SemiHand: Semi-supervised Hand Pose Estimation with Consistency

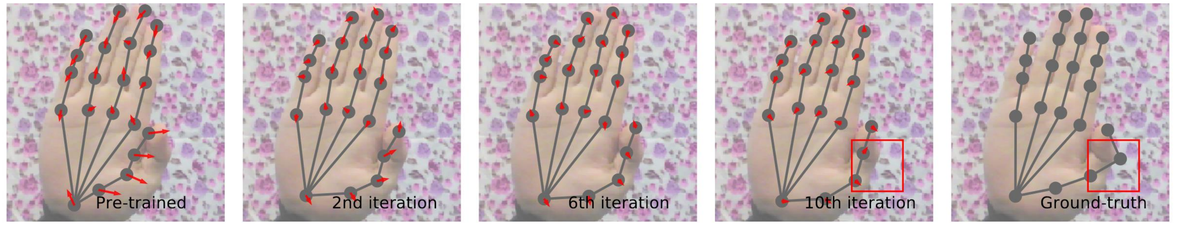
Linlin Yang[1,2], Shicheng Chen[1], Angela Yao[1]
[1]National University of Singapore, [2]University of Bonn

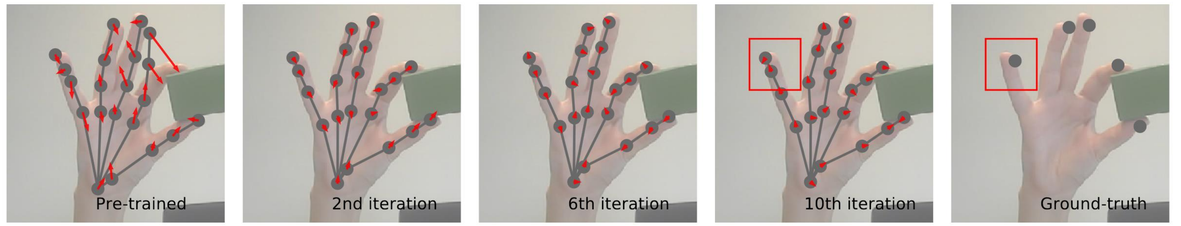
2021 ICCV OCTOBER 11-17 VIRTUAL

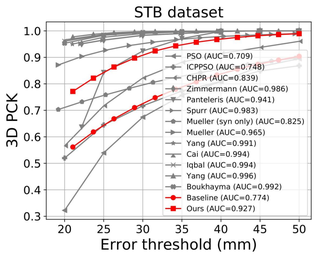
Motivation

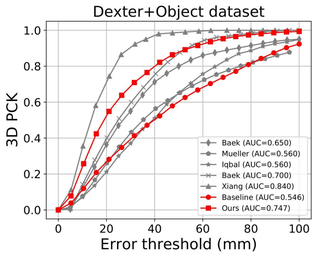
- **Synthesizing training data** is considered an easy alternative to get accurate labels. Yet there exists a significant domain gap between synthetic and real-world images.
- Formulation-wise, pose estimation is **a regression problem** that critically depends on **spatial information**. Moreover, there is a clear separation between **biomechanical feasible** versus infeasible poses.
- We propose a novel RGB-based 3D hand pose estimation framework using labelled synthetic data and unlabelled real-world data; it is the first semi-supervised framework that combines **pseudo-labelling** with **consistency training** for **3D RGB-based hand pose**.

Results

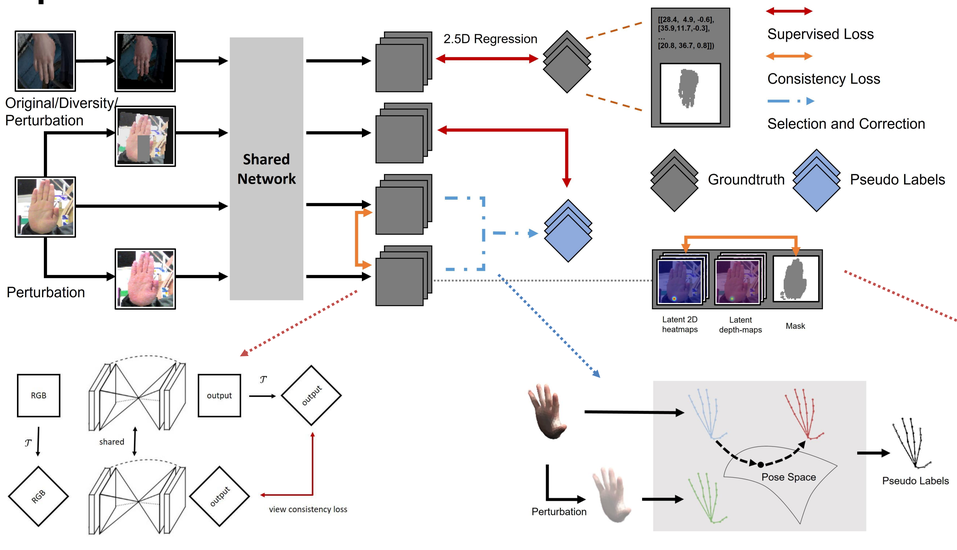


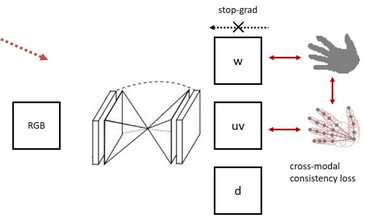






Pipeline





1. Weak-strong augmentation strategy
2. Multi-task learning with consistency for perturbations and auxiliary modalities.
3. Label correction based on the biomechanical feasibility of hand pose.
4. Label confidence based on the feasibility and the consistency.

Algorithm

Require: Pre-trained model θ_0 based on $\mathcal{L}_{sup}$, threshold τ, epoch number K, labelled synthetic data X_L and unlabelled real data X_U

Ensure: Final model θ and pseudo labels $\hat{\mathbf{p}}$

1: Initialize the pseudo-labels $\hat{\mathbf{p}}$ for X_U based on θ_0
2: Initialize the corrected pseudo-labels $\mathbf{r}$ for X_U based on label correction
3: **for** $t = 1, \ldots, K$ epochs **do**
4: Calculate confidence of $\hat{\mathbf{p}}$ based on the plausibility and stability
5: Update θ via gradient ascent based on labels/pseudo-labels and two consistency losses
6: Update $\hat{\mathbf{p}}$ and $\mathbf{r}$ accordingly
7: **end for**