

# MURAUER: Mapping Unlabeled Real Data for Label AUstERity

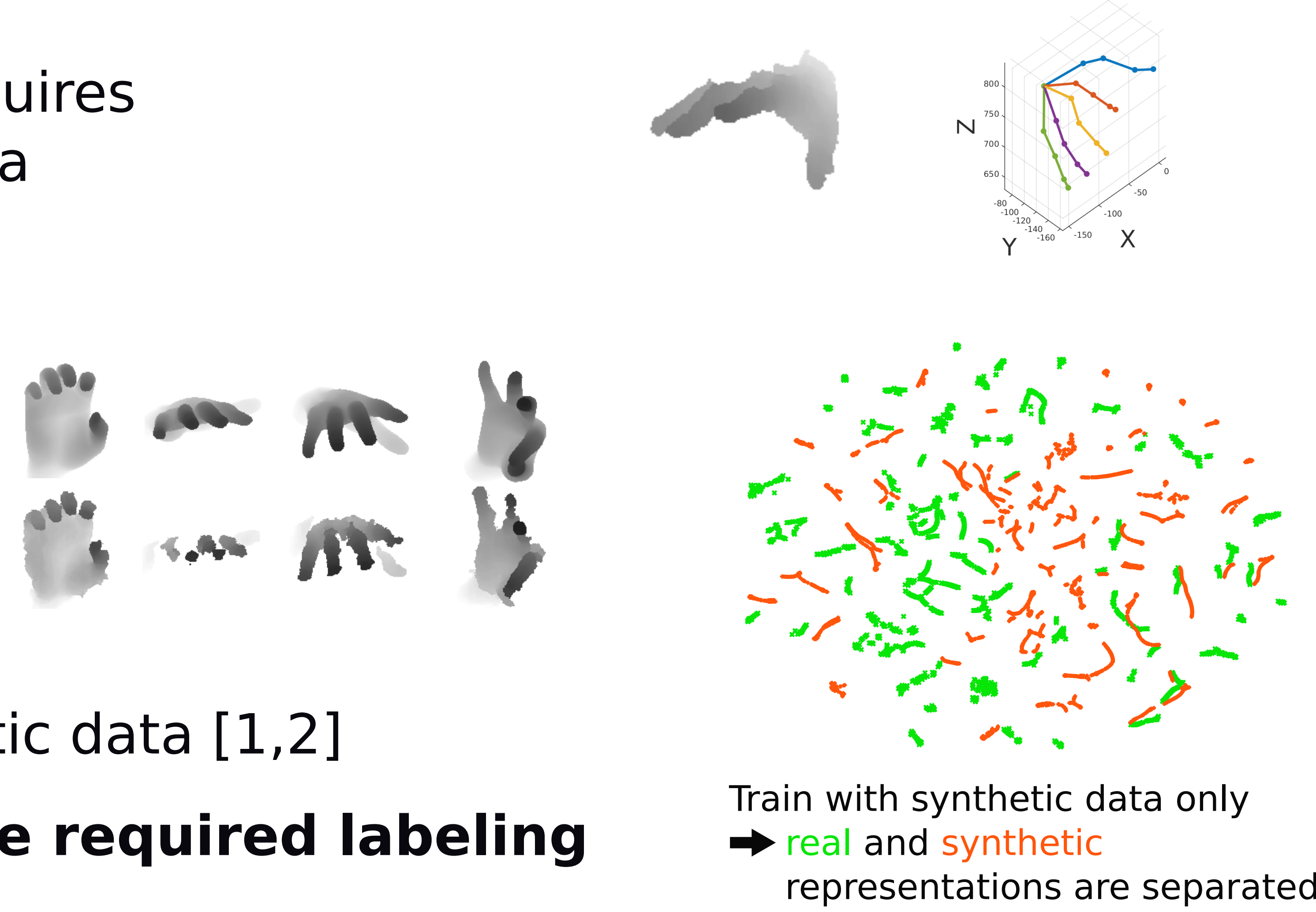
Georg Poier   Michael Opitz   David Schinagl   Horst Bischof

