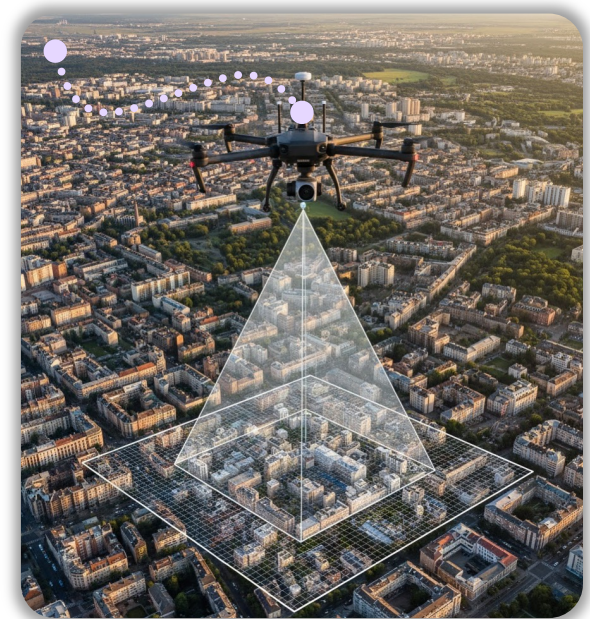




Motivation

Why UAV camera localization & calibration?



Large-scale mapping & reconstruction



Inspection



Search & rescue in GPS-denied environments

But existing localization and calibration often rely on **dense 3D models** or **large databases**.

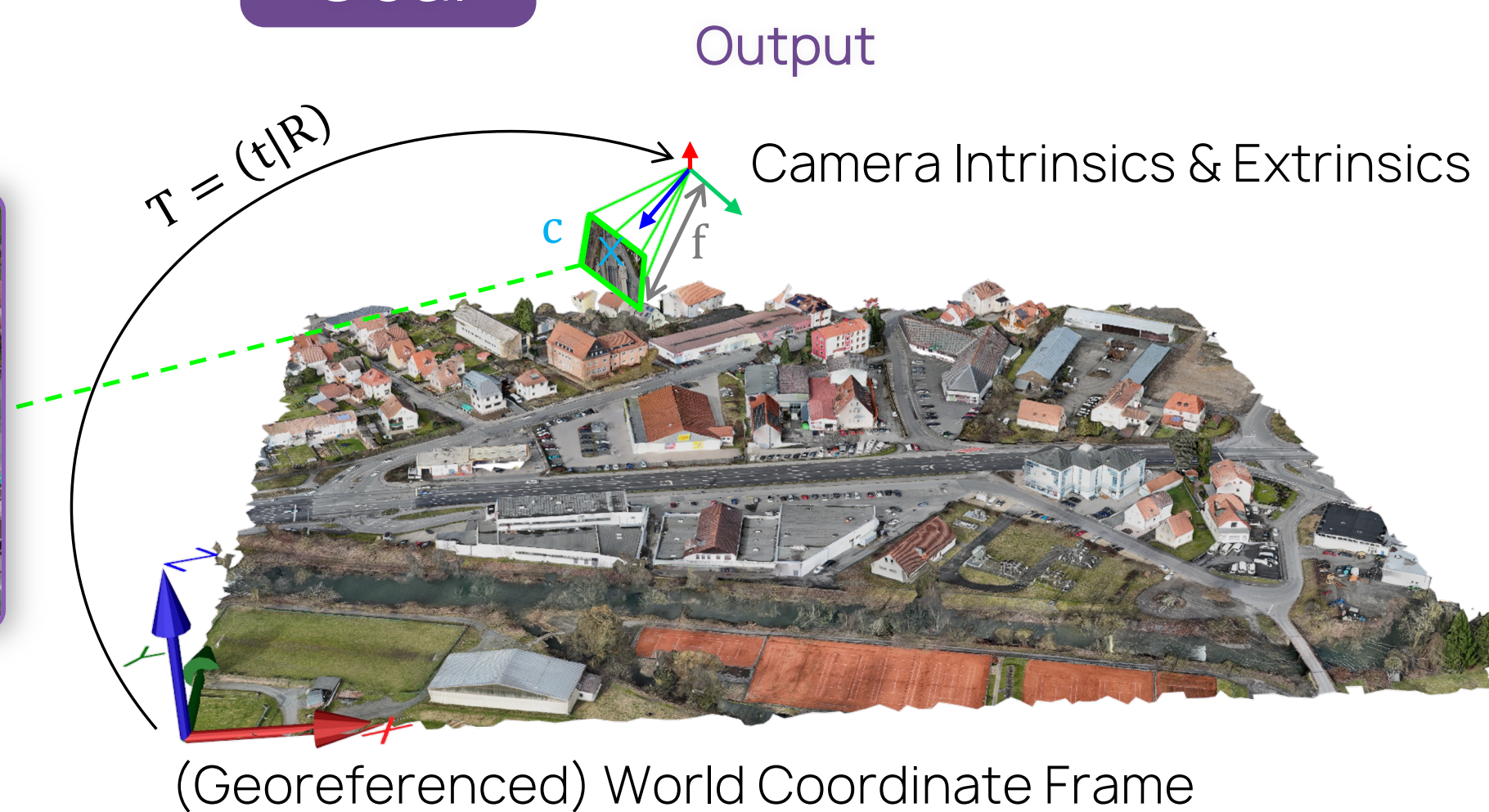
Why **orthographic geodata**?

