G33: Single Person 2-D Pose Estimation from RGB images using Convolutional Neural Networks

Akhil Bommadevara (abommad) | Ishan Bhatt (ivbhatt) | Aishwarya Seth (aseth)

Motivation:

Human Pose-Estimation(HPE) has a wide range of applications:

- Activity recognition, Person tracking, Robust Person Detection
- VR gaming, Animation (Style Transfer based approaches)
- Physical Therapy/Rehabilitation, Posture improvement, Assisted Living

Problem Definition:

We are trying to build a HPE system that:

- Takes RGB images as input.
- Assumes images to have a subject person without significant occlusions.
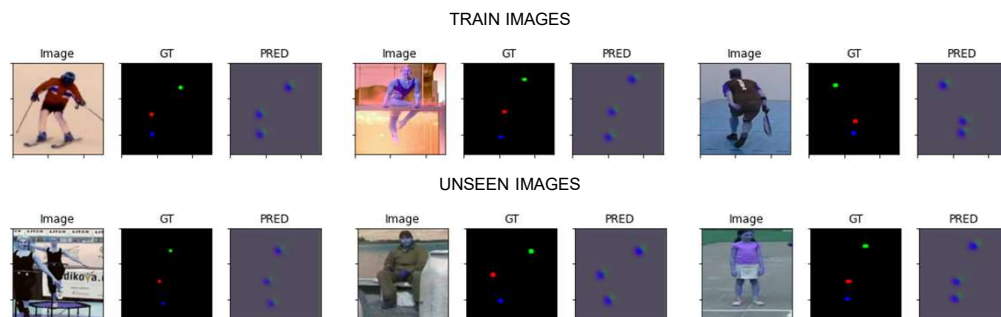- Detects the visible key points of the subject person.

Data[6]:

- MPII Human Pose dataset (25K images) of over 40K people with annotated body joints **(left/right) (ankle, knee, hip, shoulder, elbow, wrist)** + **head** + thorax(upper-neck) + pelvis.
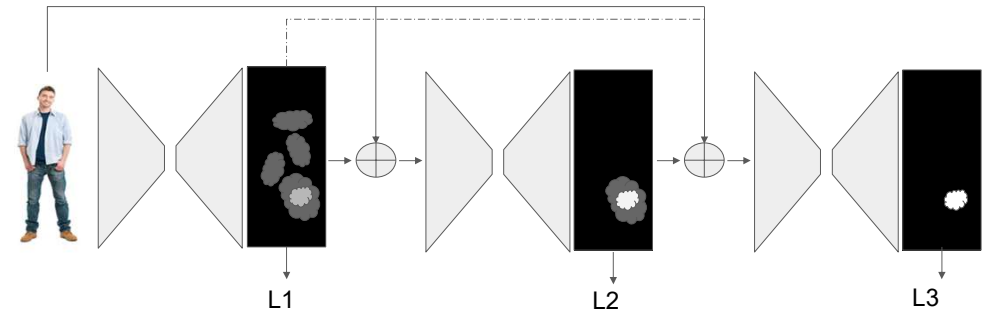- About 4K people (bounding-boxes) with all 13 key-points visible.

Results:

- Used about 20 images (10 epochs) to train; **right-ankle, right-knee, left-shoulder** on single hourglass module .
- State-of-the-Art PCK-h@0.5 = 92.5 [7]

TRAIN IMAGES

UNSEEN IMAGES

Model[4][5]:

i. Stacked Hourglass:

L1   L2   L3

ii. Hourglass unit:

(WIDTH/1) x (HEIGHT/1) x 3
(WIDTH/2) x (HEIGHT/2) x (S*1)

(WIDTH/4) x (HEIGHT/4) x (S*2)

(WIDTH/4) x (HEIGHT/4) x (S*4)

(WIDTH/2) x (HEIGHT/2) x (S*2)
(WIDTH/1) x (HEIGHT/1) x F

(WIDTH/1) x (HEIGHT/1) x C
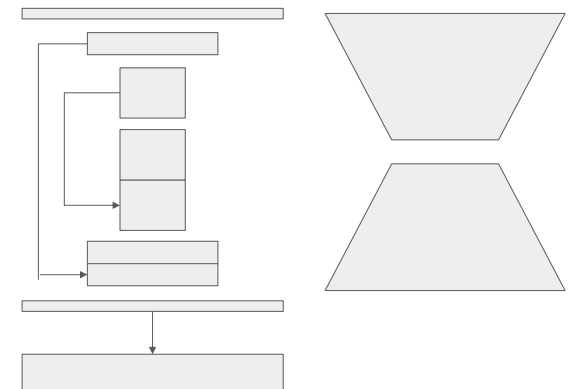
References:

1. A. Yao, et. al., ''Coupled action recognition and pose estimation from multiple views,'' Int. J. Comput. Vis., Oct 2012
2. M. B. Holte, et. al., ''HPE and activity recognition from multi-view videos: Comparative explorations of recent developments,'' IEEE J. Sel. Topics Signal Process., Sep 2012
3. Spagnolo & Paolo, ''Proceedings IEEE conference on advanced video & signal based surveillance. AVSS 2003,'' in Proc. IEEE Conf. Adv. Video Signal Based Surveill.
4. S.-E. Wei, et. al.. "Convolutional pose machines" in CVPR, 2016
5. A. Newell, 'Stacked Hourglass Networks for Human Pose Estimation", arxiv, July 2016
6. M. Andriluka, et. al., "2D Human Pose Estimation: New Benchmark and State of the Art Analysis", CVPR, 2014
7. Zhang, Hong, et al. "HPE with spatial contextual information." arxiv, 2019