Project page:  
poier.github.io/murauer



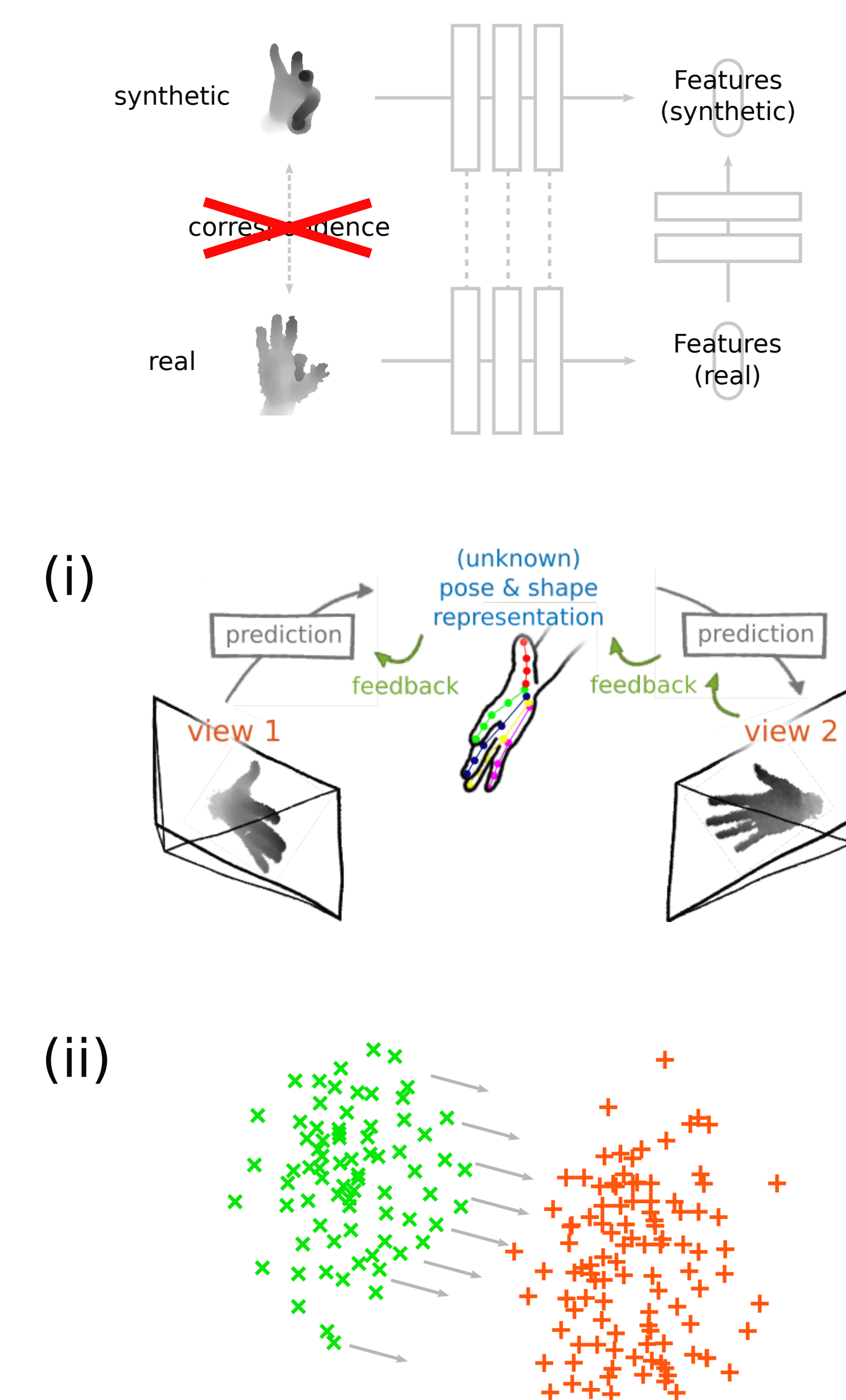
## Motivation

- Learning accurate models requires a large amount of labeled data
- Accurate labeling vital
- Synthetic data can help  
→ **But: domain gap**
- Mitigated using corresponding real ↔ synthetic data [1,2]  
→ **But: using correspondence required labeling**

