Can synthetic drone data replace real flights?

Literature and resources review, and the first experiment we propose

Data-efficient UAV–satellite geo-localization through synthetic-to-real training · 17 November 2026

The question we want to answer

Can synthetic UAV–satellite pairs replace or supplement part of the real training locations, while Recall@1 on real regions never seen in training stays within an agreed margin δ?

1 Replacement

Swap a share of real locations for synthetic ones, keeping the total fixed.

2 Augmentation

Keep 75% of real locations and add synthetic ones on top.

3 Counting unit

Unique locations, not images, so the real budget is measured fairly.

What prior work already tells us

92.7%

R@1

Sample4Geo (ICCV 2023)

A plain contrastive model with hard negatives reaches the top of University-1652 drone→satellite retrieval.

24.9 vs 18.8

zero-shot R@1

Game4Loc (AAAI 2025)

Pre-training on synthetic GTA-UAV transfers to real UAV-VisLoc better than pre-training on real DenseUAV.

85.6%

R@1 with 2% pairs

CDIKTNet (2025)

Paired data needs can drop sharply, here with unpaired real images rather than synthetic ones.

The gap our experiments fill

Game4Loc	Fine-tunes on all real data and tests on the same areas it trained on.
CDIKTNet	Varies the amount of paired data, but never uses synthetic data or a region-disjoint test.
Sample4Geo	Assumes full real supervision and one positive per query.

Missing so far

A controlled curve of real training locations against Recall@1 on unseen real regions, with synthetic data used to replace or to augment the real set.

Baseline: Sample4Geo, with a few targeted changes

Keep as is

- Shared encoder for drone and satellite views
- Symmetric InfoNCE, positives only
- GPS and similarity hard negatives
- Lighter backbone: ConvNeXt-Nano, Tiny as fallback

What we add, mostly ported from Game4Loc

1	**Loaders**	GTA-UAV LR and UAV-VisLoc dataset classes
2	**Multi-positive**	Several valid satellite tiles per drone image
3	**Batch sampler**	No two matches of one query in a batch (MES)
4	**Evaluation**	Hit = any positive tile in top-K; also Dis@1 in metres
5	**Data config**	Build training sets by unique locations and real/synthetic mix

Datasets: one synthetic source, one real source

GTA-UAV LR

Synthetic · GTA V

33,763 drone images, 14,640 satellite tiles, 80–650 m altitude, IoU-graded positive tiles

Watch out: Game-world look, one fictional region, low resolution

UAV-VisLoc

Real · 11 sites in China

6,742 nadir drone images, 0.3 m satellite maps, GPS, altitude and heading

Watch out: Few regions, no published pairing, height above sea level in hilly sites

Worth discussing: UAV-VisLoc v1.1 (Zenodo, Nov 2026) has 21 real flights plus 9 synthetic ones and ships ground truth. More real regions means more stable test folds.

How we test: regions the model has never seen

1 Group by region

Every VisLoc site stays whole in one split

2 Balance terrain

StratifiedGroupKFold keeps terrain mix similar across folds

3 Average over folds

Report the mean and spread, not one lucky split

Primary metric

Recall@1 on held-out real regions. Also Recall@5, AP and Dis@1 in metres.

Unique location

In UAV-VisLoc, one satellite tile with at least one drone image at IoU ≥ 0.39. The GTA-UAV definition is being finalised.

First experiment, in four steps

0

Prepare the real data

Height above ground for hilly sites, multi-scale tiles, validation and test sets

1

B100 reference

All real training locations, 3 seeds. Gives the baseline and its noise

2

Synthetic only

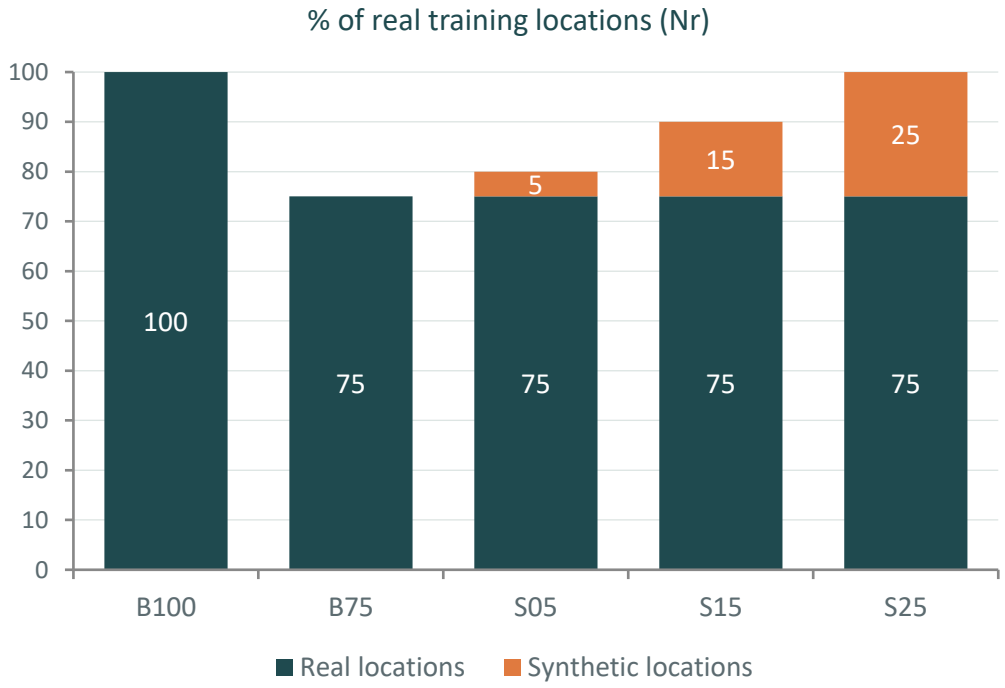
Train on GTA-UAV LR, test zero-shot on the same real regions

3

H1 / H2 pilot

B75, S05, S15, S25, 3 seeds each, same test regions

The pilot runs and what they test



H1 · Fewer real locations

Does B75 stay within δ of B100?

H2 · Synthetic fills the gap

Do S05, S15 and S25 recover what B75 loses, and how much synthetic data is enough?

Open questions

- δ for H1: proposed at least twice the seed standard deviation, set after the B100 runs
- Stay on UAV-VisLoc 2024 or switch to v1.1 with 21 real regions
- Drone images below IoU 0.39: drop, lower the threshold, or keep as semi-positives
- Number of folds, and whether the pilot starts on one fixed fold
- A fixed training budget (epochs or steps) for every run