TransPathNet: A Novel Two-Stage Framework for Indoor Radio Map Prediction

ICASSP 2025 Indoor Pathloss Radio Map Prediction Challenge

Xin Li, Ran Liu, Saihua Xu, Sirajudeen Gulam Razul, Chau Yuen

Nanyang Technological University, Singapore

xin019@e.ntu.edu.sg, {ran.liu, shxu, esirajudeen, chau.yuen}@ntu.edu.sg

March 18, 2025

Outline

1. **Problem Statement and Challenges**
2. **Proposed Method: TransPathNet**
3. **Implementation Details**
4. **Experimental Results**
5. **Conclusion and Future Work**

Problem Statement and Challenges

Indoor Pathloss Radio Map Prediction

- Given a 2D floor plan (geometry, materials), antenna info, frequency, etc.
- Predict pathloss (in dB) at each pixel.

ICASSP 2025 Challenge Setup

- Dataset: extensive ray-tracing simulations.
- Tasks:
  1. New geometry
  2. New frequency
  3. New antenna patterns
- Evaluation metric: **RMSE (dB)**.

Key Difficulties

- Complex interactions (reflection, diffraction).
- Structural variations in building design.
- Wide range of frequencies & antenna characteristics.

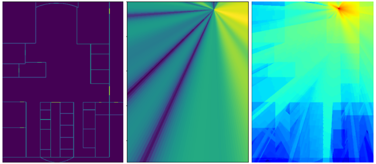


Figure: Dataset example.

System Overview

Two-Stage Coarse-to-Fine Framework

- **Stage 1:** Coarse prediction.
- **Stage 2:** Fine-grained refinement (focus on residuals, finer details).

TransPathNet Architecture

- **Transformer-based encoder** + **multiscale convolutional attention decoder**.
- Enhanced input channels for better representation.

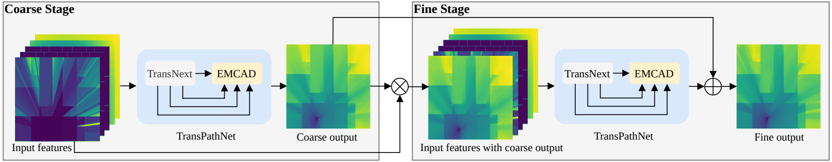


Figure: Overview of the TransPathNet two-stage architecture.

TransPathNet Architecture

Encoder (TransNeXt)

- Transformer + convolution layers.
- Learns hierarchical features from complex indoor layouts.

Decoder (EMCAD)

- Multiscale upsampling.
- Convolutional attention refines pathloss maps.

Skip Connections

- Retain high-resolution details.
- Ensure effective gradient flow.

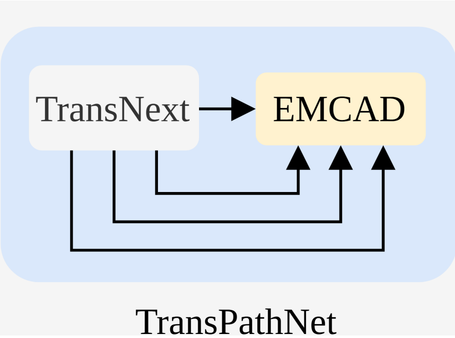


Figure: Network architecture.

Two-Stage Coarse-to-Fine

Coarse Stage

- Generates an initial (rough) pathloss map.
- Captures global structure but may lack fine detail.

Fine Stage

- Takes coarse output + original input features.
- Learns residual corrections to refine the final pathloss prediction.

Result

- Refined prediction for complex indoor conditions and subtle variations.

Enhanced Input Features

Default 3-Channel Inputs

- Reflectance, transmittance, distance maps.

Enhanced Inputs

- Free Space Pathloss (FSPL) estimate.
- Transmission Ray Encoding (direct multipath emphasis).
- Antenna Embeddings (pattern, angle, position).
- Spatial-Frequency Encodings (positional + frequency embeddings).

Motivation

- Richer representation $\Rightarrow$ more robust performance across varied conditions.
- Improves generalization to unseen geometry/frequency/antenna scenarios.

Implementation Details

Training Configuration

- **Framework:** PyTorch
- **Loss:** MSE
- **Optimizer:** Adam, LR $= 10^{-4}$ (halved later)
- **Data Augmentation:** random flips, rotations

Hyperparameters

- Input resolution: 384×384
- Batch size: 4
- Epochs: 30
- **Post-Processing:** Flip + rotate test-time ensembling
- Hardware: NVIDIA RTX 4090
  - Inference time: ≈ 43.8 ms per sample

Overall Performance

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Evaluation Metric: RMSE (dB)

- Tests on:
  1. **Kaggle Subset**
  2. **Full Test Set**
- Weighted tasks: geometry (30%), frequency (30%), antenna (40%).

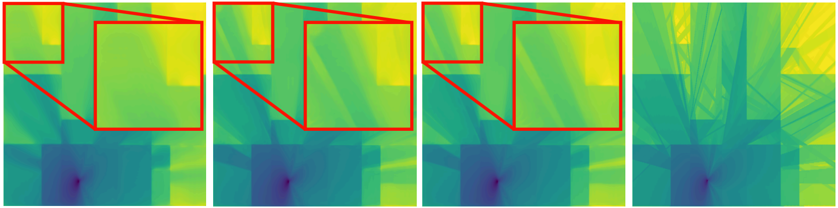
TransPathNet Results

- RMSE = 9.73 dB on Kaggle subset
- RMSE = 10.40 dB on full test set

Case	**Two-Stage**	**Post-Proc.**	**RMSE(dB): Kaggle** $\downarrow$	**RMSE(dB): full** $\downarrow$
Coarse only	$\times$	$\times$	9.93	10.327
+ Two-Stage Training	$\checkmark$	$\times$	9.75	10.430
Full pipeline	$\checkmark$	$\checkmark$	9.73	10.397

Table: Performance across different configurations.

Qualitative Results



(a) Coarse only RMSE 5.01 | (b) Two-stage training RMSE 4.26 | (c) Full pipeline RMSE 4.24 | (d) Target

Figure: (a) Ground truth, (b) Coarse prediction, (c) Fine-stage output, (d) Final post-processed.

- **Observations:**
  - Fine stage visibly corrects errors in complex regions.
  - Post-processing brings additional smoothness and consistency.

Conclusion

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Key Points

- Introduced **TransPathNet**: a two-stage coarse-to-fine deep network for indoor pathloss map prediction.
- **Architecture**: Transformer-based encoder (TransNeXt) + EMCAD decoder with skip connections.
- **Enhancements**: Extra input features (FSPL, antenna embeddings, freq. encodings).
- Achieves strong results ($\approx$ 10 dB RMSE) on the ICASSP 2025 Challenge set.

Limitations

- Inaccuracies under heavy reflection scenarios.
- Possible overfitting to certain floorplan styles.

Project Page

- `https://lixin.ai/TransPathNet/`

Future Work

Reflection-Aware Modules

- Integrate explicit multipath or advanced ray-based features.

3D Extensions

- Handle true volumetric data or multi-floor/3D building models.

Real-World Validation

- Compare with measurements from real indoor deployments.

Thank You for Your Attention!

Questions?

https://lixin.ai/TransPathNet/