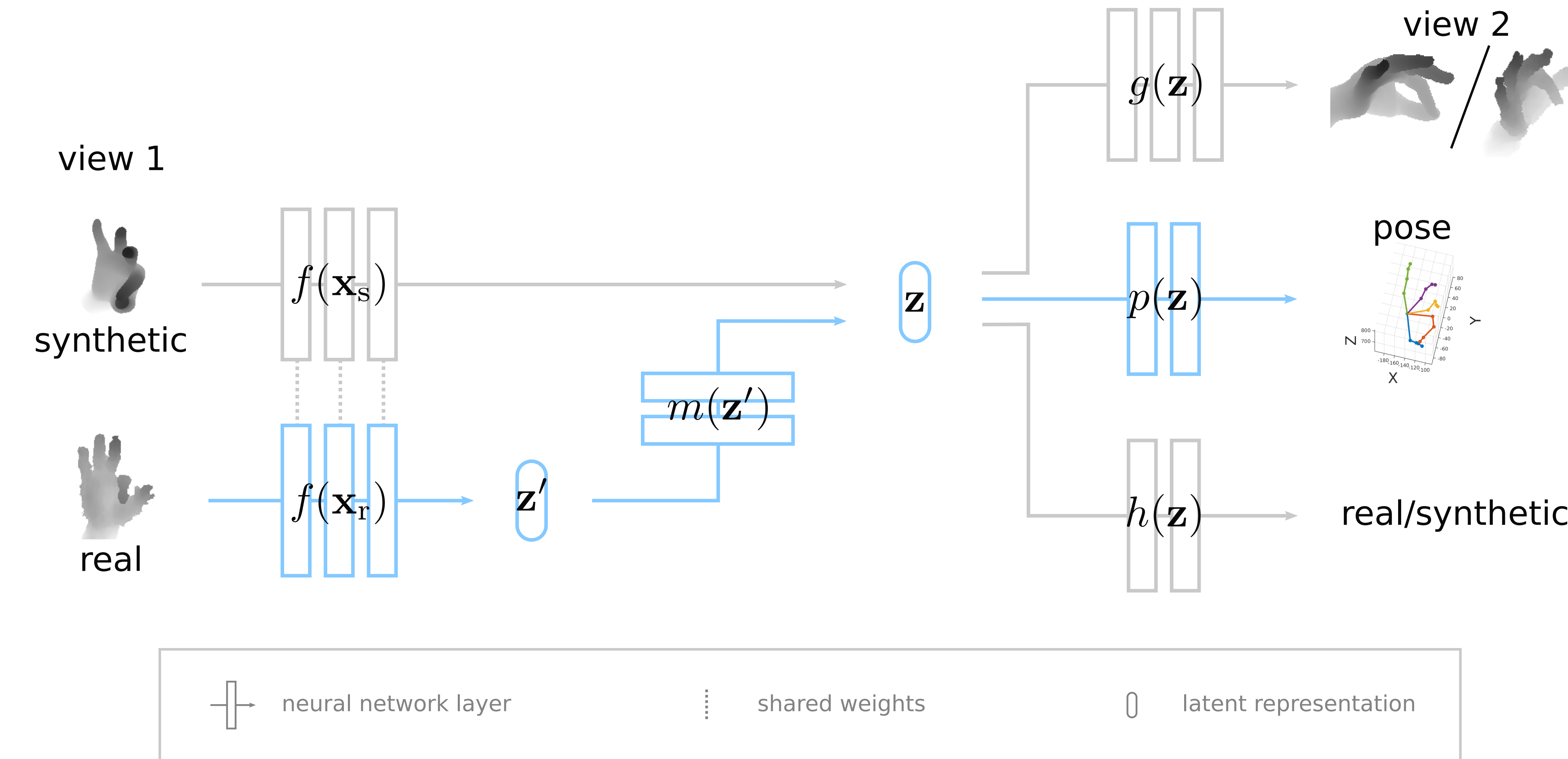


## Idea

- Map features real to synthetic without labels/correspondence
- Using two auxilliary objectives computed from unlabeled data:
  - enforcing pose specificity [3] (use 2 views, predict one from the other)
  - enforcing to align real and synthetic samples (make distributions indistinguishable)



