VL4Pose: Active Learning Through Out-Of-Distribution Detection For Pose Estimation

Megh Shukla
Roshan Roy *
Pankaj Singh *
Shuaib Ahmed
Alexandre Alahi

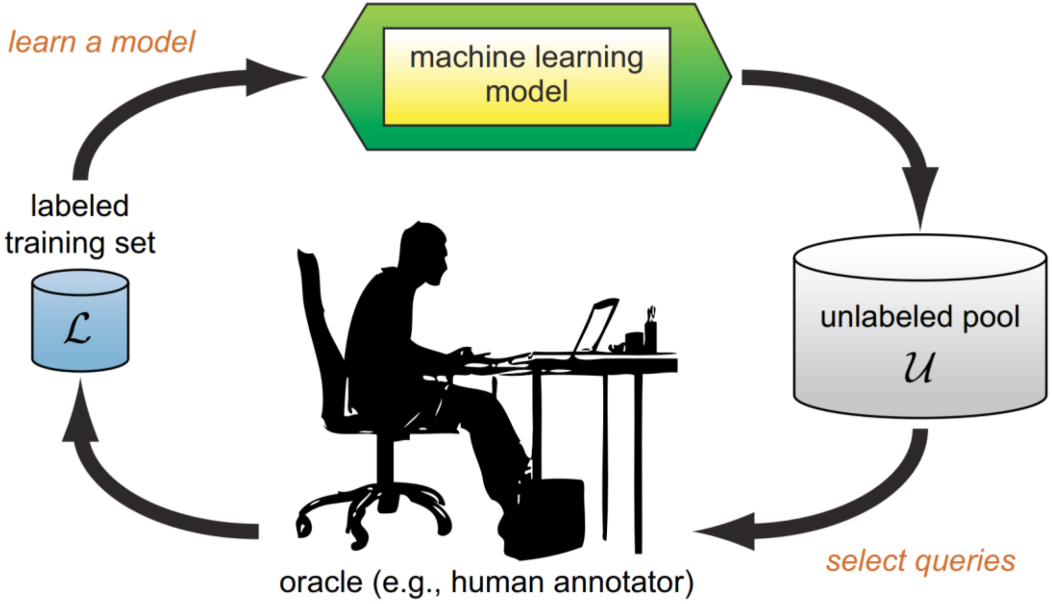
November 2022

Researcher: Out-Of-Distribution For Pose Estimation?

Engineer: How can we improve pose estimation in production?

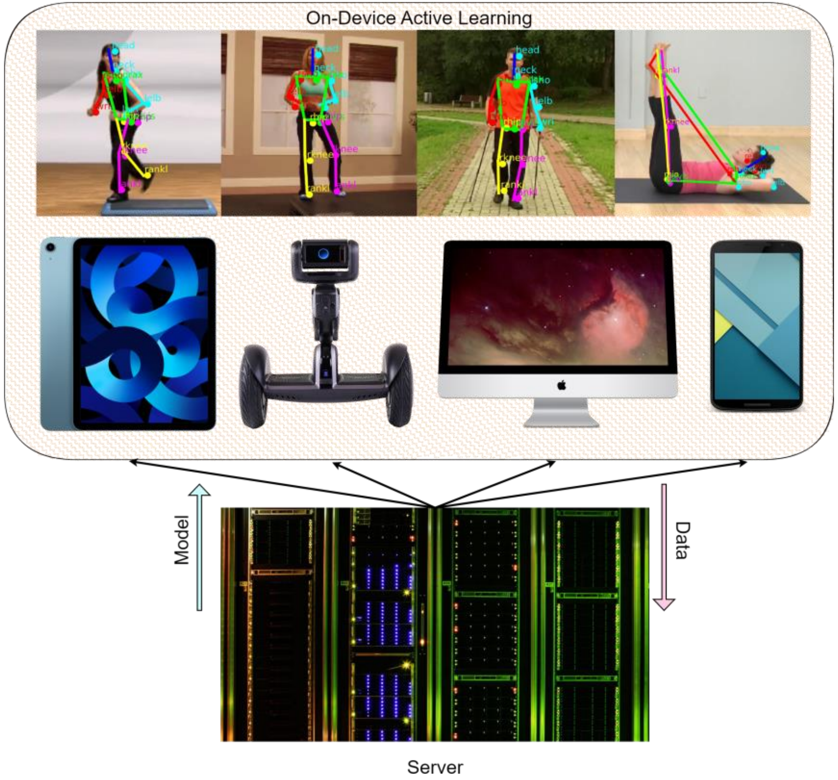
Introduction

Active Learning



Settles, Burr. "Active learning literature survey." (2009).

