

# Physics-Informed Map-Conditioned UE Localization

Coarse-to-fine Transformer posterior with differentiable radio-map physics

Nir Tzur

January 19, 2026

## Story arc

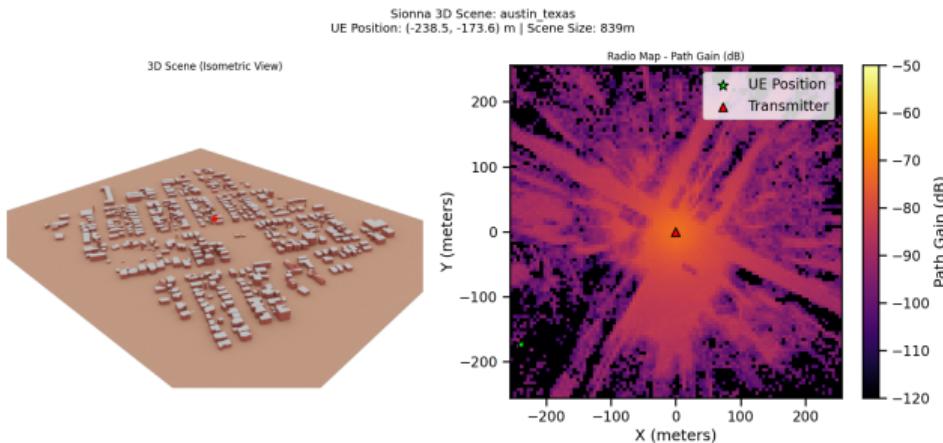
- Problem framing: localization is multi-modal and physics-constrained.
- Pipeline: OSM to scenes to radio maps to training and inference.
- Model: dual encoder, cross-attention fusion, coarse-to-fine posterior.
- Evidence: qualitative results, error statistics, and training dynamics.

## Problem and goals

- Input: sparse, irregular radio measurements across multiple protocol layers.
- Challenge: inverse mapping is ill-posed and multi-modal in urban scenes.
- Goal: map-conditioned probabilistic posterior with calibrated uncertainty.
- Constraint: predictions must remain consistent with propagation physics.

# End-to-end pipeline

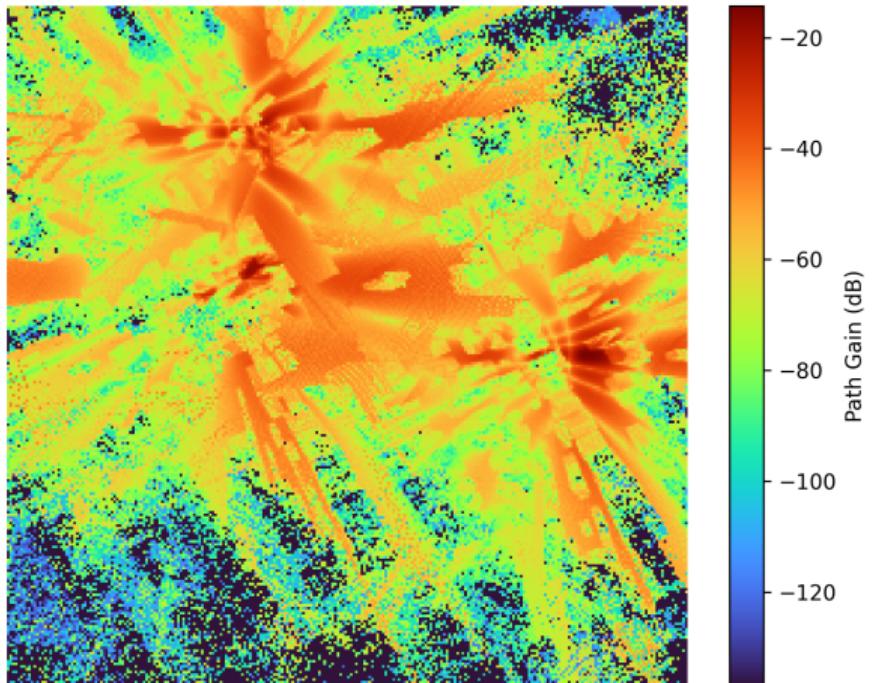
- OSM and GIS data define scene geometry and semantics.
- Differentiable ray tracing yields radio maps and measurements.
- Train a transformer to amortize inference over UE position.
- Optional refinement improves low-confidence predictions.