## Implementation



- Overall loss: 
$$\ell = \underbrace{\ell_p + \lambda_c \ell_c}_{\text{labeled}} + \underbrace{\lambda_g \ell_g + \lambda_m \ell_m}_{\text{unlabeled}}$$
  - Enforce pose specificity (by learning to predict/reconstruct other view [3]):
- Enforce feature distribution alignment (adversarial; LS-GAN [4]):

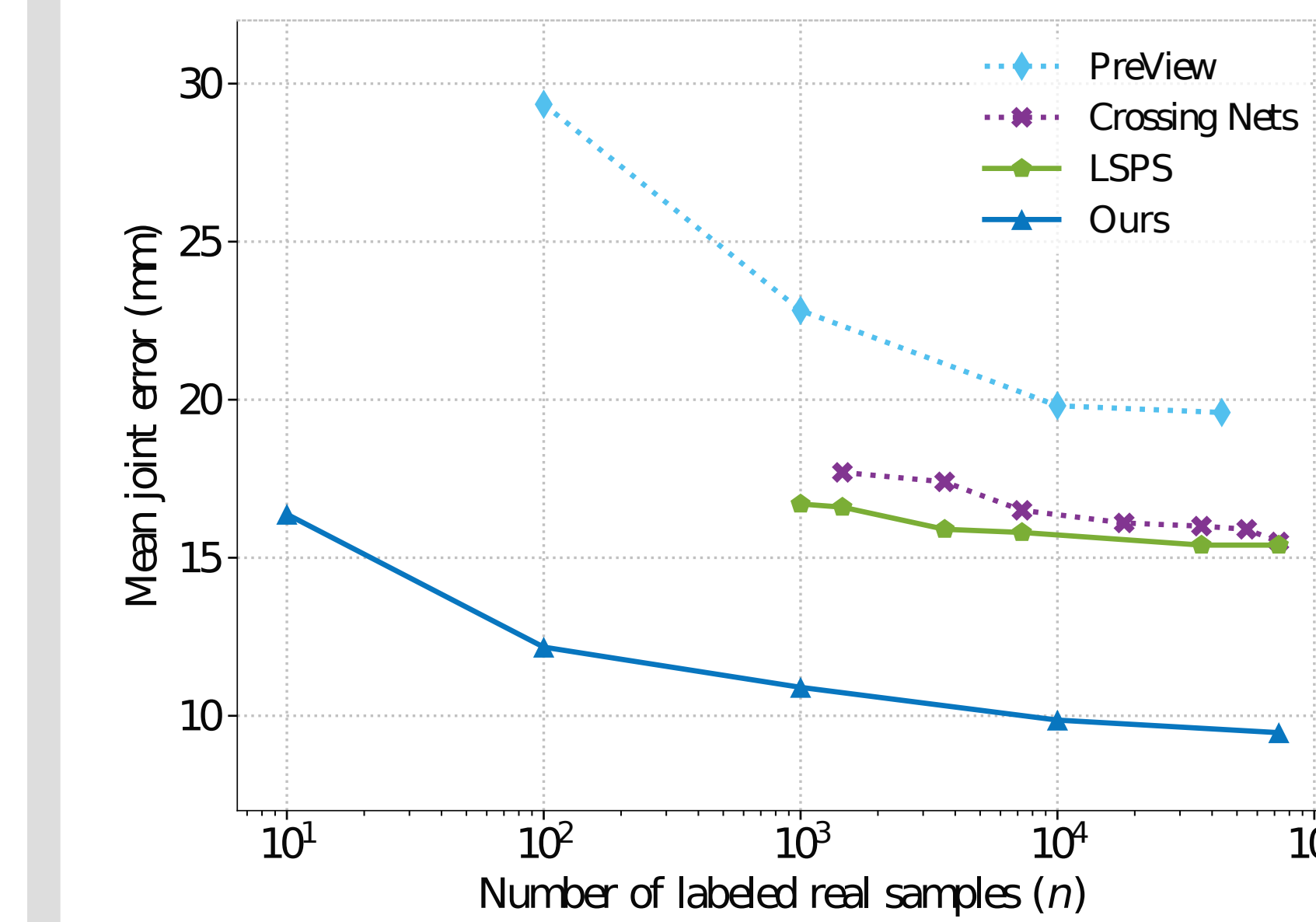
$$\hat{l} = h(z), \quad \hat{l} \in \mathbb{R} \quad \text{Discriminator output: real valued label}$$

