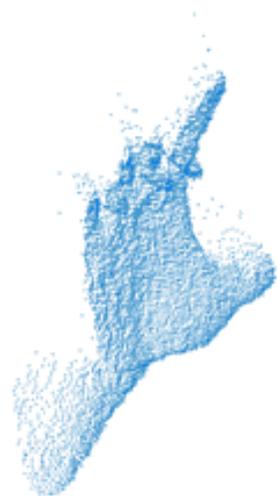


SO-HandNet: Depth-based 3D Hand Pose Estimation using Self-Organizing Network and Semi-supervised Learning



Depth Image



Point Cloud



3D Hand Pose

手势估计方法存在的问题:

1. 基于CNN的手势估计方法虽然取得良好效果，但太过依赖于使用标注数据进行训练，且手势标记过程复杂。
2. 如何高效地表达和处理三维数据?

我们的解决思路:

1. 采用半监督的训练策略：除了使用带标注的数据，使用未标注的数据来增强训练；
2. 从深度图像中恢复出点云，使用点云自编码器（Point Cloud Autoencoder）处理点云数据。

SO-HandNet – 概述

- 流程：
 - 深度图像 → 三维点云；
 - 编码器（Encoder）：提取特征；
 - 解码器（Decoder）：从特征中恢复三维点云； *[仅训练阶段]*
 - 估计器（Estimator）：回归三维手势。
- 半监督训练策略：
 - 不带标签的数据：训练自编码器部分（Encoder, decoder）；
 - 带手势标签的数据：优化整个网络。