On-Device Active Learning



Compute Hardware ✗
Real-Time Data ✓

Compute Hardware ✓
Real-Time Data ✗

Related Work

VL4Pose: Active Learning Through Out-Of-Distribution Detection For Pose Estimation

A Mathematical Analysis of Learning Loss for Active Learning in Regression

Megh Shukla ✉ Shuaib Ahmed
Mercedes-Benz Research and Development India

Sequential Graph Convolutional Network for Active Learning

Razvan Caramalau[1], Binod Bhattarai[1] and Tae-Kyun Ki...
[1]Imperial College London, UK
[2]KAIST, South Kore...

Meta Agent Teaming Active Learning for Pose Estimation

Jia Gong[1] Zhipeng Fan[2] Qiuhong Ke[3] Hossein Rahmani[4] Jun Liu[1*]
[1]Singapore University of Technology and Design, Singapore; [2]New York University, United States
[3]The University of Melbourne, Australia; [4]Lancaster University, United Kingdom

Active Learning for Human Pose Estimation

Buyu Liu
University of Edinburgh
Vittorio Ferrari
University of Edinburgh

Bayesian Uncertainty and Expected Gradient Length - Regression: Two Sides Of The Same Coin?

Megh Shukla
Mercedes-Benz Research and Development India

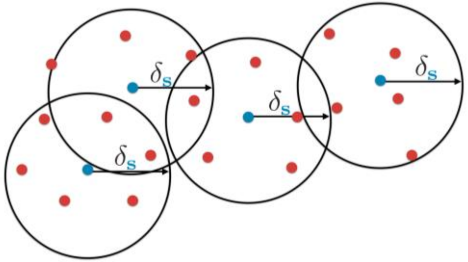
Learning Loss for Active Learning

Donggeun Yoo[1,2] and In So Kweon[2]
[1]Lunit Inc., Seoul, South Korea.
[2]KAIST, Daejeon, South Korea.

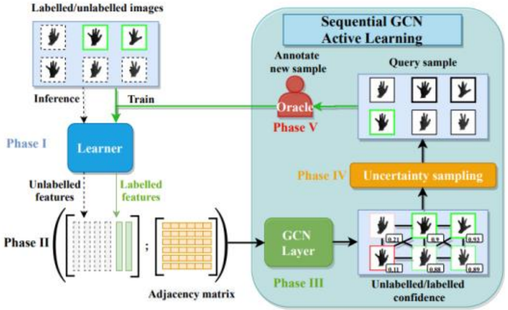
Active Learning for Bayesian 3D Hand Pose Estimation

Razvan Caramalau[1], Binod Bhattarai[1], and Tae-Kyun Kim[1,2]
[1]Imperial College London, UK
[2]KAIST, South Korea

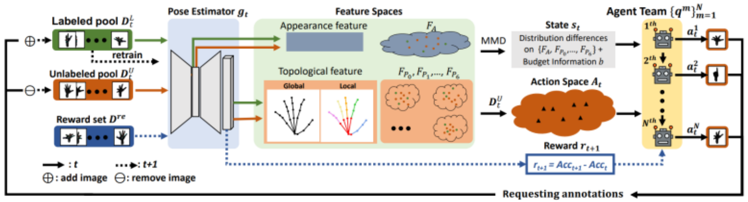
Non Real Time



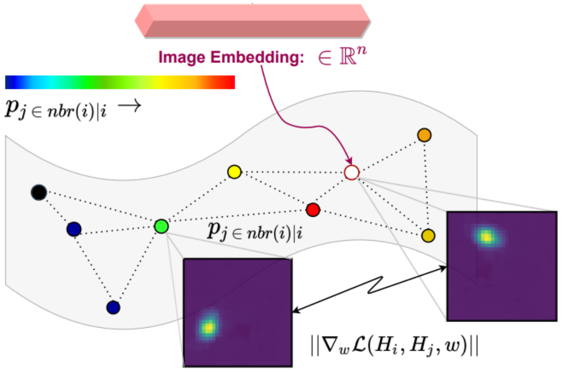
Sener and Savarese. *Active Learning for Convolutional Neural Networks*, ICLR 2018