$$\ell_h = \frac{1}{2} \sum_{k \in \mathcal{R}} (\hat{l}_k - l_r)^2 + \frac{1}{2} \sum_{k \in \mathcal{S}} (\hat{l}_k - l_s)^2 \quad \text{Discriminator between real and synthetic}$$

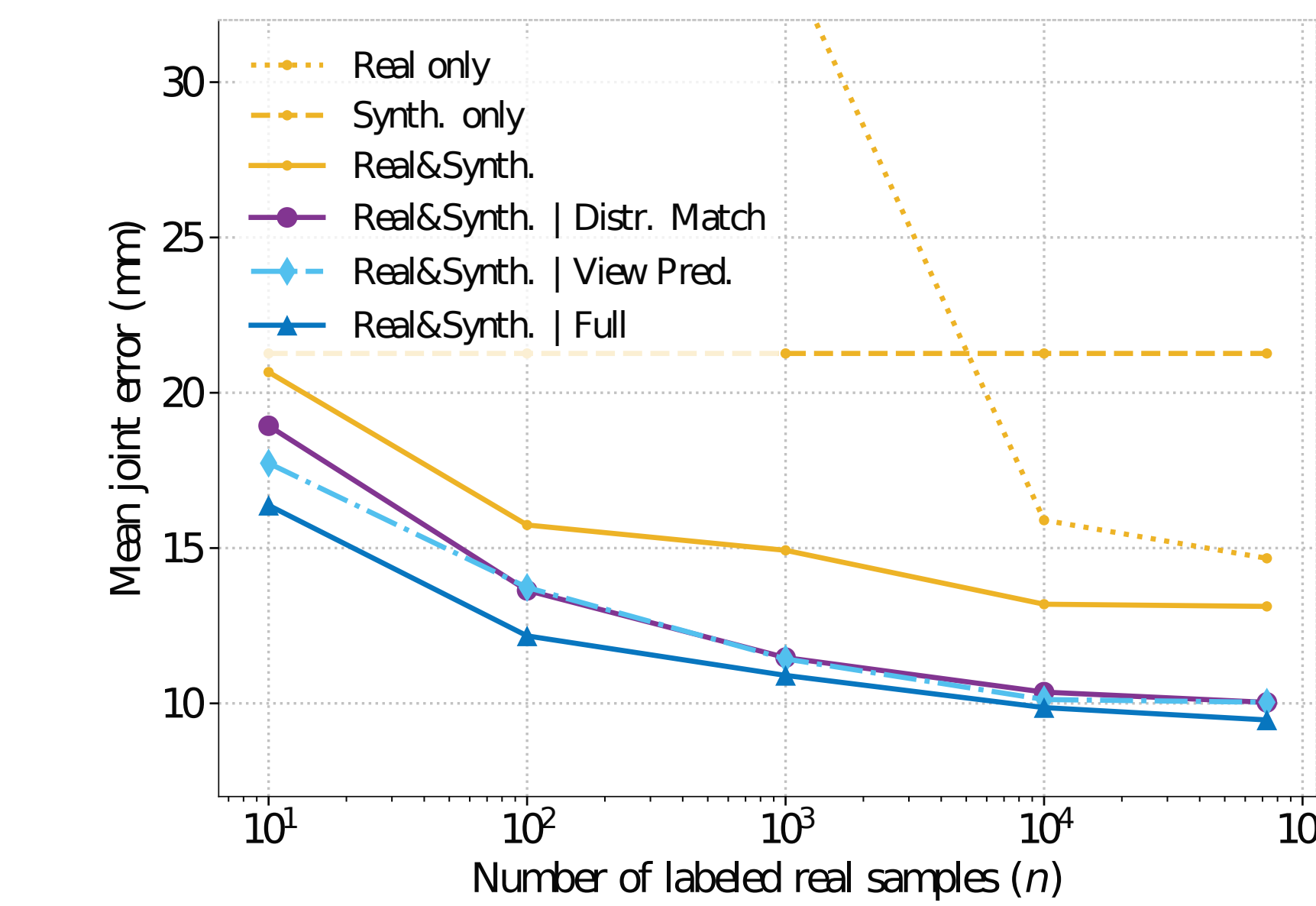
$$\ell_m = \frac{1}{2} \sum_{k \in \mathcal{R}} (\hat{l}_k - l_s)^2 \quad \text{Mapping tries to make real indistinguishable from synthetic}$$