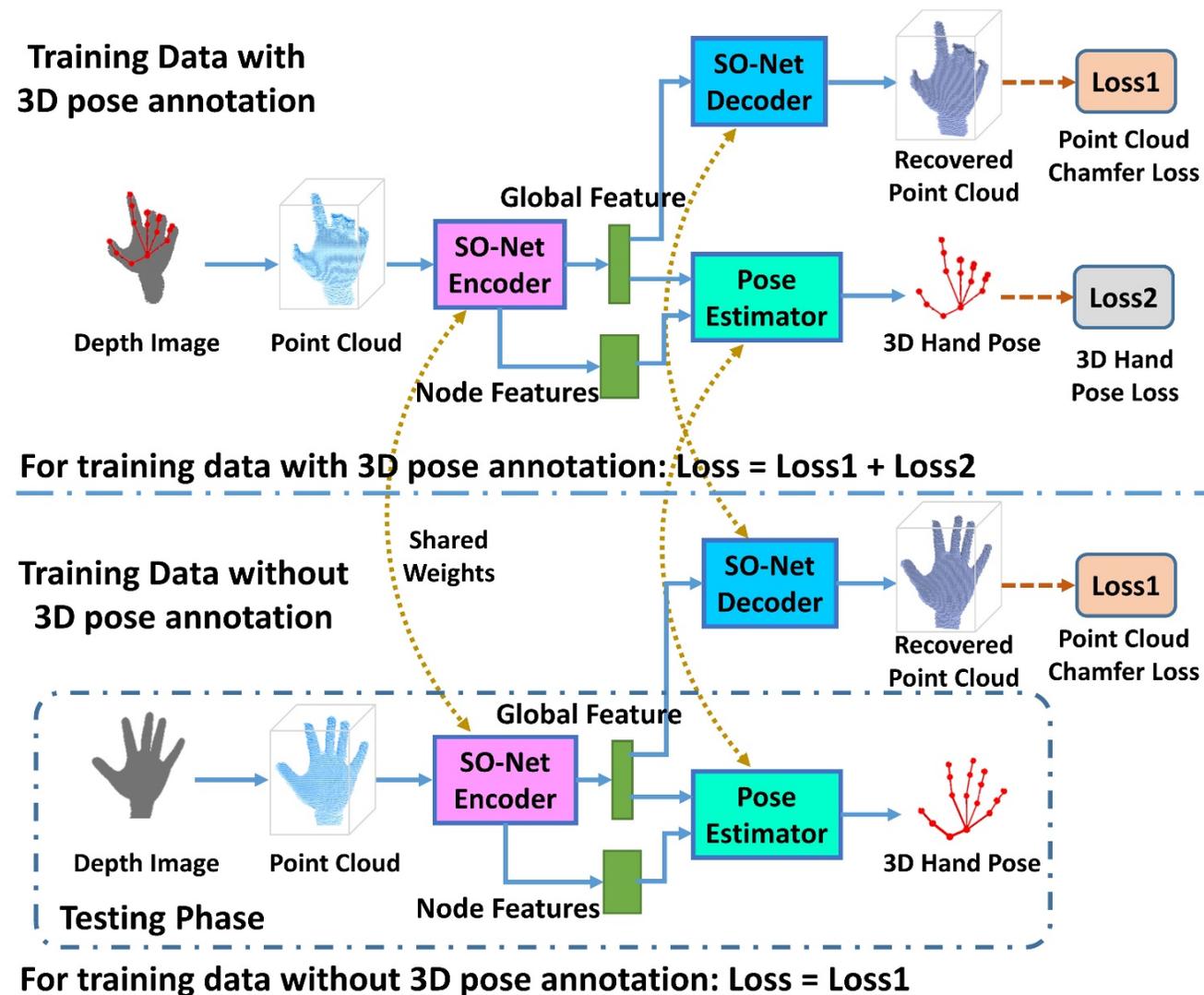


图1. SO-HandNet流程图

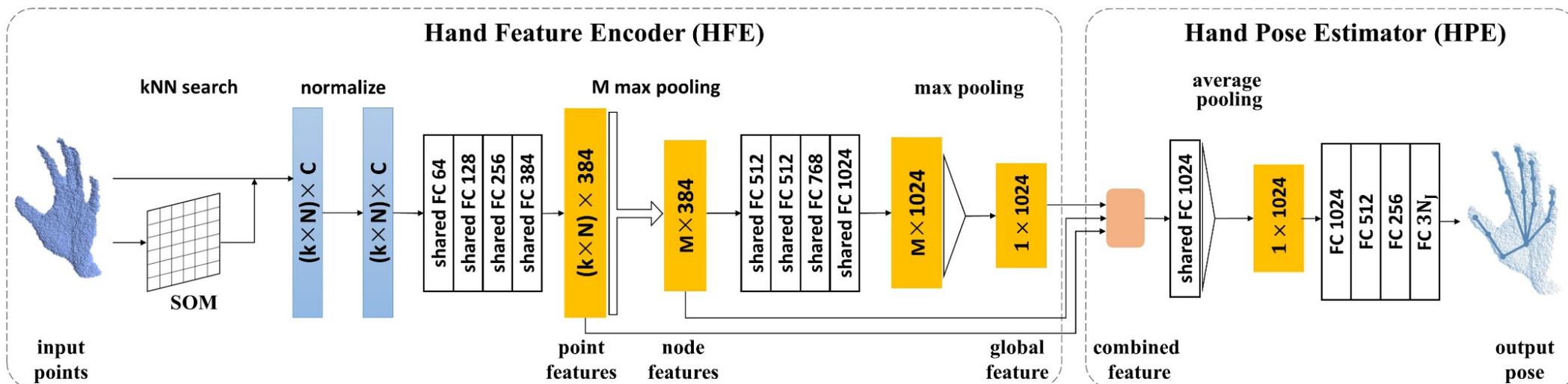


图3. SO-HandNet结构图（编码器和估计器）

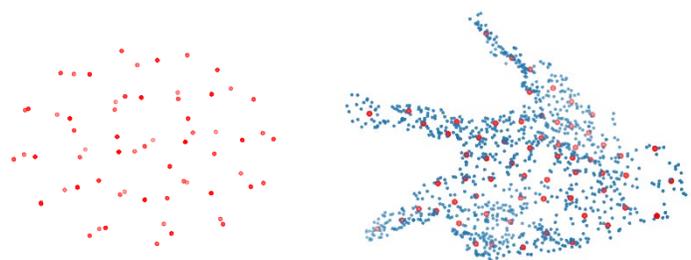


图2. 初始化SOM（Self-organizing map）与非监督训练后的SOM.

