	0.631	0.873	0.687	0.578	0.714
G-RMI COCO-only	0.649	0.855	0.713	0.623	0.700
CPN	0.721	0.914	0.800	0.687	0.772
Bottom-up:					
CMU-Pose (+refine)	0.618	0.849	0.675	0.571	0.682
Assoc. Embed(single-Scale)	0.630	0.857	0.689	0.580	0.704
Assoc. Embed(mscale, refine)	0.655	0.879	0.777	0.690	0.752
PersonLab (single-scale)	0.665	0.880	0.726	0.624	0.723
PersonLab (multi-scale)	0.687	0.890	0.754	0.641	0.755
MultiPoseNet	0.696	0.863	0.766	0.650	0.763
StrongPose:					
ResNet101	0.708	0.889	0.752	0.652	0.753
ResNet152	0.725	0.891	0.778	0.671	0.762

- 9.4 % in AP compare to Mask-RCNN
- 7.6 % in AP compare to G-RMI
- 🔏 0.4 % in AP compare to CPN
- 10.7 % in AP compare to CMU-Pose

- 7.0 % in AP compare to Assoc
- 3.8 % in AP compare to PersonLab
- 2.9 % in AP compare to MultiPoseNet





#### Conclusion

- The proposed system jointly tackle the problem of pose estimation by using Strong Keypont Heat Map and person detection by using Body Heat Map
- Significant improvement of Average Precision is delivered on the COCO 2017 keypoint challenging dataset

#### Future Work

- Understanding human body language
- Activity recognition in live environment





# **Thank You**

**Paper ID: 2353** 