Widely available orthophotos and 2.5D digital surface models offer a **lightweight alternative**.

Goal



UAV Image



Previous Work

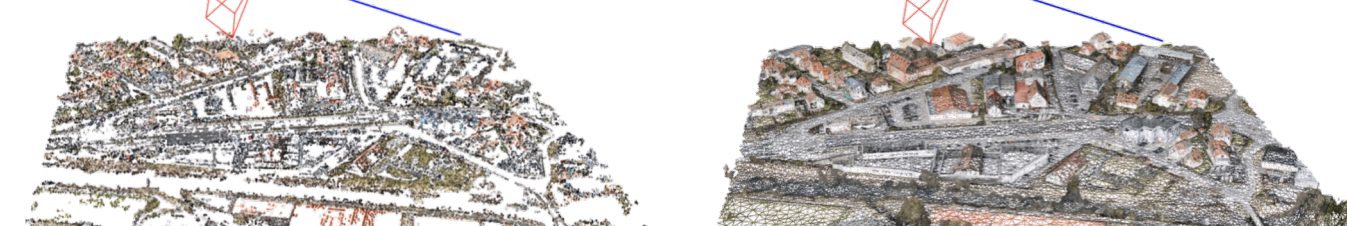
Image Retrieval



Match posed images
e.g. ComplexUAV^[1]

- × Sensitive to lighting
- × Viewpoint-dependent
- × Requires dense coverage

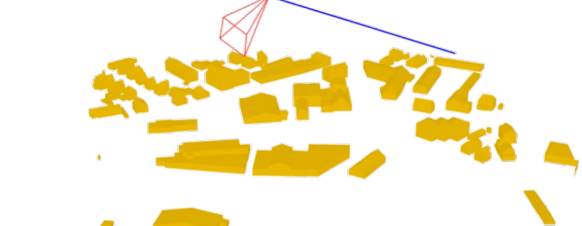
3D Models



SfM Point Cloud, LiDAR, (textured) meshes
e.g. UAVD4L^[2], CrossLoc^[3]

- × Expensive data acquisition
- × Large data size
- × Difficult real-time matching

Abstract Models



Level-of-Detail (LoD)
e.g. LoD-Loc^[4]

- × Requires strong priors
- × Loss of detail
- × Ambiguous in repeats

TL;DR

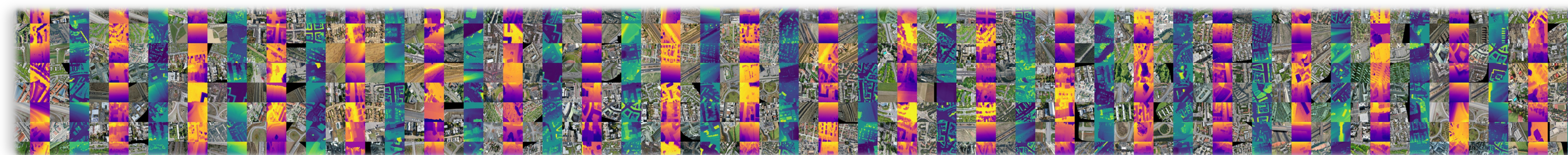
OrthoLoC is a dataset and plug-and-play framework for UAV 6-DoF camera localization and calibration using readily available orthophotos and digital surface models