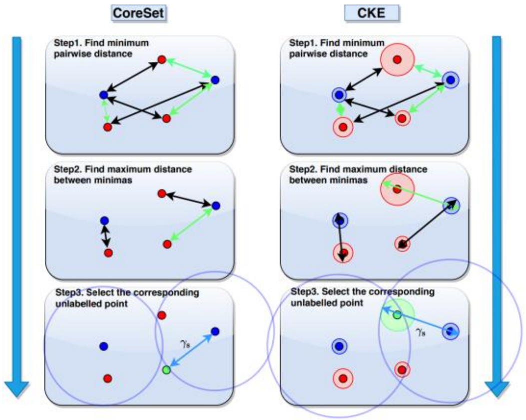
Caramalau, Bhattarai, and Kim. *Sequential graph convolutional network for active learning*, CVPR 2021.



Gong, Jia, et al. *Meta agent teaming active learning for pose estimation.* CVPR 2022.

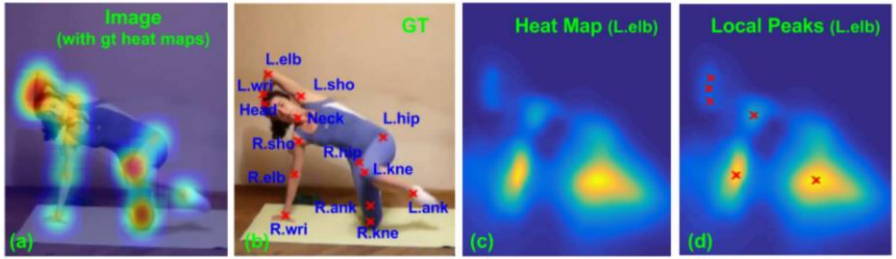


Shukla. *Bayesian Uncertainty and Expected Gradient Length – Regression: Two Sides of the Same Coin?*, WACV 2022

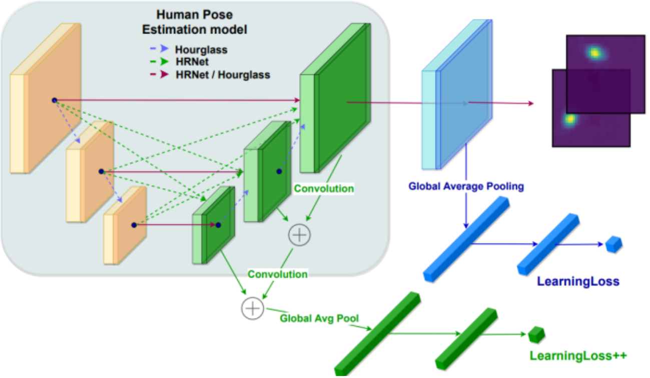


Caramalau, Bhattarai, and Kim. *Active learning for bayesian 3d hand pose estimation*, WACV 2021

Real Time



Liu and Ferrari. *Active learning for human pose estimation*. CVPR 2017



Shukla and Ahmed. *A mathematical analysis of learning loss for active learning in regression*. CVPR Workshops 2021
Yoo and Kweon. *Learning loss for active learning*. CVPR 2019.

$$\mathrm{Var}(\mathbf{y}) \approx \frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{y}}_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T}\hat{\mathbf{y}}_t\right)^2 + \frac{1}{T}\sum_{t=1}^{T}\hat{\sigma}_t^2$$

