

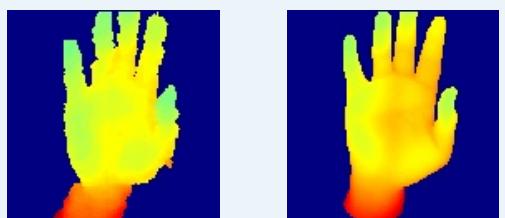
Crossing Nets: Combining GANs and VAEs with a shared latent space for Hand Pose Estimation

Chengde Wan¹, Thomas Probst¹, Luc Van Gool^{1,3}, Angela Yao²

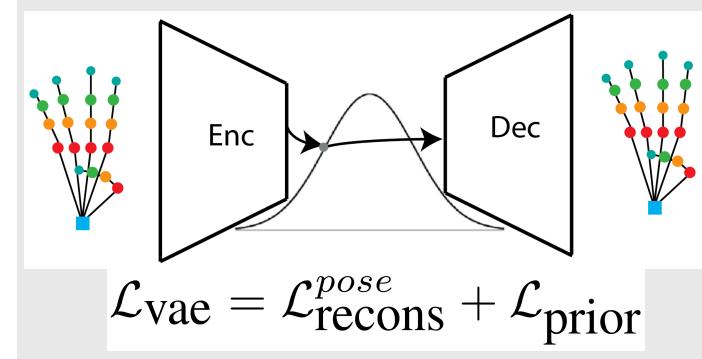
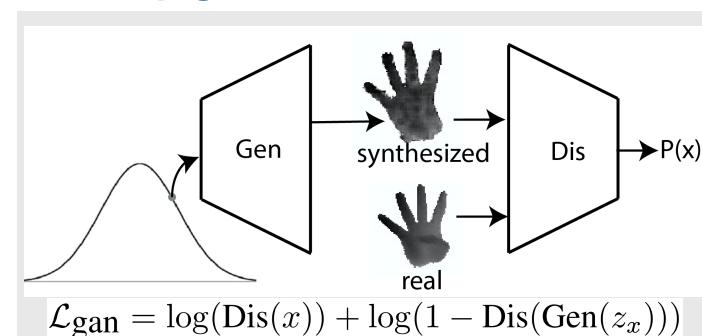
¹Computer Vision Laboratory, ETH Zurich; ²Institute of Computer Science, University of Bonn; ³VISICS, EAST, KU Leuven

1 Motivation

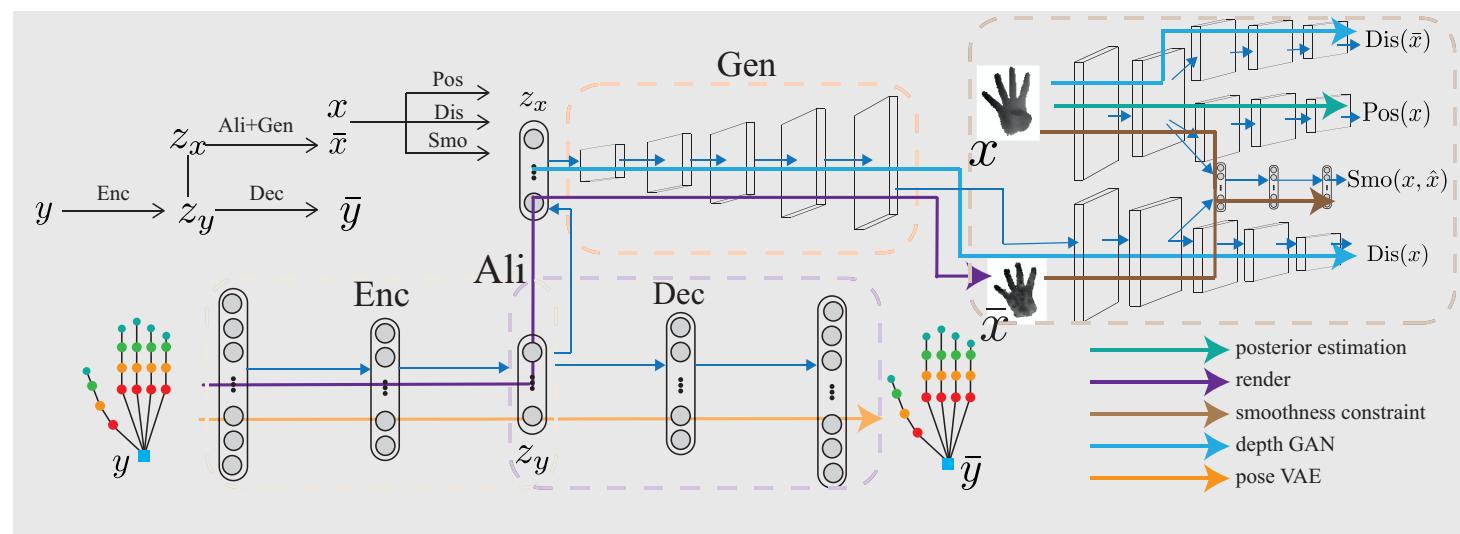
- Annotation is expensive
- Synthesized samples are not realistic
- Unlabeled data is cheap to collect
- Recent powerful deep generative models