4 Gaps → 4 Main Contributions

- (1) No UAV ↔ orthographic geodata dataset with full GT
- (2) Calibration under-explore
- (3) Cross-view: perspective (oblique) ≠ orthographic (nadir)
- (4) Cross-domain: Different times → appearance shifts

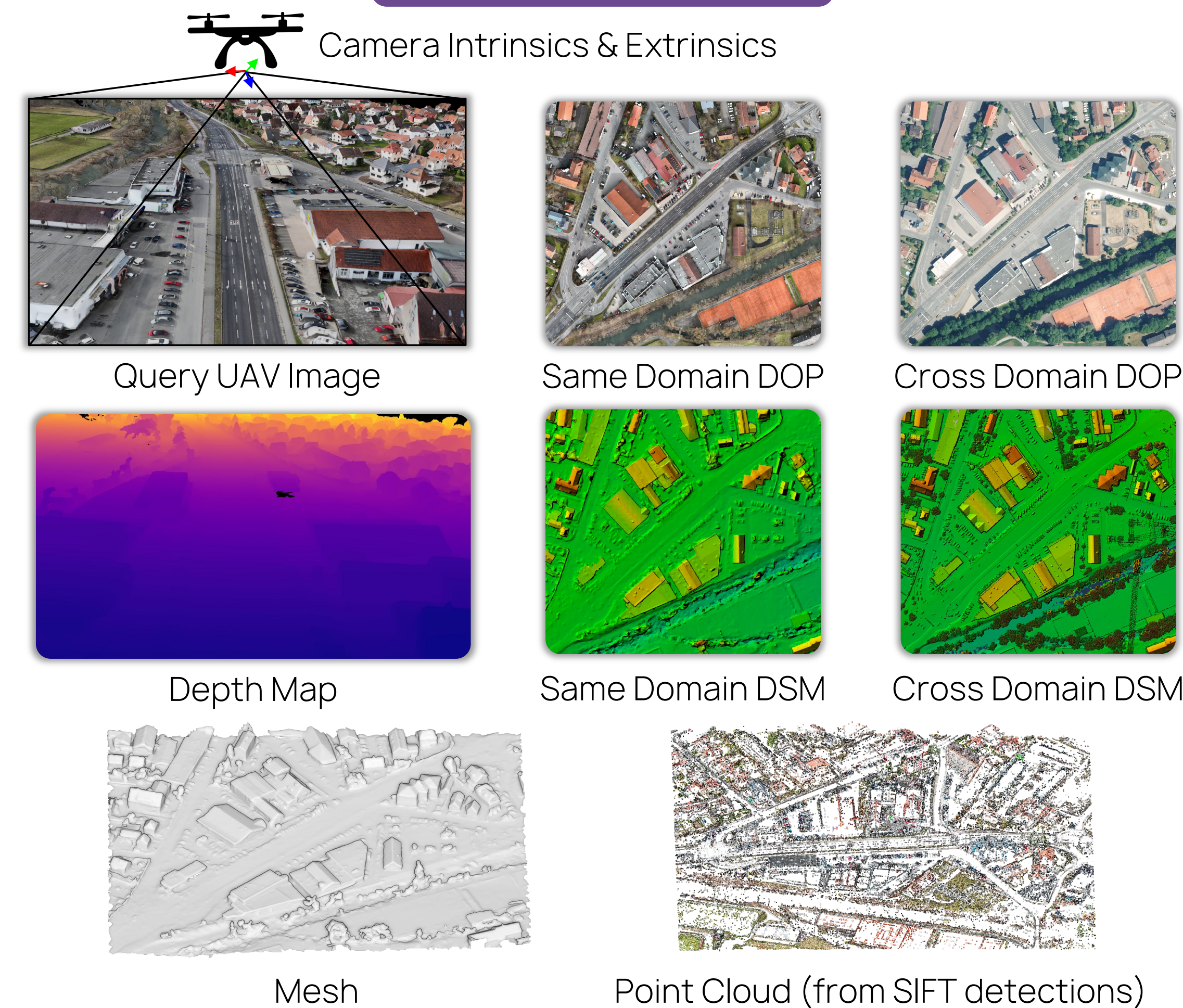
- OrthoLoC dataset
- Joint localization & calibration framework
- Adaptive Homography Preconditioning (AdHoP)
- Benchmark across domain shift

