Alireza Limooie

Roof Graph Edge Classification Progress Report

Date: 07/07/2025

1. Introduction

This report details the progress on project focused on the geometric reconstruction of roof structure plans. The primary objective is to develop a model that can take noisy or imperfect roof graph data—such as that produced by an automated line detector—and iteratively refine it into a clean, geometrically correct representation. Recent work has focused on developing a denoising model, addressing key performance challenges, and establishing a robust training and data augmentation pipeline.

Initial versions of the model, while effective at removing high levels of noise, exhibited convergence issues and underperformed on inputs with low levels of noise, which is a critical use case. Key accomplishments in this reporting period include a re-architecting of the training process and loss function to solve this issue. Specifically, the introduction of a **Connectivity Loss** to preserve topological integrity and an expansion of the **diffusion schedule (from T=3 to T=10)** have yielded significant improvements. Promising initial results have been achieved after 10 epochs of training on the new configuration.

2. Model Overview

The system is built around a **Transformer Coordinate Reconstructor**, an iterative denoising model designed to process and refine entire roof graphs.

- **Functionality:** The model operates within a diffusion framework. It is trained to reverse a noise-adding process. During inference, it takes a noisy graph and predicts a less-noisy version of the graph.

- **Input:** The model receives a sequence of line segments, where each segment is described by a set of 72 geometric features (e.g., noisy coordinates, length, angle) and the current time step t. The model uses a K-Nearest Neighbors (k=16) attention mechanism to focus on local context.

- **Output:** The model is configured to predict the *offsets* required to transform the noisy input coordinates into the clean, ground-truth coordinates (denoise-to-clean). This direct-to-clean prediction strategy simplifies the inference chain.

3. Data Preparation and Enhancement

A sophisticated data preparation and augmentation pipeline is used to create a robust training environment.

- **Source Data:** The pipeline starts with JSON files containing clean, normalized coordinates and connectivity information for thousands of roof plans.

- **Denoising Task Generation:** For each clean roof graph, a "diffusion process" is simulated by generating multiple noisy versions. With a schedule of **T=10**, this creates 10 distinct training tasks per graph, each corresponding to denoising from a different noise level (t=10 -> t=0, t=9 -> t=0, etc.).

- **Online Geometric Augmentation:** To improve the model's robustness against real-world detector errors, aggressive geometric augmentations are applied on-the-fly during training. This includes:

    - **Edge Subdivision:** Splitting single edges into multiple smaller segments.

    - **Edge Duplication:** Creating slightly offset copies of existing edges.

- **Node & Edge Deletion:** Randomly removing nodes (and their connected edges) or individual edges.

This process ensures the model learns to reconstruct correct geometry even from topologically imperfect inputs.

4. Addressing Convergence and Low-Noise Performance

A primary challenge was the model's difficulty in making fine-grained adjustments when the input was already close to the ground truth (low t values). Two key strategies were implemented to mitigate this.

- **Connectivity Loss:** A LogConnectivityLoss term was added to the total loss function. This term penalizes the model if endpoints that should be connected in the final graph are predicted to be far apart. This forces the model to learn and preserve the topological structure of the roof, preventing it from breaking apart valid junctions during the final, fine-tuning steps of denoising.

- **Expanded Diffusion Schedule:** The number of diffusion steps (T) was increased from 3 to 10. This creates a more granular noise schedule and a more gradual denoising process. By training on many intermediate noise levels, the model is better able to learn the subtle adjustments required at the low-noise end of the spectrum, leading to significantly better convergence and final accuracy.

5. Model Training and Performance Evaluation

The model was trained using the new configuration (hpa76492744b) on the prepared dataset.

- **Training Data:** The model was trained on a dataset of **247,189** tasks (expanded from ~29,000 unique graphs) and evaluated against a validation set of **43,621** tasks.
- **Training Process:** The training was conducted for **10 epochs**. Due to the complexity of the model and the large, augmented dataset, training times were substantial. The model's performance was tracked using a combined loss (Reconstruction + Connectivity) on the validation set.
- **Results:** The model demonstrated strong and consistent improvement across the 10 epochs. The final validation loss reached **3.81e-02**. The steadily decreasing loss curve indicates that the model has not yet fully converged and has significant room for improvement with further training. The primary constraint has been the long training time per epoch.

