Data-Efficient UAV–Satellite Geo-Localization

Synthetic-to-Real Training

Supervisor: Roman Ilechko
Team:
Grzegorz Jeziorowski (SUT), Svyatoslav Borysyuk (LPNU), Rafał Szafraniec (SUT)

UAV visual geo-localization

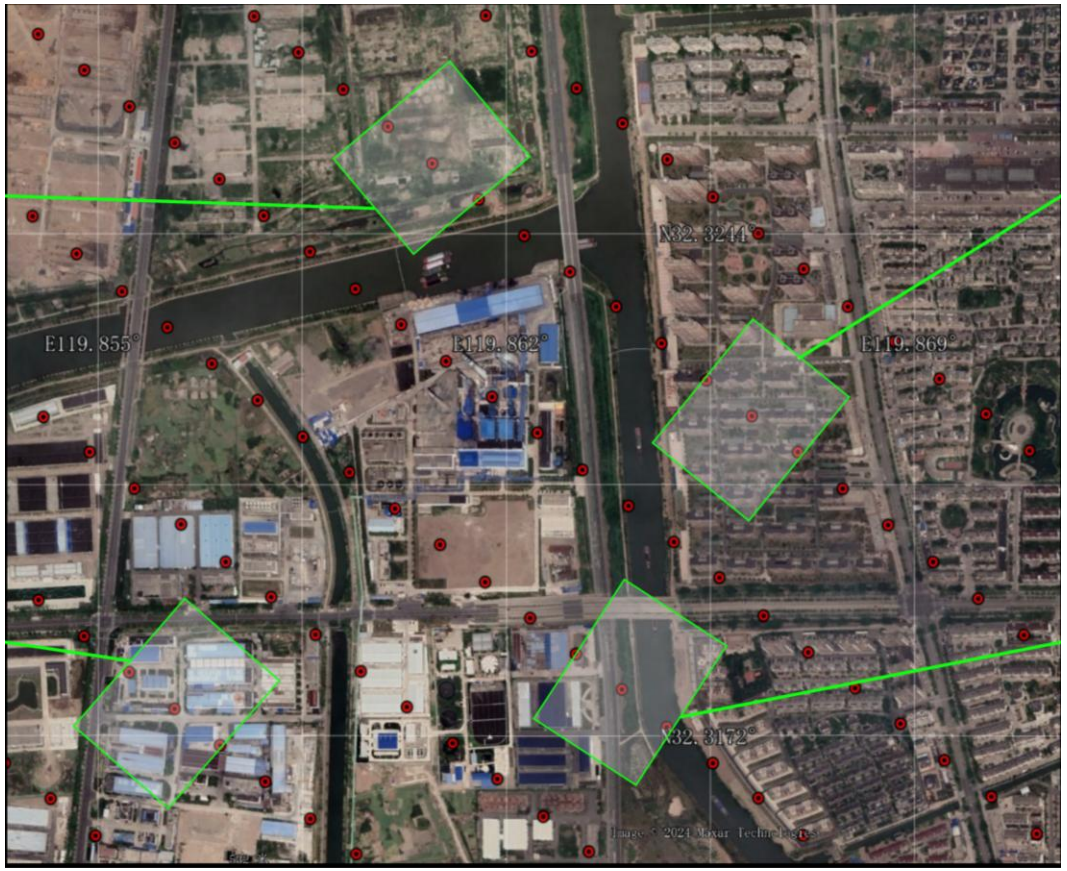
Match the scene observed by a UAV to a georeferenced map.

Real UAV query



The query image shows the scene whose location we want to estimate.