The OrthoLoC Dataset at a Glance



- 16,425 UAV images
- 47 locations
- Germany and the U.S.
- centimeter-accuracy
- temporal gaps
- diverse altitudes
- diverse viewpoints
- paired structure
- plug-and-play ready

Dataset Modalities



Domain Shift Analysis

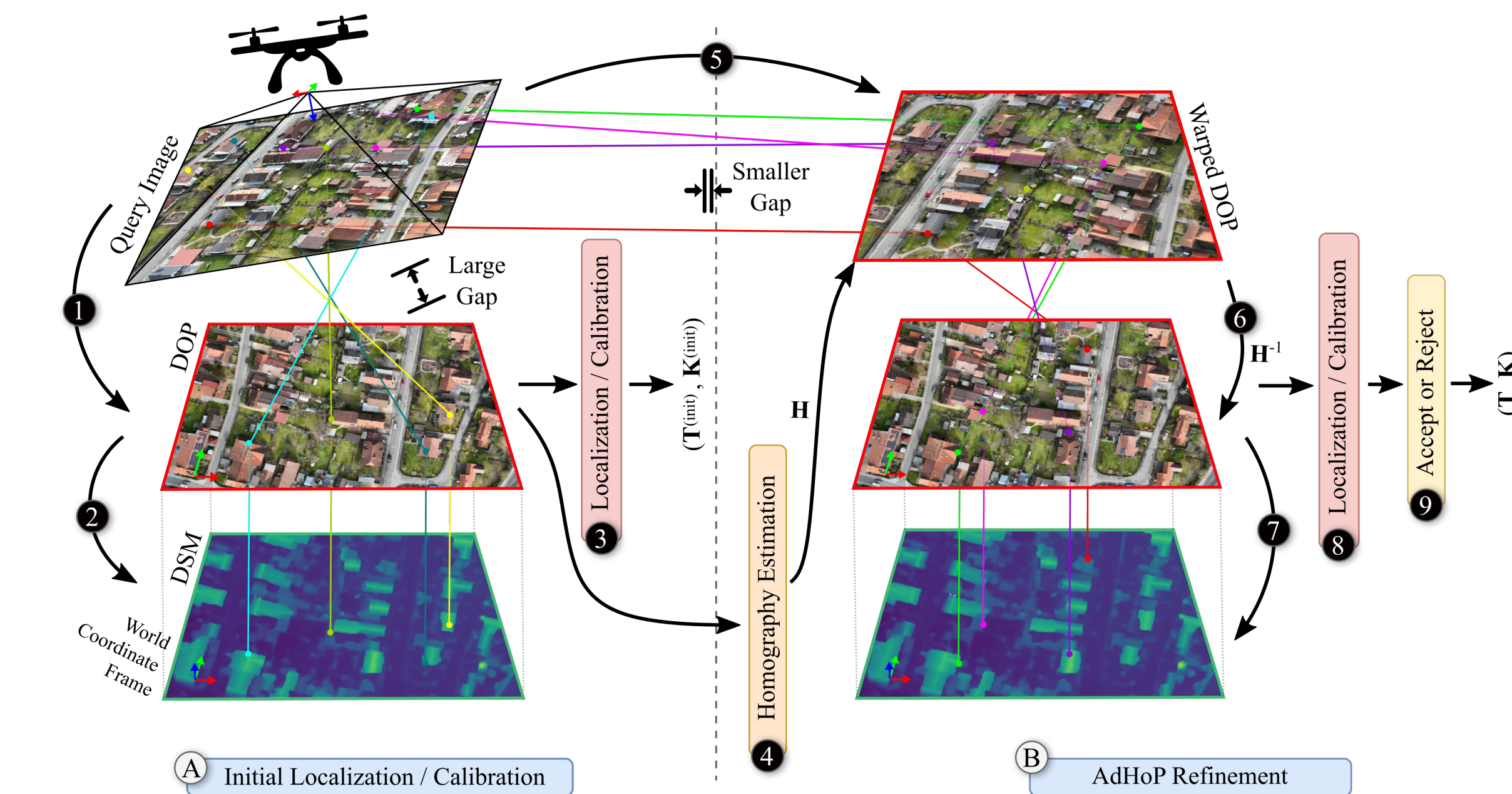


→ **Temporal (visual and geometric) discrepancies** are key **bottlenecks** for cross-domain localization

Takeaways

- OrthoLoC: dataset and plug-and-play framework for UAV localization & calibration from orthophotos and digital surface models
- AdHoP bridges the perspective ↔ orthographic gap
- Temporal domain shift challenges existing matchers