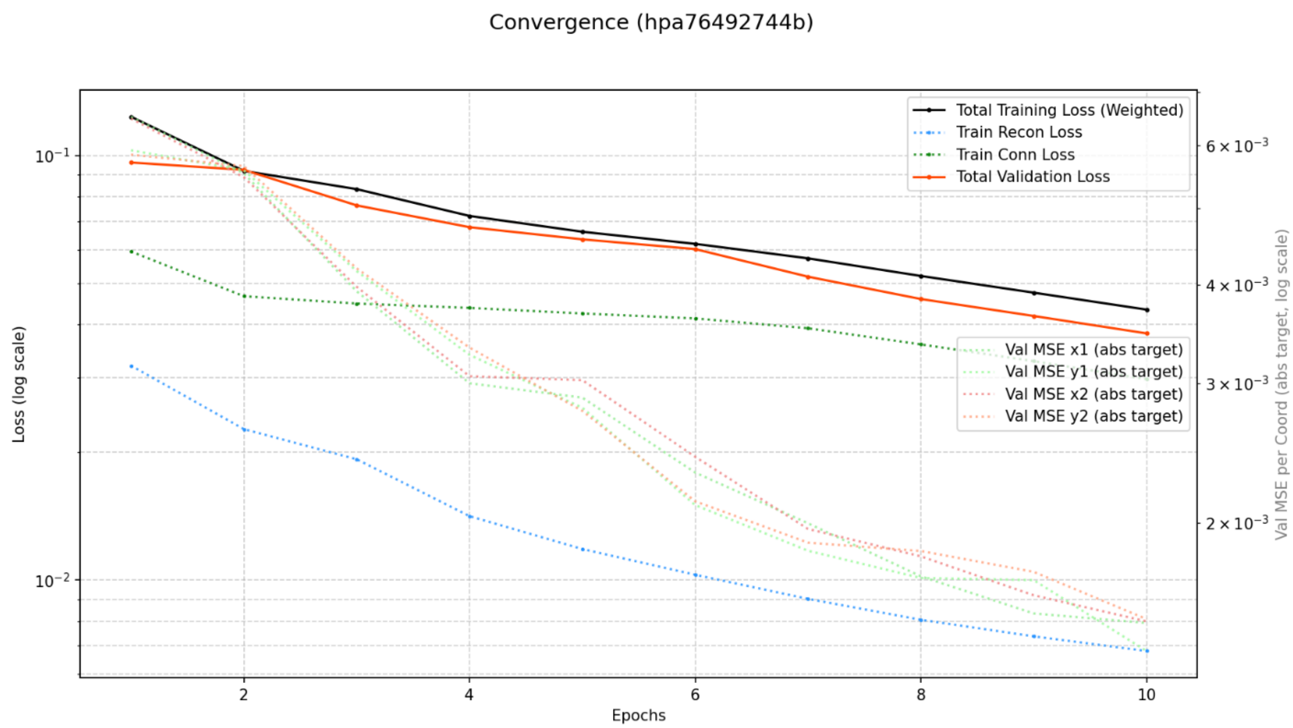
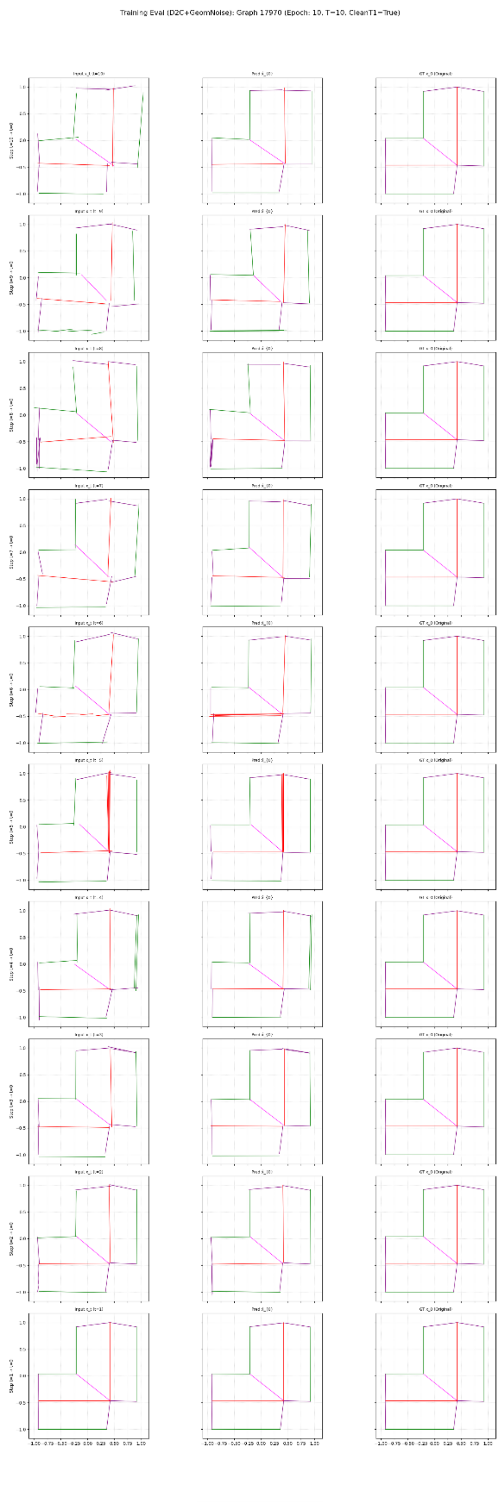