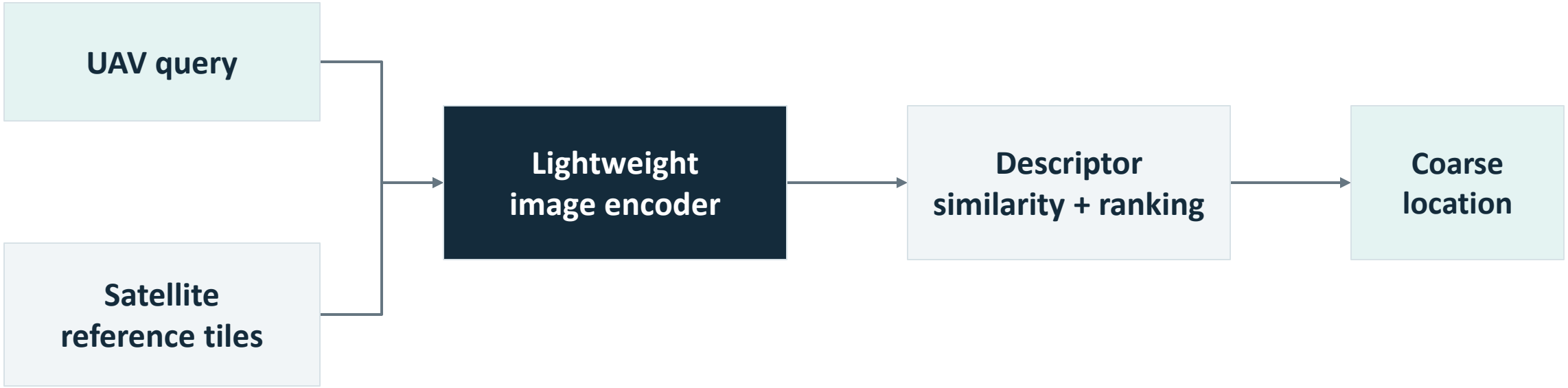
Georeferenced reference map



Source: UAV-VisLoc, Fig. 2

Image retrieval pipeline

Both views use a shared encoder to produce comparable image descriptors.



Satellite descriptors can be precomputed. Each UAV image requires one encoder pass.

The best matching reference tile supplies approximate coordinates.

The real-data bottleneck

Synthetic imagery can expand training coverage, but transfer to real flights needs testing.

Aspect	Real UAV imagery	Synthetic UAV imagery
Acquisition	Flights, location metadata and quality checks	Rendered images with generated annotations
Coverage	Limited by accessible locations and capture conditions	Controllable viewpoints within a virtual environment
Main challenge	Cost of expanding real training coverage	Appearance and geometry can differ from real flights

Research motivation

Reduce the real UAV training data required for reliable retrieval.

Research question

Starting point

Game4Loc already tests GTA-UAV pretraining followed by real-data fine-tuning.
Our study will compare replacement and augmentation at the level of unique locations.

Under a fixed training budget, how does replacing or supplementing real training locations with synthetic locations affect localization performance in unseen real regions?

Target evidence: fewer real training locations and gains from synthetic augmentation.

Proposed data sources

GTA-UAV provides synthetic training pairs. UAV-VisLoc provides real observations.

Source	UAV imagery	Reference imagery	Planned use
GTA-UAV LR	33,763 simulated UAV images	14,640 game-world overhead tiles	Synthetic training
UAV-VisLoc 2024	6,742 real UAV images	11 georeferenced regional maps	Real training, validation and test

Pair preparation

Retain each source's own matches. Adapt the existing Game4Loc tiling and pairing tools for real maps.

Training dataset composition

Replacement keeps the total fixed. Augmentation keeps the real subset fixed.

N_r = 1,000 unique real training locations (illustrative example)

Each location groups UAV observations with their matching reference imagery.

Replacement

Keep α × 1,000 real locations. Replace the remaining locations with synthetic ones.

$$N_{\text{total}} = \alpha N_r + (1-\alpha)N_r = N_r$$

Real	Synthetic	Total
1,000	0	1,000
750	250	1,000
500	500	1,000

Can synthetic data preserve performance when fewer real locations are available?

Augmentation

Keep for example the same 750 real locations. Add β × 1,000 synthetic locations. Total grows.

$$N_{\text{total}} = 0.75N_r + \beta N_r = (0.75+\beta)N_r$$

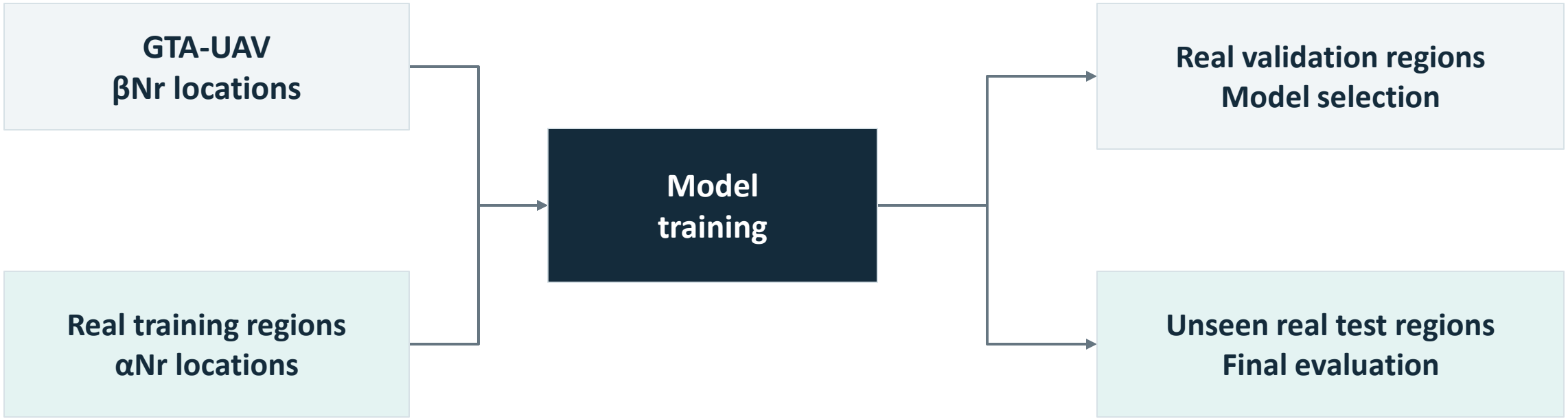
Real	Synthetic	Total
750	0	750
750	50	800
750	150	900
750	250	1,000

How much does extra synthetic data help, and where does the gain level off?

