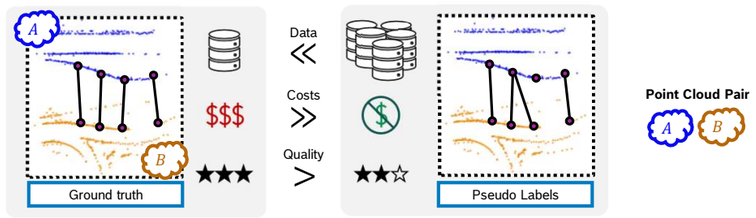
Unsupervised Point Cloud Registration (PCR) with Self-Distillation

Christian Löwens[1,2], Thorben Funke[1], André Wagner[1], Alexandru Paul Condurache[2,3]

(1) Bosch Research (2) University of Lübeck (3) Bosch, Automated Driving

BMVC 2024

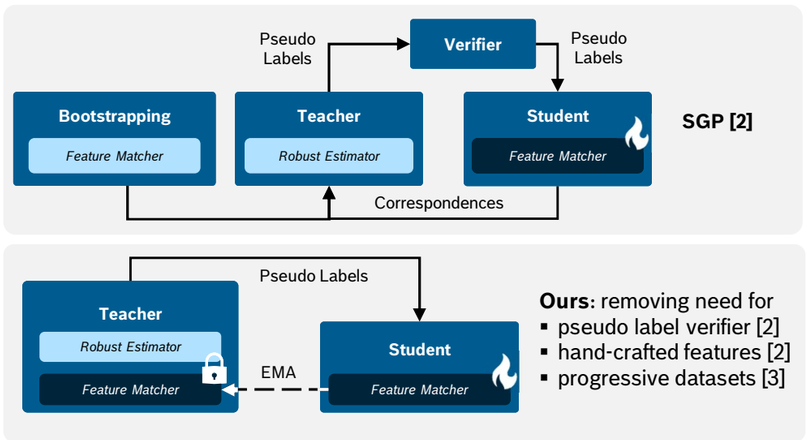
1. Motivation for Unsupervised PCR



- Collection of ground truth poses is **costly** and limited in size in the automotive context
- Crowd-sourced data from consumer-grade cars contains **orders of magnitude more unlabeled data**
- **Common unsupervised algorithms** rely on non-trivial initial **hand-crafted features or progressive datasets** to bootstrap the learning process

2. Contribution

1. **Simplify** unsupervised PCR compared to related work:



2. We show that common data augmentation can impede the bootstrapping phase
3. Generalizable: Experiments on **RGB-D** and **automotive radar point clouds**

References & Code

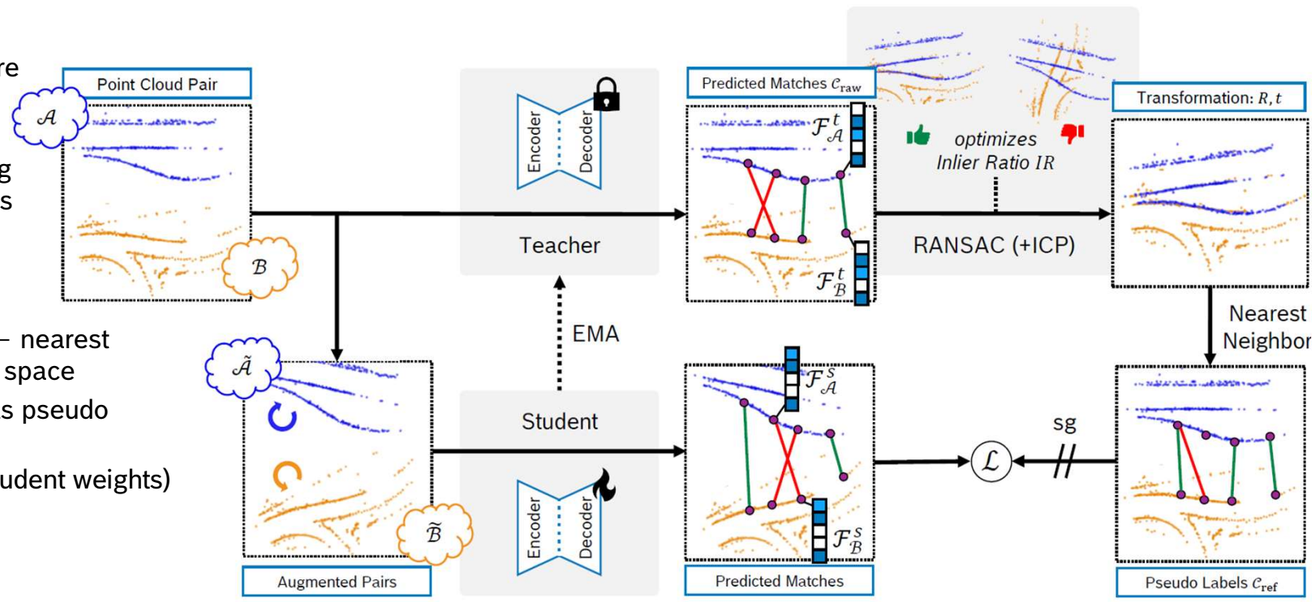
[1] Choy et al., 2019, Fully convolutional geometric features
[2] Yang et al., 2021, Self-supervised geometric perception
[3] Liu et al., 2024, Extend your own correspondences: ...