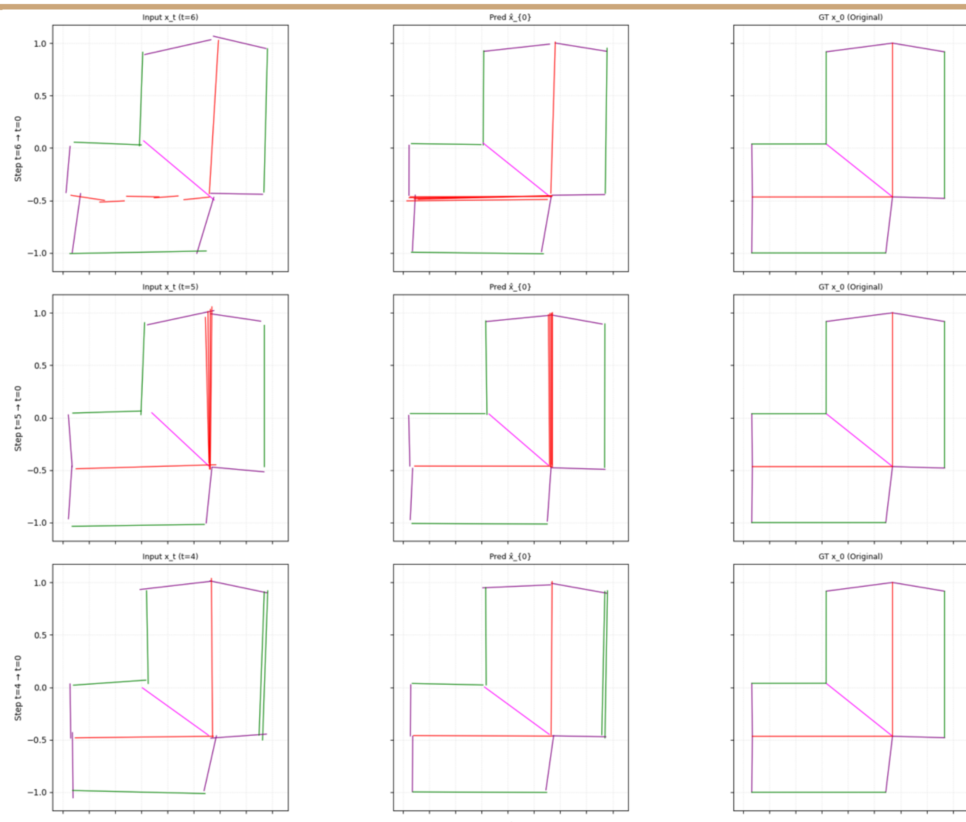
Table 1: Training Run Summary (hpa76492744b)