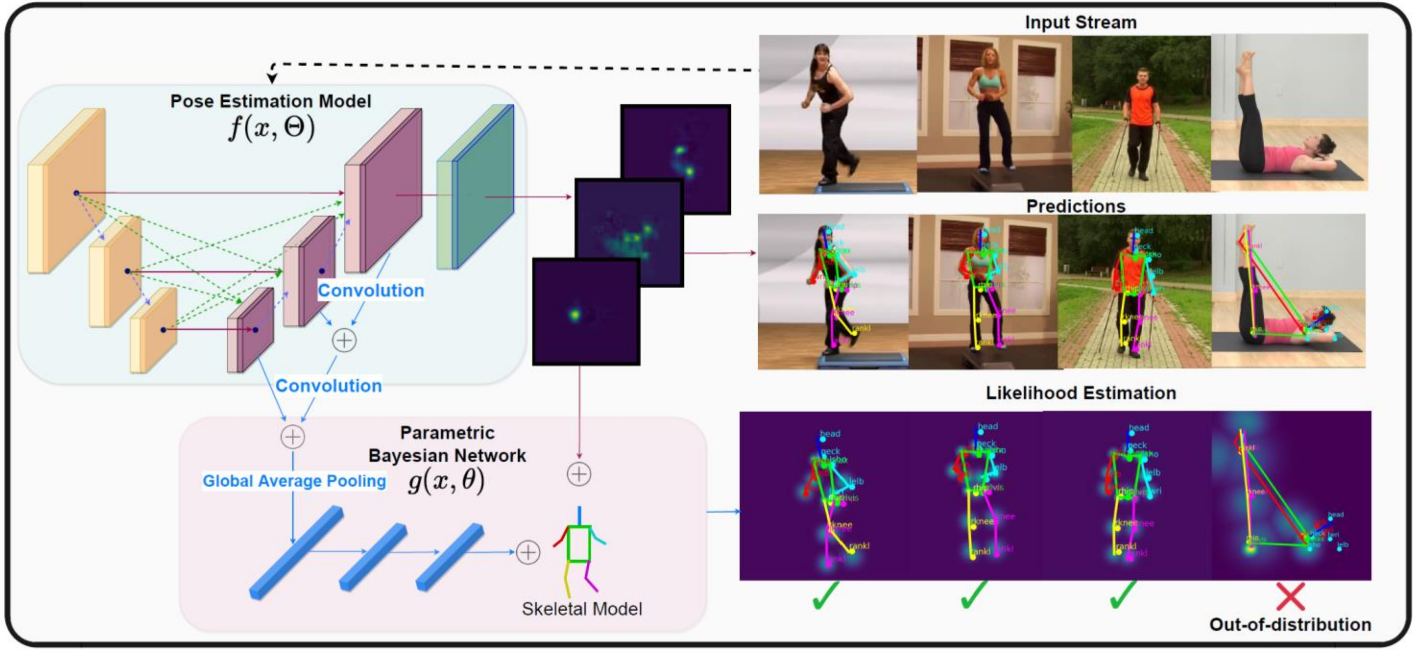
Kendall and Gal. *What uncertainties do we need in bayesian deep learning for computer vision?.* NeurIPS 2017

Megh Shukla

Methodology

VL4Pose: A First Principles Approach to Active Learning for Pose Estimation

VL4Pose: Intuition



Out-of-Distribution detection = Maximize likelihood of training distribution

VL4Pose: Likelihood

- Distribution over joints (Y_i)

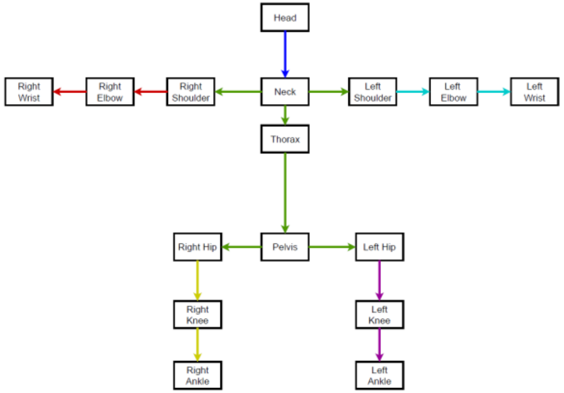
$$q_{BN}(y_1, y_2 \ldots y_N | x, \theta)$$