## Findings

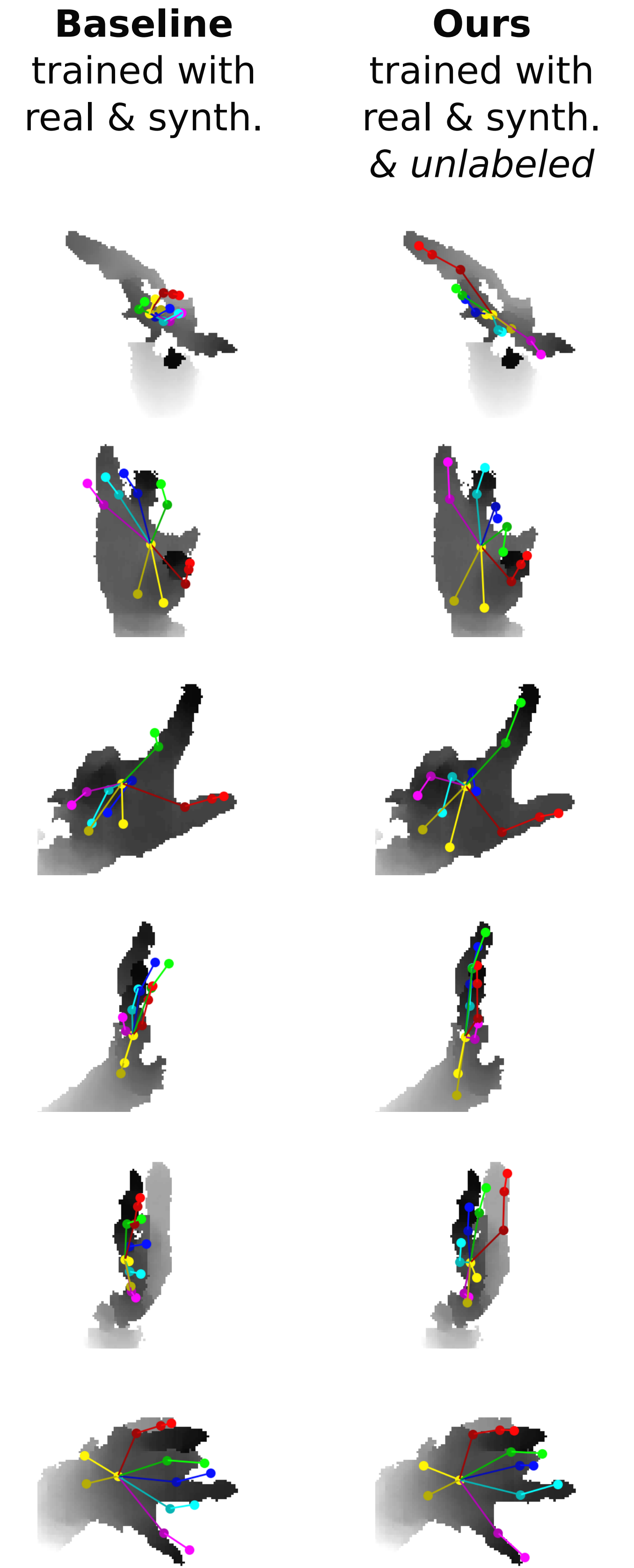
### Comp. to state-of-the-art



### Ablation

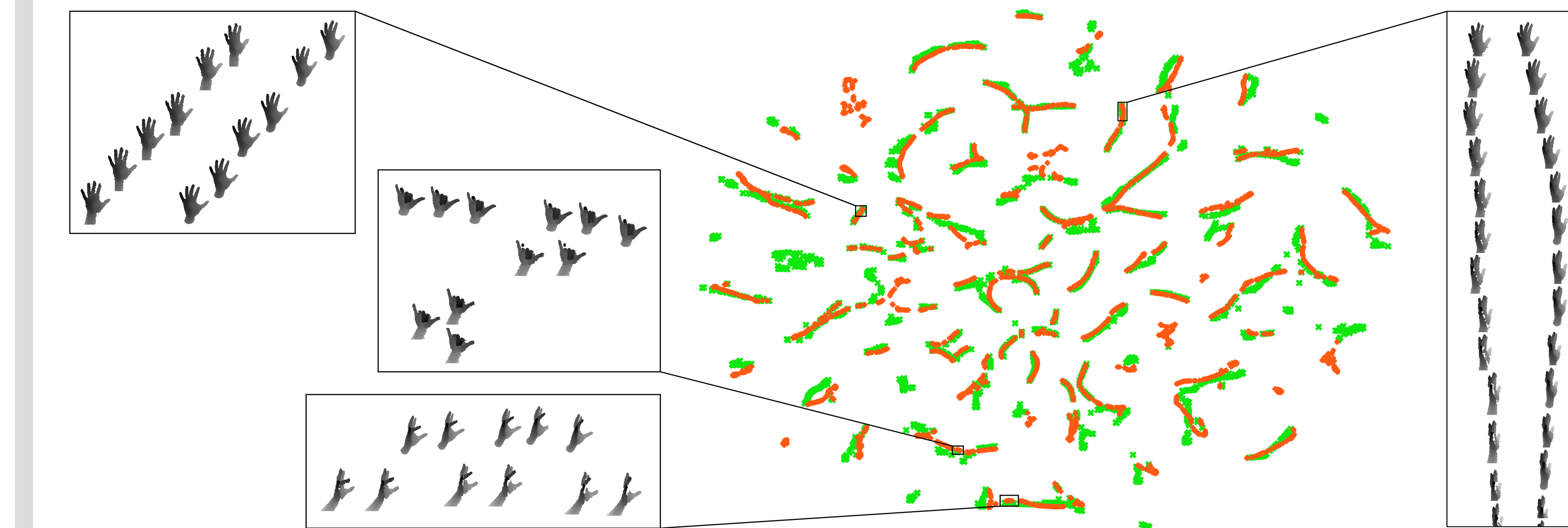


### Qualitative results



### Mapped latent representation

well aligned real and synthetic samples trained with only 100 labeled real & unlabeled & synthetic samples



## References

- [1] F. Massa, B. C. Russell, and M. Aubry. Deep exemplar 2d-3d detection by adapting from real to rendered views. In Proc. CVPR, 2016.
- [2] M. Rad, M. Oberweger, and V. Lepetit. Feature mapping for learning fast and accurate 3d pose inference from synthetic images. In Proc. CVPR, 2018.
- [3] G. Poier, D. Schinagl, and H. Bischof. Learning pose specific representations by predicting different views. In Proc. CVPR, 2018.
- [4] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. P. Smolley. Least squares generative adversarial networks. In Proc. ICCV, 2017.