Scene and radio map from Sionna RT.

# Map context for localization

Radio Map (Signal Propagation)



Semantic Map (Buildings)

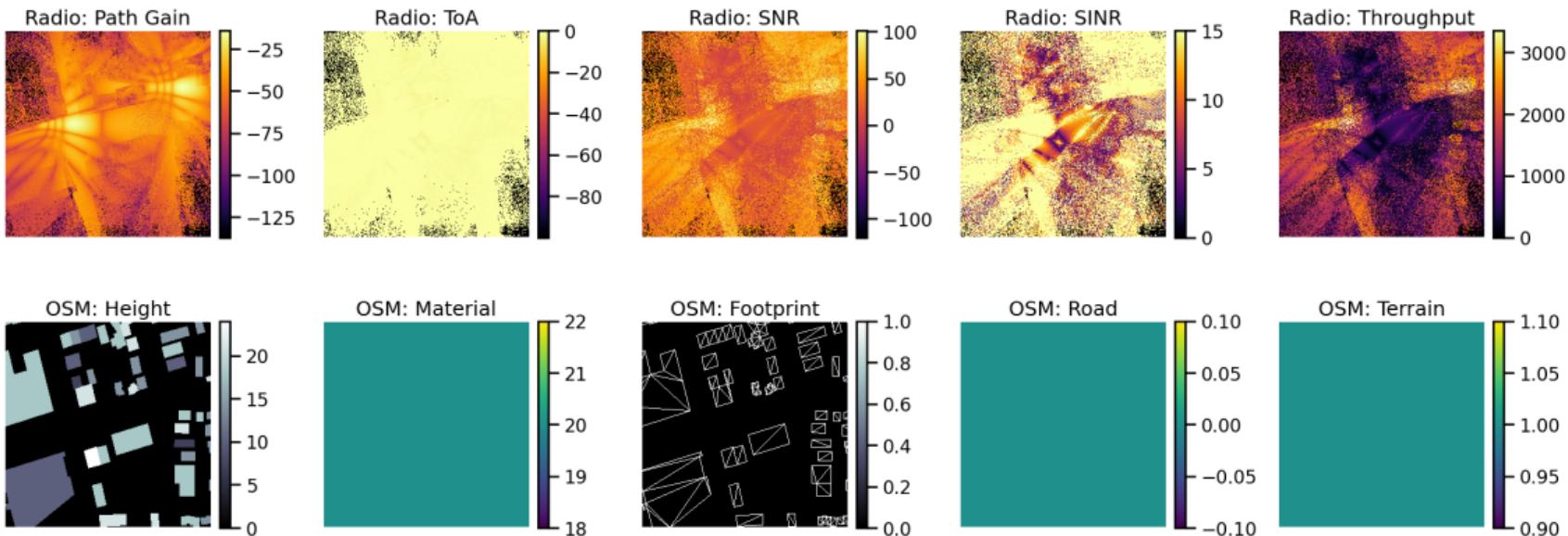


Radio propagation map and semantic building context.

# Inputs and map channels

- Radio features: RT, PHY/FAPI, and MAC/RRC measurements.
- Map features: radio channels (path gain, ToA, SNR, SINR, throughput) and OSM channels (height, material, footprint, road, terrain).