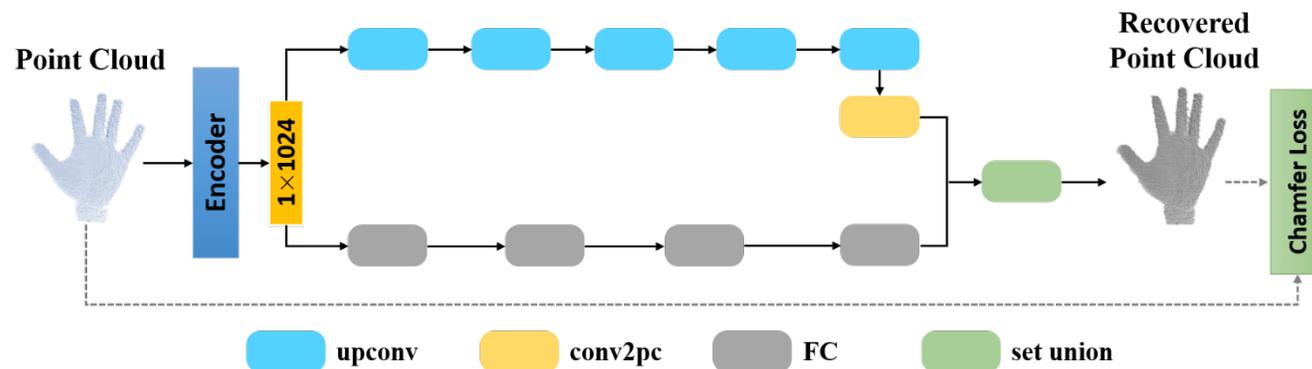


图4. 解码器结构

自比较:

- 特征融合策略的比较;
- 半监督策略的效果验证。

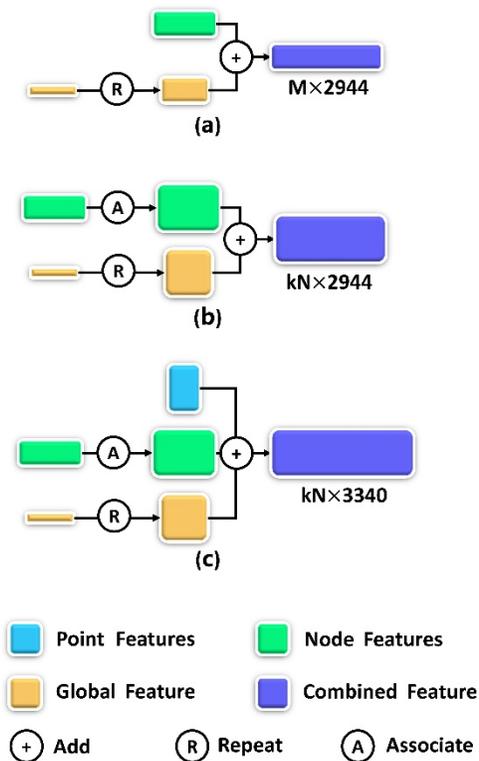


图5. 三种不同的特征融合方法

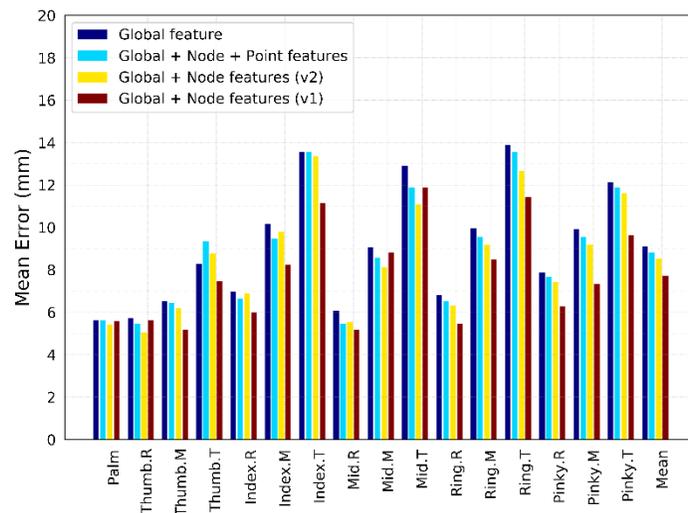


图6. 不同融合方法的各关节点误差比较

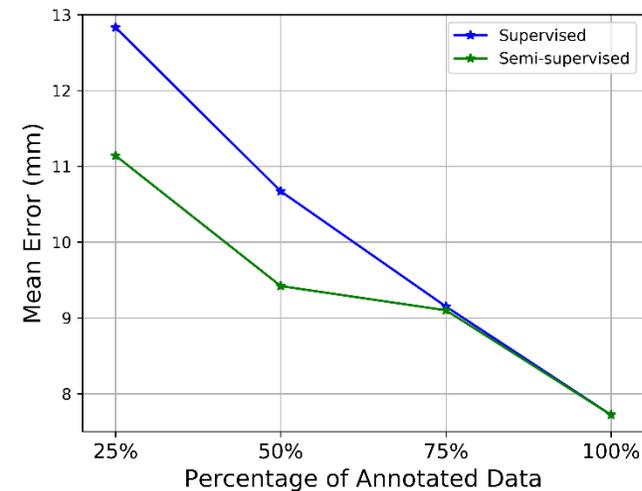


图7. 半监督方法与使用同等标注数据的全监督方法的平均误差比较

SO-HandNet – ICCV2019



Yujing Chen, Zhigang Tu*, *et. al.* SO-HandNet: Self-Organizing Network for 3D Hand Pose Estimation with Semi-supervised Learning. ICCV2019.