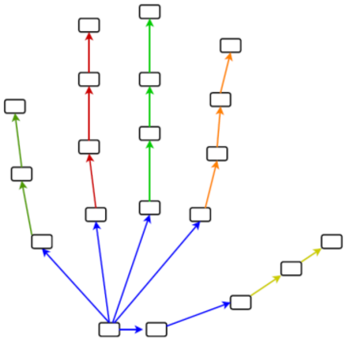
- Applying Chain Rule

$$q(y_1 | y_2 \ldots y_N, x, \theta) q(y_2 | y_3 \ldots y_N, x, \theta) \ldots q(y_N | x, \theta)$$

- Markov Blanket

$$q_{BN}(y_1, y_2 \ldots y_N | x, \theta) = \left[\prod_{i=1}^{N-1} q(y_i | y_{i+1}, x, \theta) \right] q(y_N | x, \theta)$$





VL4Pose: *Expected* Likelihood

- But wait … what if Y is a random variable given X?
  - For instance in human pose … $p_{pose}(Y) = p_{pose}(y_1, y_2 \ldots y_N) = \prod_{i=1}^{N} p(y_i)$

- We get Expected Log-Likelihood!

$$\mathbb{E}_Y \left[\log\, q_{BN}(y_1, y_2 \ldots y_N | x, \theta) \right]$$

- A bit of solving gives us:

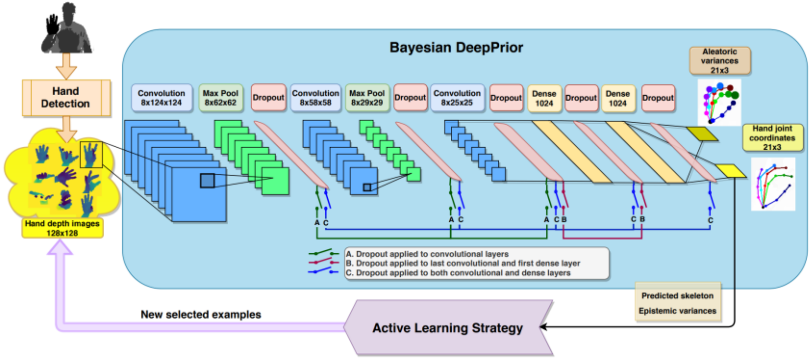
$$\sum_{Y} \left[p_{pose}(y_N) \log\, q_{BN}(y_N | x, \theta) + \sum_{i}^{N-1} p_{pose}(y_i) \log\, q_{BN}(y_i | y_{i+1}, X, \theta) \right]$$

Joint uncertainty

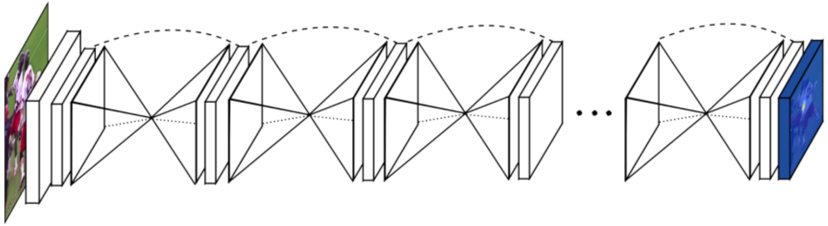
Pose uncertainty

Megh Shukla

VL4Pose: *Expected* Likelihood



Caramalau, Bhattarai, and Kim. *Active learning for Bayesian 3d hand pose estimation*, WACV 2021