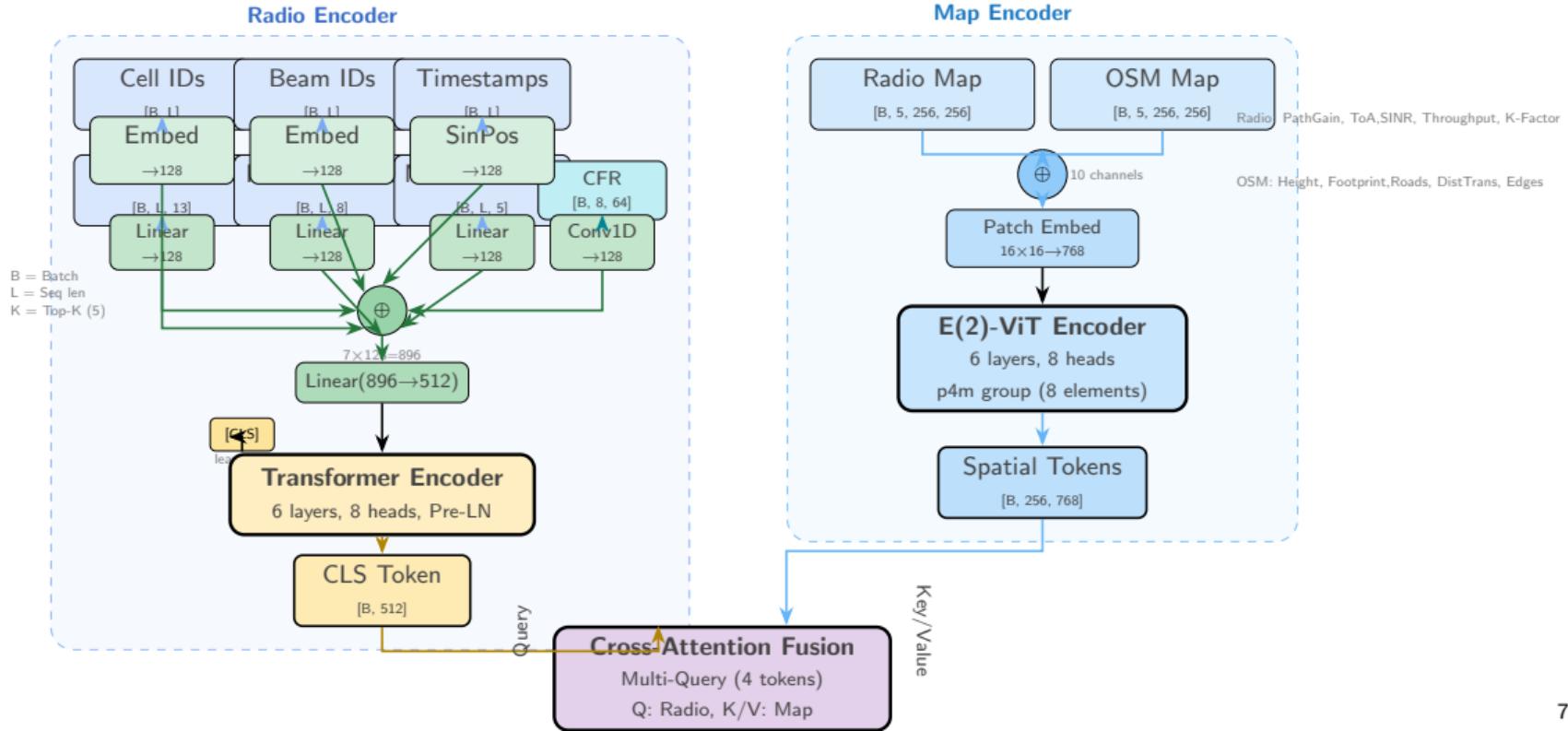
Map Layers (Sample 0)