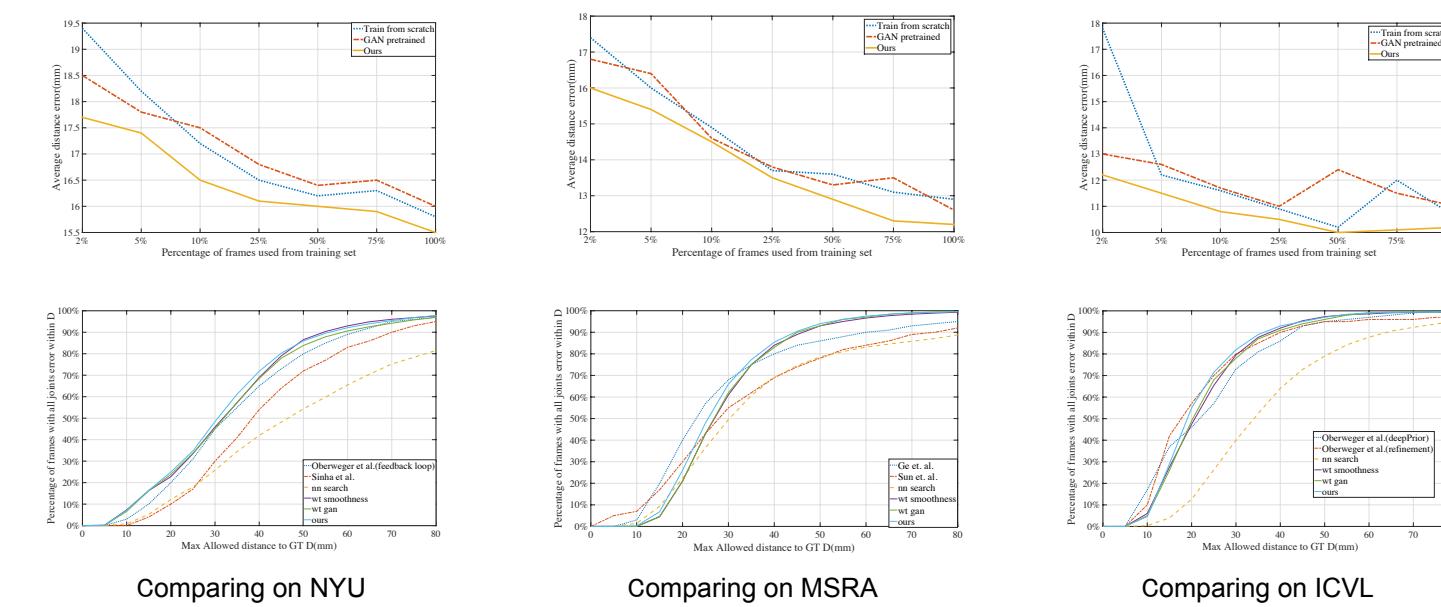
2 Deep generative model



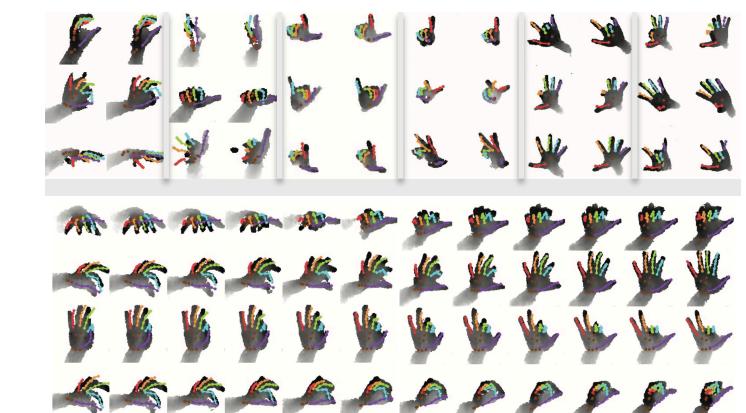
3 Method



4 Quantitative results



5 Qualitative results



6 Conclusion

- Extend GAN to semi-supervised setting for real valued structured prediction
- Synthesize highly realistic depth maps given hand pose
- A multi-task setting for pose estimation to regularize the pose regressor

7 Contacts

- Chengde Wan (wanc@vision.ee.ethz.ch)
- <http://www.vision.ee.ethz.ch/~wanc/>
- Source code and pretrained model available:**
- github.com/melonwan/crossingNet