github.com/boschresearch/direg

3. Method

Self-distillation algorithm:

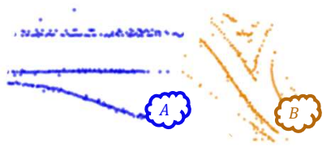
- Both student and teacher are FCGF feature extractors [1]
- Initial correspondences ← nearest neighbors among the extracted feature vectors
- RANSAC estimates a final transformation to align both clouds
- Refined correspondences ← nearest neighbors in the coordinate space
- Refined correspondences as pseudo labels to train student
- Teacher weights ← EMA(student weights)



4. Results

Automotive Radar

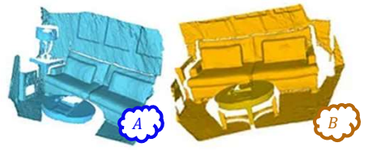
- private dataset
- frames from different drives
- long-range consumer-grade radar



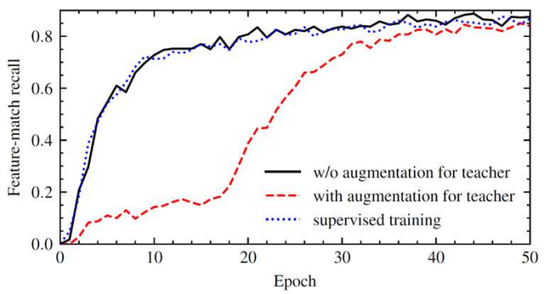
Method	RR (%) ↑	RTE (cm) ↓	RRE (°) ↓
Supervised [1]	**96.6**	**0.355**	0.160
SGP (Unsupervised) [2]	90.2	0.596	0.219
EYOC (Unsupervised) [3]	91.7	0.487	0.166
Ours, Teacher = Student	96.1	0.390	0.181
Ours, Teacher = EMA(Student)	95.8	0.413	**0.156**

3DMatch

- public PCR benchmark
- indoor scenes
- RGB-D dataset



Method	FMR (%) ↑	RR (%) ↑	IR (%) ↑
Supervised [1]	**93.5**	**92.0**	**24.3**
SGP (Unsupervised) [2]	91.3	90.8	22.4
EYOC (Unsupervised) [3]	62.5	76.8	10.6
Ours, Teacher = Student	92.3	91.1	22.4
Ours, Teacher = EMA(Student)	92.7	91.6	24.1



No augmentation for teacher

- A common data augmentation technique is to randomly rotate both point clouds to force the network to become rotation invariant [1]
- It makes the registration problem harder for the teacher during the bootstrap phase and makes training unstable
- By keeping the augmentation only on student side, the training becomes more stable without loosing the rotation invariance