# Model architecture

## UE Localization Model Architecture

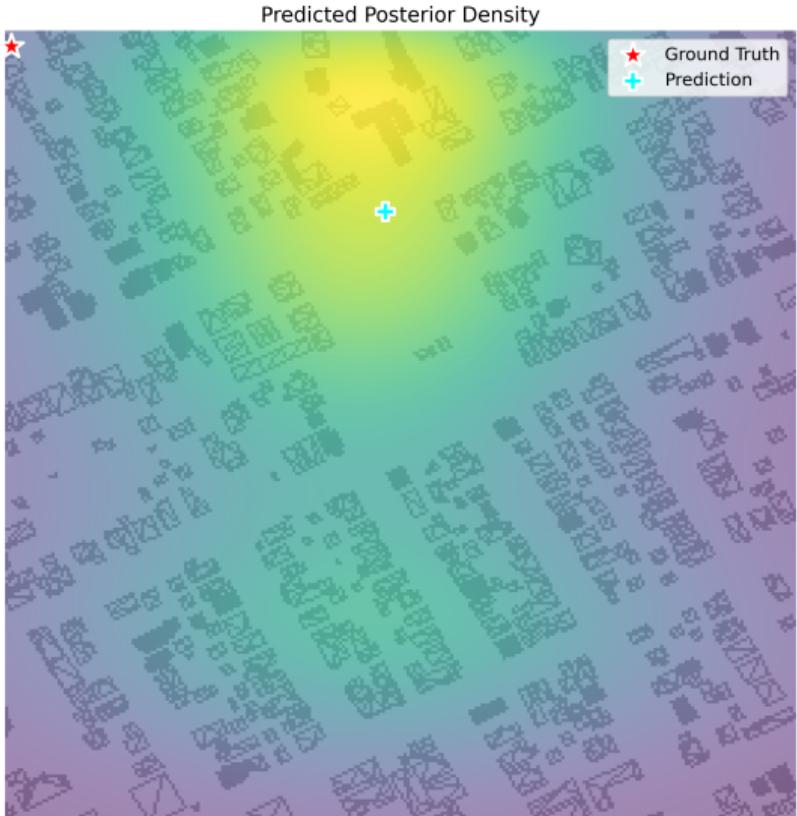
Coarse-to-Fine Map-Conditioned Transformer



## Coarse-to-fine posterior

- Coarse grid predicts a heatmap over spatial cells.
- Top-K cells refine to Gaussian components.
- Mixture posterior encodes multi-modality and uncertainty.

$$p(\mathbf{y}) = \sum_{k=1}^K \pi_k \mathcal{N}(\mathbf{y}; \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$$



# Training objective and physics loss

## Total objective

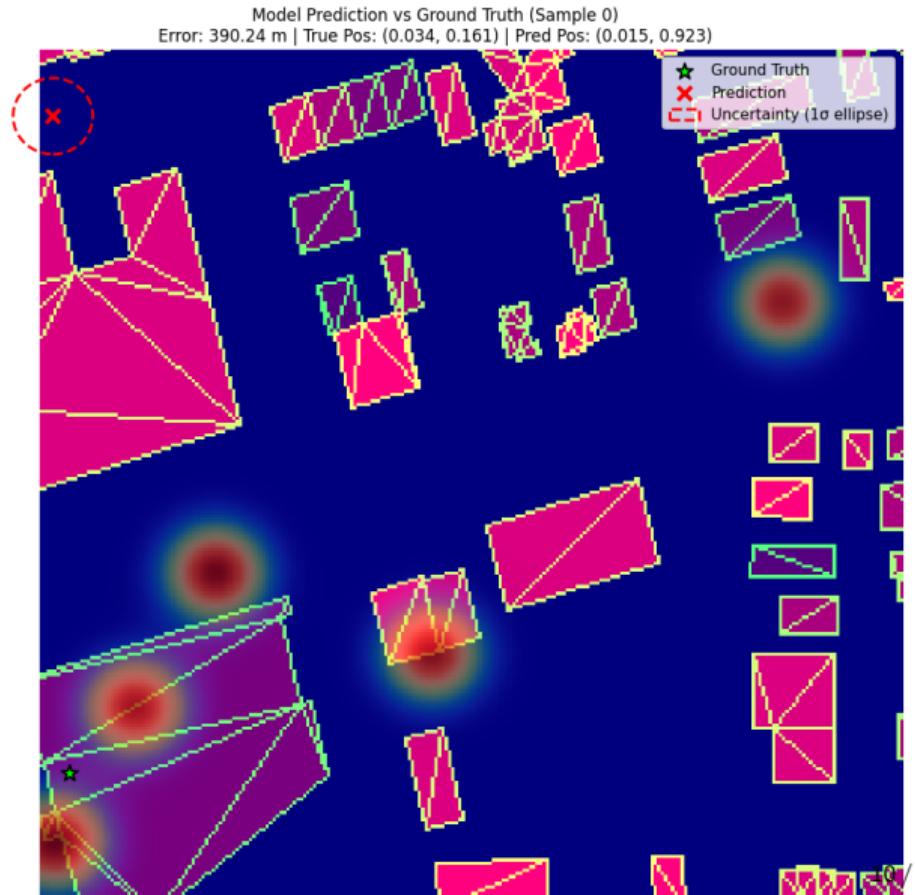
$$\mathcal{L} = \mathcal{L}_{\text{coarse}} + \lambda_f \mathcal{L}_{\text{fine}} + \lambda_p \mathcal{L}_{\text{phys}}$$