All fractions use the original 1,000-location reference. Example: 5% synthetic = 50 added locations.

Proposed geographic split

Real regions remain separate. Validation and test sets stay fixed across training compositions.



Test-region satellite tiles form the reference gallery. Test UAV images enter final evaluation only.

Core research questions and hypotheses

H1–H2 form the initial study and share the five-configuration pilot.

Focus	Research question	Working hypothesis
H1 Replacement	RQ1. Can synthetic locations compensate for reducing real training locations?	Synthetic locations can compensate for removing real locations while keeping Recall@1 within an agreed margin of the full-real baseline.
H2 Augmentation	RQ2. At a fixed real-data budget, does adding synthetic locations improve retrieval?	Adding synthetic locations improves Recall@1 over training on the same real subset alone.

H1–H2 are core. Gain saturation remains exploratory.

Real-data availability and model capacity

H3 is the first extension. H4 compares gains in a smaller and a larger model.

Focus	Research question	Working hypothesis
H3 **Real-data availability**	RQ3. Does the benefit of synthetic data change with real-data availability?	The benefit of synthetic data depends on how many real training locations are available.
H4 **Model capacity**	RQ4. Does model capacity change the benefit of synthetic training data?	Smaller and larger models obtain different gains from synthetic data.

Compare the synthetic gain within each real-data budget or model.

Training schedule and real-source transfer

Optional work packages H5–H6 use the common data and evaluation setup.

Focus	Research question	Working hypothesis
H5 Training schedule	RQ5. Does gradually introducing real data improve synthetic-to-real training?	Progressively introducing real data outperforms an abrupt switch when training steps and source exposure are matched.
H6 Real-source transfer	RQ6. Does synthetic data still help after pretraining on a separate real source?	Synthetic training improves target-domain retrieval even after pretraining on a separate real dataset.

H5 matches training exposure. H6 uses the same real-source checkpoint.

Pilot experiment matrix

Five configurations per seed. Counts illustrate 1,000 real training locations. B – baseline, S - Synthetic-assisted

Run	Real	Synthetic	Total	Purpose
B100	1,000	0	1,000	Full-real reference
B75	750	0	750	Reduced-real baseline
S05	750	50	800	Small synthetic addition
S15	750	150	900	Larger synthetic addition
S25	750	250	1,000	Shared by both tracks

Replacement (H1): S25 vs B100

Augmentation (H2): S05, S15 and S25 vs B75

Evaluation and possible conclusions

Primary metric: Recall@1 on geographically separate real test regions.

Outcome	Comparison	Interpretation
Replacement	S25 vs B100	A loss within δ, supported by uncertainty estimates, indicates that 25% fewer real locations may be sufficient.
Augmentation	S05 / S15 / S25 vs B75	Repeatable gains support synthetic additions at 75% real-data availability. Report each tested addition.
Uncertainty	Repeated seeds and results by region	Large variability or imprecise differences leave the outcome inconclusive. A clear failure of its criterion rejects that tested setting.

Scope and experiment mapping

The full plan contains six hypotheses. The initial commitment remains H1–H2.

Hypotheses	Priority	Experiment
H1–H2	Core	Five pilot configurations with one lightweight model.
H3	First extension	Repeat at a second real-data budget, with its real-only baseline.
H4	Extension	Nano vs Base at the same resolution. Compare synthetic gain.
H5	Extension	Progressive vs abrupt transition, with matched training exposure.
H6	Extension	Same real-source checkpoint: adaptation with vs without a synthetic stage.

Preliminary research plan

Six stages connect the data setup to the evidence needed for H1–H6.

01 Data audit and split

Verify real/synthetic pairs, define unique locations and freeze geographically separate train, validation and test regions.

02 Shared baseline

Run real-only baselines with one lightweight model. Fix settings and the acceptable loss margin.

03 Composition pilot

Keep the same real subset. Add nested synthetic subsets and compare replacement and augmentation for H1–H2.

04 Core replication

Repeat core comparisons across seeds. Check validation gains and variability.

05 Selected extensions

Prioritise H3: another real-data budget. Add H4 model size, H5 schedules or H6 transfer as resources allow.

06 Final evaluation

Freeze the protocol. Test on unseen real regions and report Recall@1, uncertainty and computational cost.

The core study comes first. Extensions depend on the measured run budget.