Newell, Yang, and Deng. *Stacked hourglass networks for human pose estimation*. ECCV 2016

$$y_i \in \mathbb{R}^{\text{joints}\times 3}$$

$$q_{BN}(y_i|y_{i+1},x,\theta) = \mathcal{N}\left(y_i - [y_{i+1} + \hat{o}_i], \Sigma_i\right)$$

$$p(y_i) = \{1 \text{ at ground truth location, } 0 \text{ otherwise}\}$$

$$h \in \mathbb{R}^{\text{joints}\times 64\times 64}$$

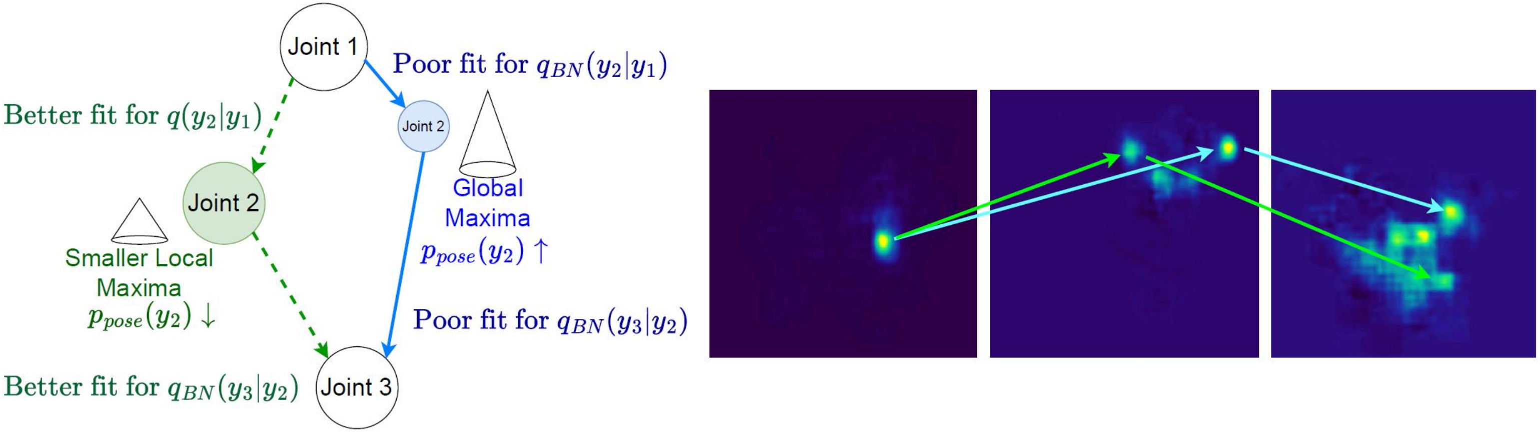
$$q(y_i|y_{i+1},x,\theta) = \mathcal{N}(\texttt{dist}(y_i,y_{i+1}) - \hat{d}_i\,,\, \sigma_i)$$

$$\hat{p}(y_i) = \texttt{softmax}(\texttt{local\_maxima}(h_i))$$

BMVC 2022

VL4Pose: Pose Refinement (Heatmaps)

$$\sum_{Y}\left[p_{pose}(y_N)\log q_{BN}(y_N|x,\theta) + \sum_{i}^{N-1} p_{pose}(y_i)\log q_{BN}(y_i|y_{i+1},X,\theta)\right]$$

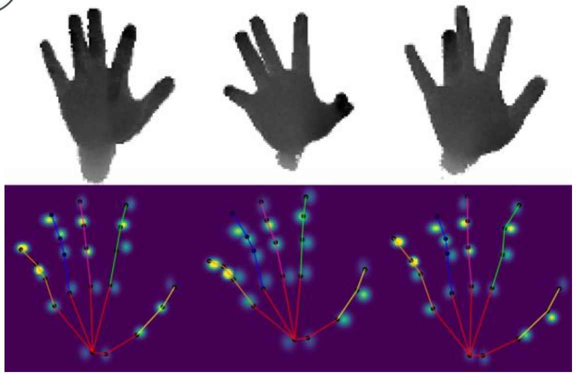


Interplay between *p* and *q*!