- Coarse: cross-entropy over grid cells.
- Fine: Gaussian NLL with heteroscedastic uncertainty.
- Physics: differentiable lookup of radio maps via bilinear sampling.

$$\mathcal{L}_{\text{phys}} = \sum_f w_f \|m_f^{\text{obs}} - R_f(\hat{\mathbf{y}})\|^2$$

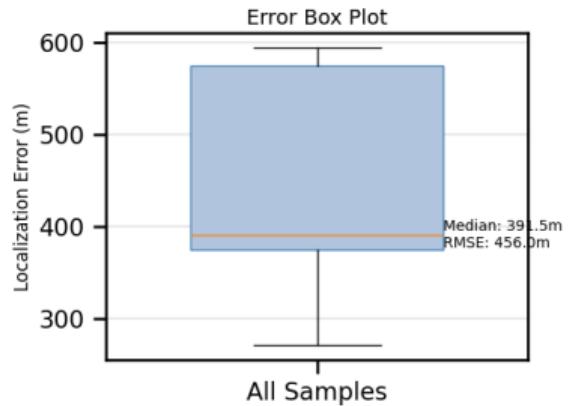
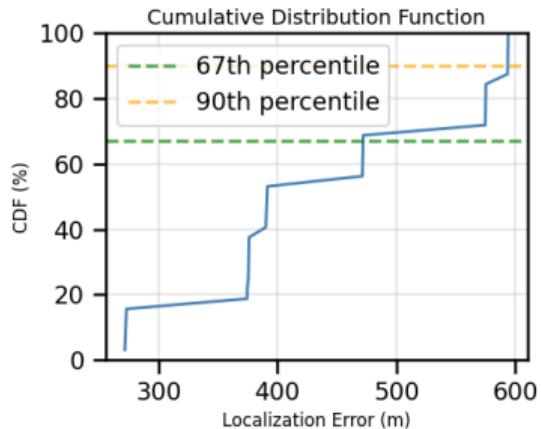
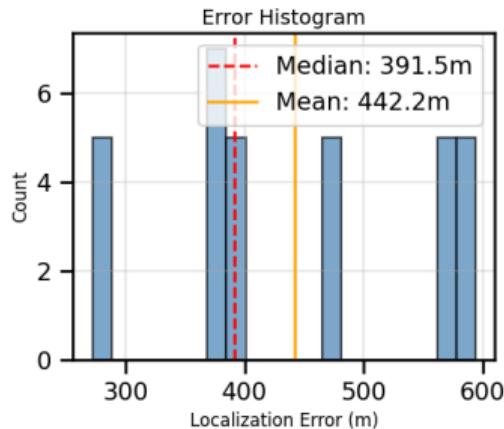
## Inference and refinement

- Predict heatmap and top-K candidate cells.
- Refine to continuous positions with uncertainty.
- Optional MAP refinement combines network posterior with radio-map likelihood.