计算机视觉&人工智能顶级会议;
在多个3D手势识别公开数据库(NYU, MSRA, ICVL)实现2019年7月精度第一

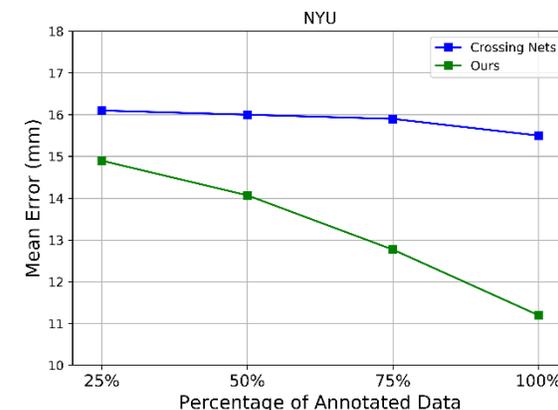
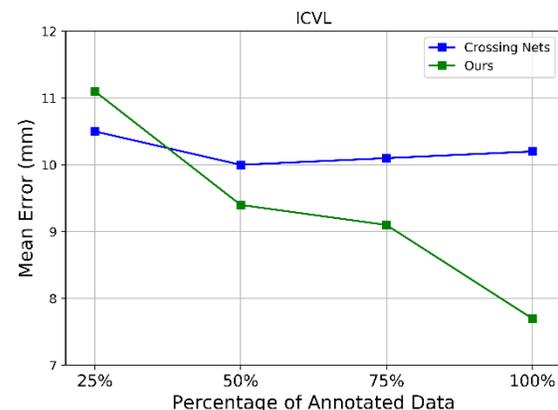


图1. ICVL和NYU数据集对比

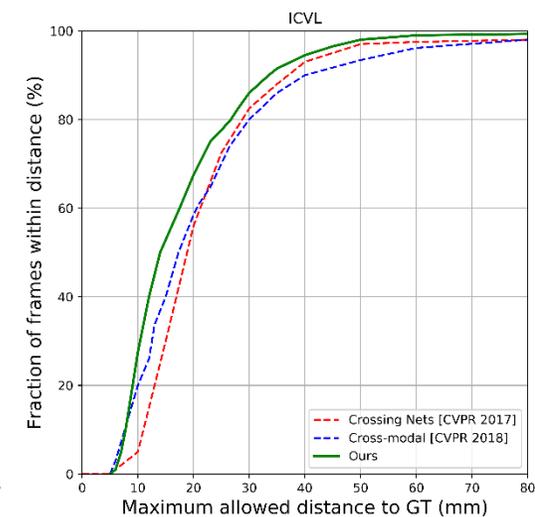
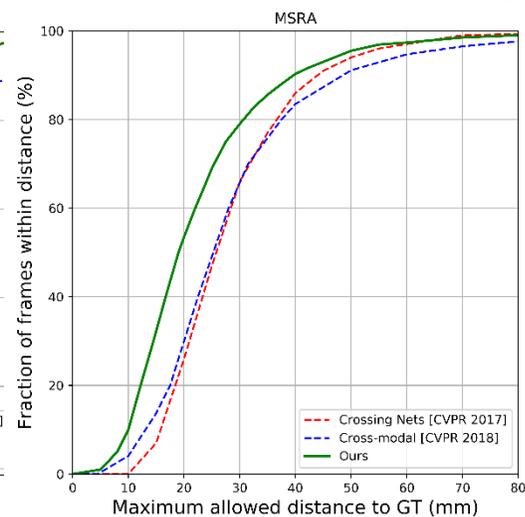
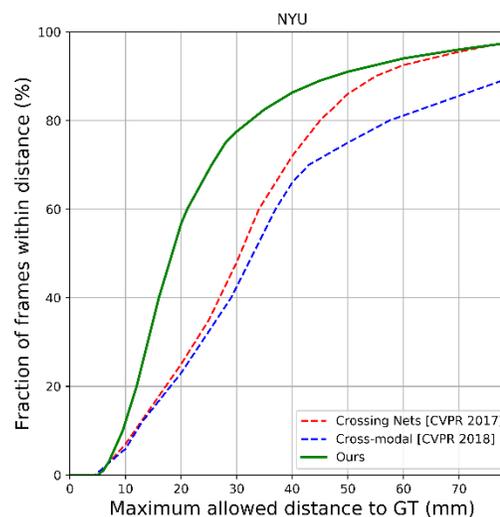


图2. SOTA数据集半监督对比