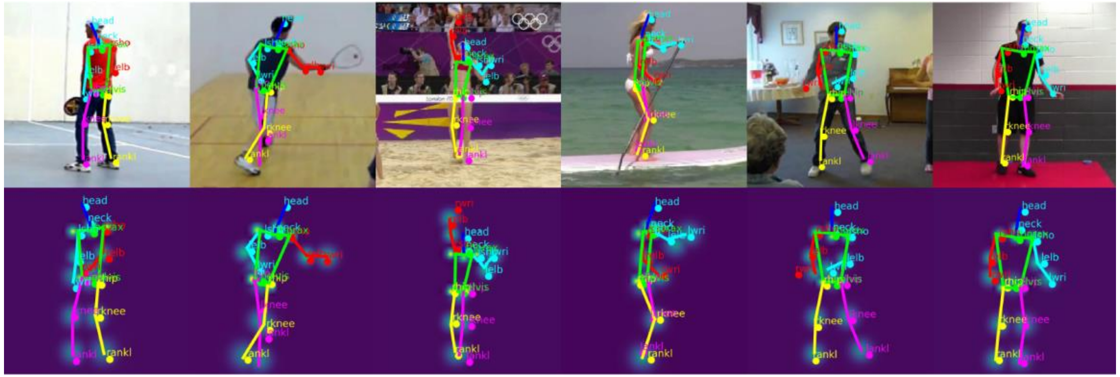
Results

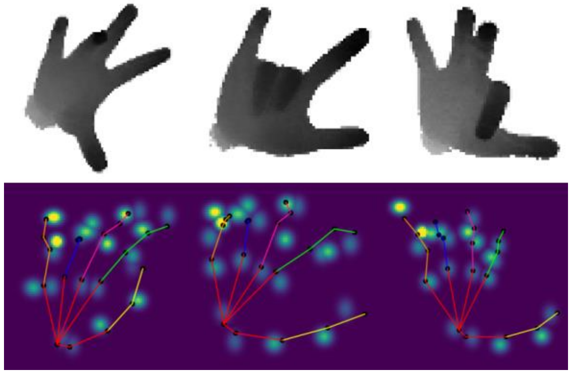
Qualitative and Quantitative Analysis

Qualitative Results

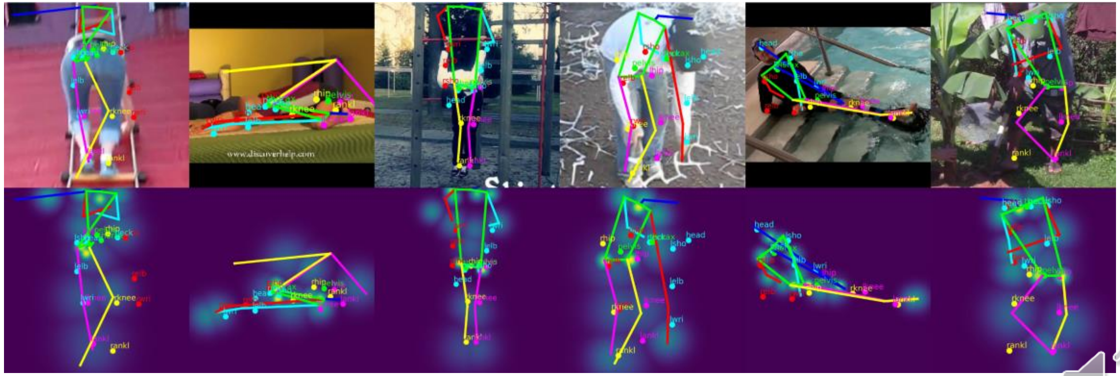


Maximum Likelihood



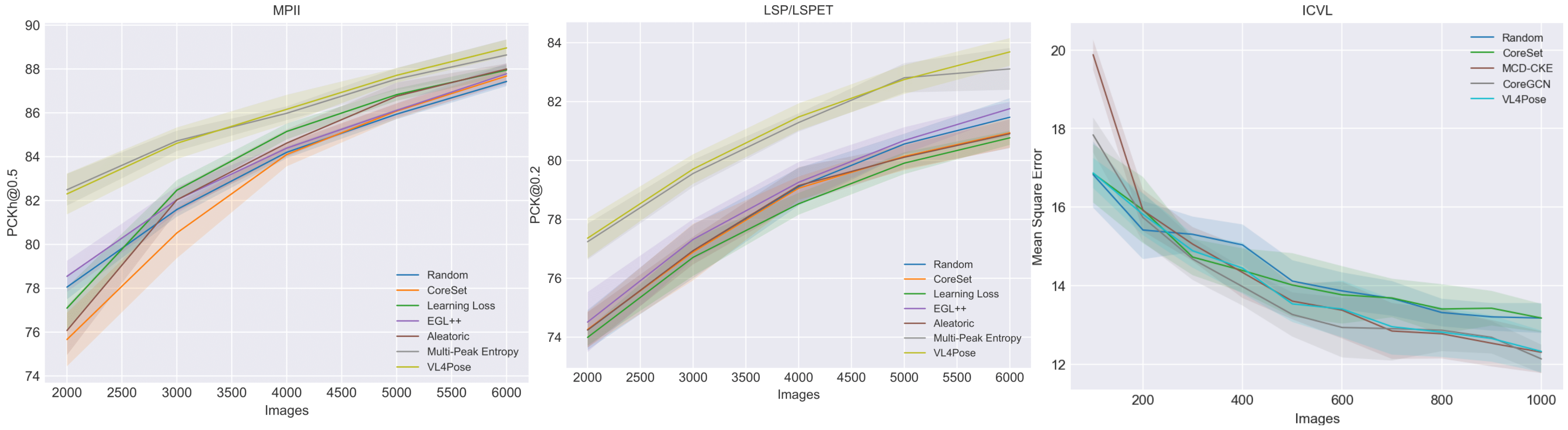


Minimum Likelihood



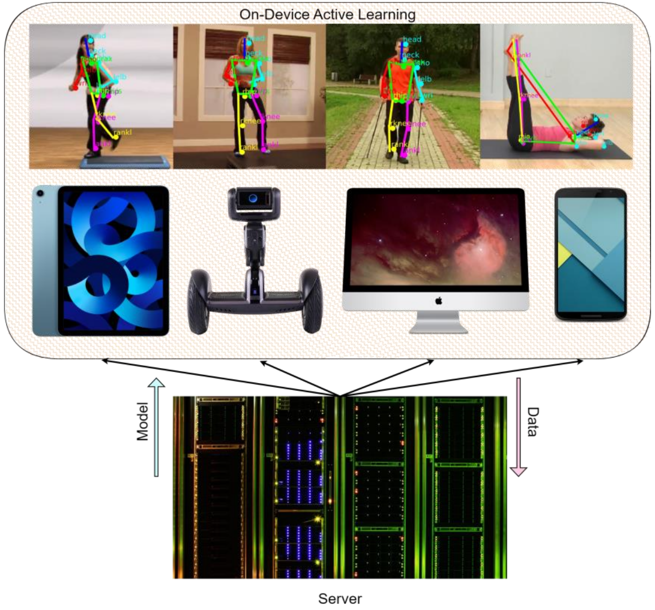
Quantitative Results

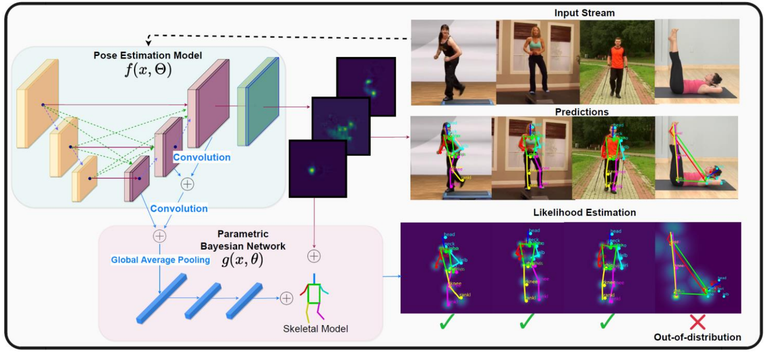
Megh Shukla



BMVC 2022

Pose Refinement





Conclusion

How far can simple domain knowledge take us?

1. Lightweight and real-time
2. Unifies joint and pose uncertainty
3. Tackles three problem statements
   1. Out-Of-Distribution
   2. Active Learning
   3. Pose refinement