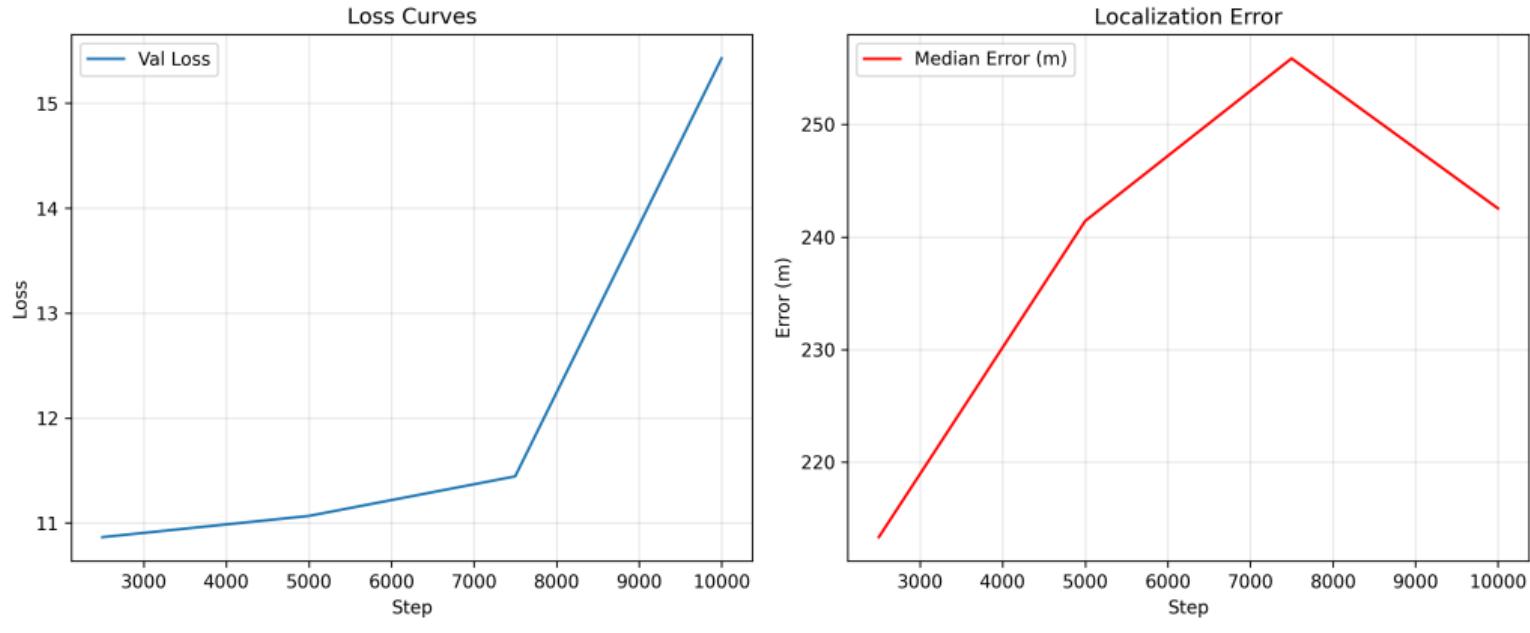
# Evaluation metrics

Batch Error Analysis



Error histogram, CDF, and box plot for batch evaluation.

# Training dynamics



curves and median localization error over steps.

## Project status and integration

- Scene generation from OSM.
- Differentiable ray tracing and dataset generation.
- Dual-encoder transformer with fusion.
- Training and evaluation pipeline.
- Streamlit web demo.

Milestone	Status
Scene generation	Complete
Data generation	Complete
Model	Complete
Training	Complete
Web UI	Complete

## Conclusion and next steps

- Map-conditioned transformer yields a calibrated, multi-modal posterior.
- Physics-informed loss aligns predictions with radio propagation.
- Pipeline is end-to-end: OSM to scenes to training to inference.
- Next: larger-scale city runs, ablations, and uncertainty calibration.