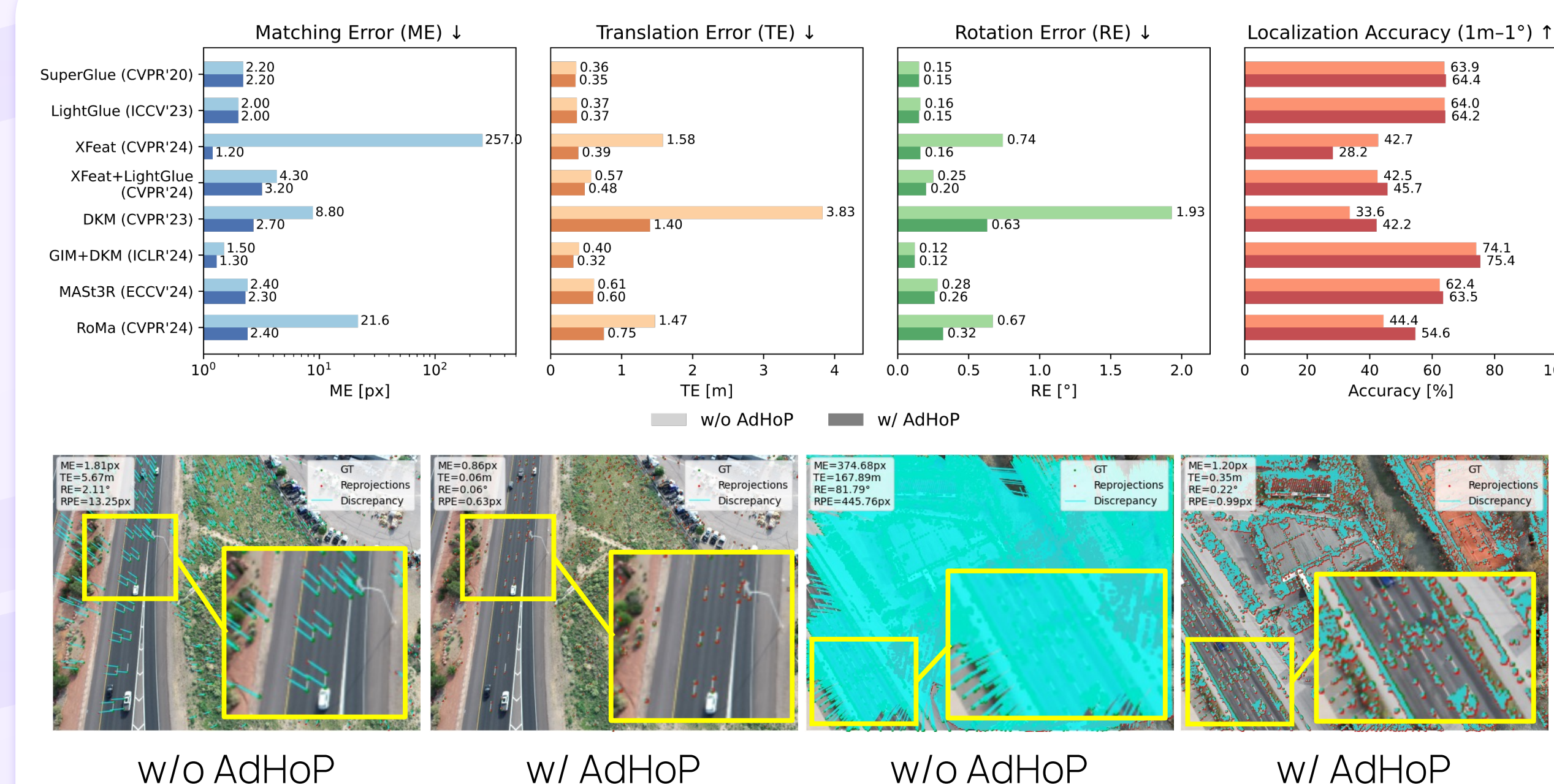
Benchmarking Framework



- (1) Feature Matching
- (2) 3D Lifting
- (3) Initial Estimation
- (4) Homography Estimation
- (5) Re-Matching
- (6) Back-Projection
- (7) 3D Lifting
- (8) Re-Estimation
- (9) Validation



6-DoF Localization Benchmarking



→ AdHoP improves pose precision by up to 60 % (**centimeter-level accuracy**)