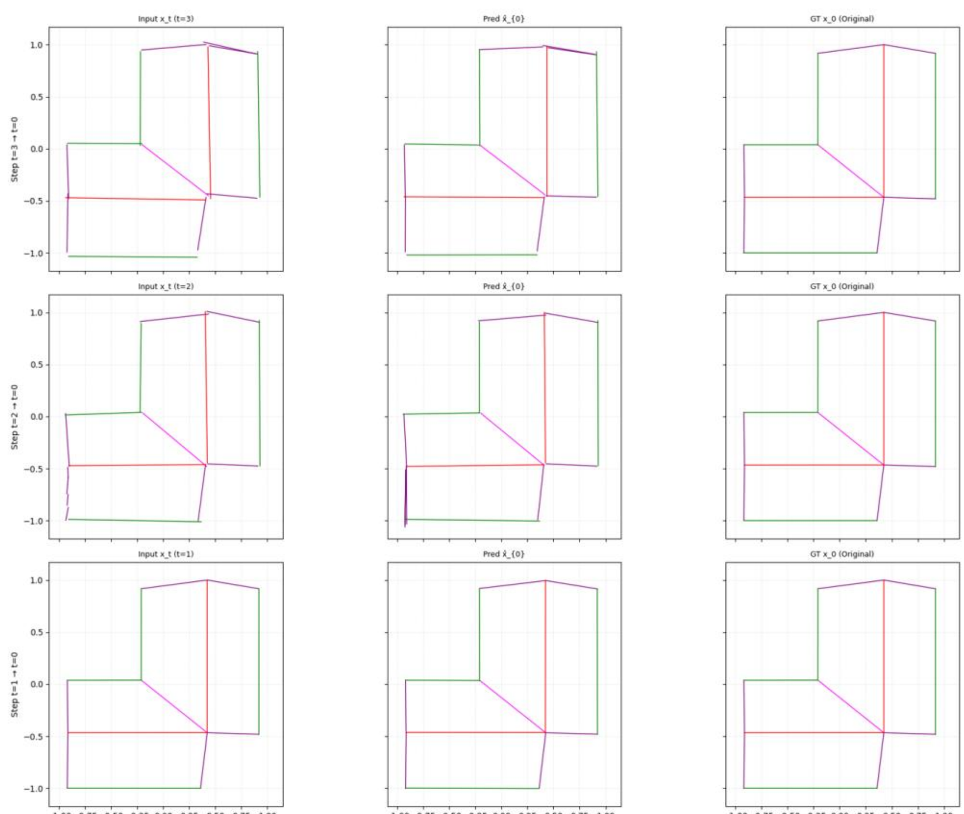
Metric	Value
Trainable Parameters	4,766,832
Model Size	18.667 MB
Number of Training Tasks	247,189
Number of Validation Tasks	43,621
Training Epochs Completed	10
Total Training Time	193.88 minutes

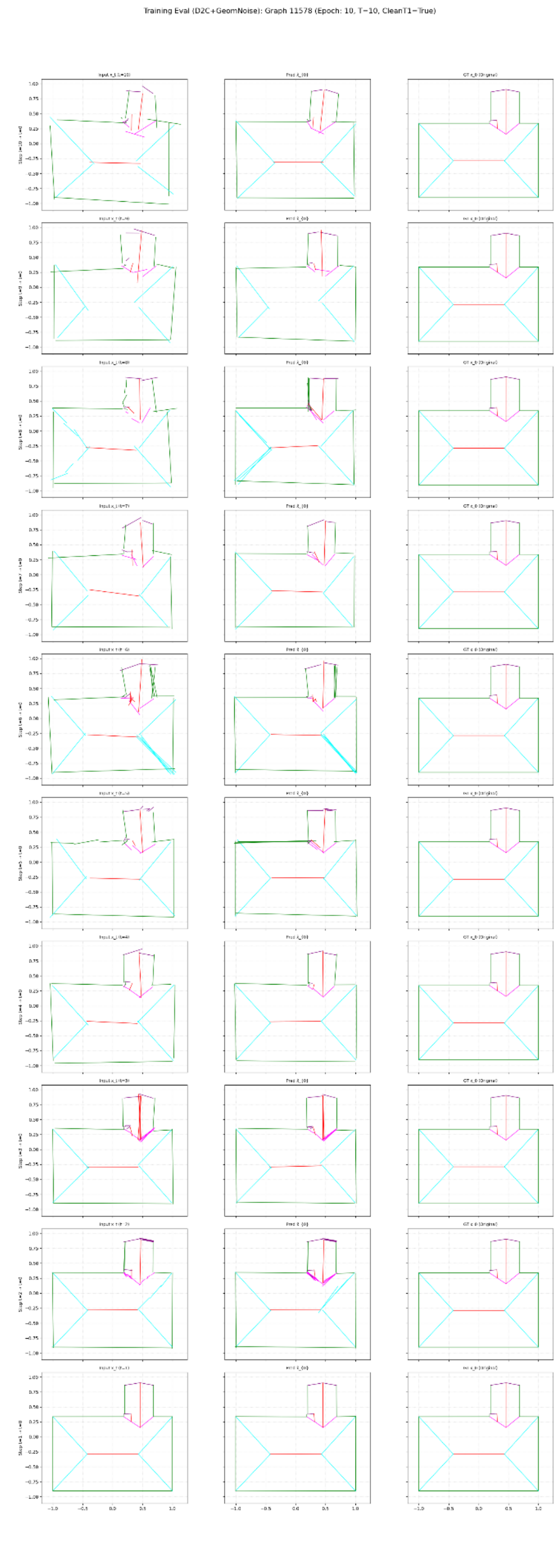
Training Samples:

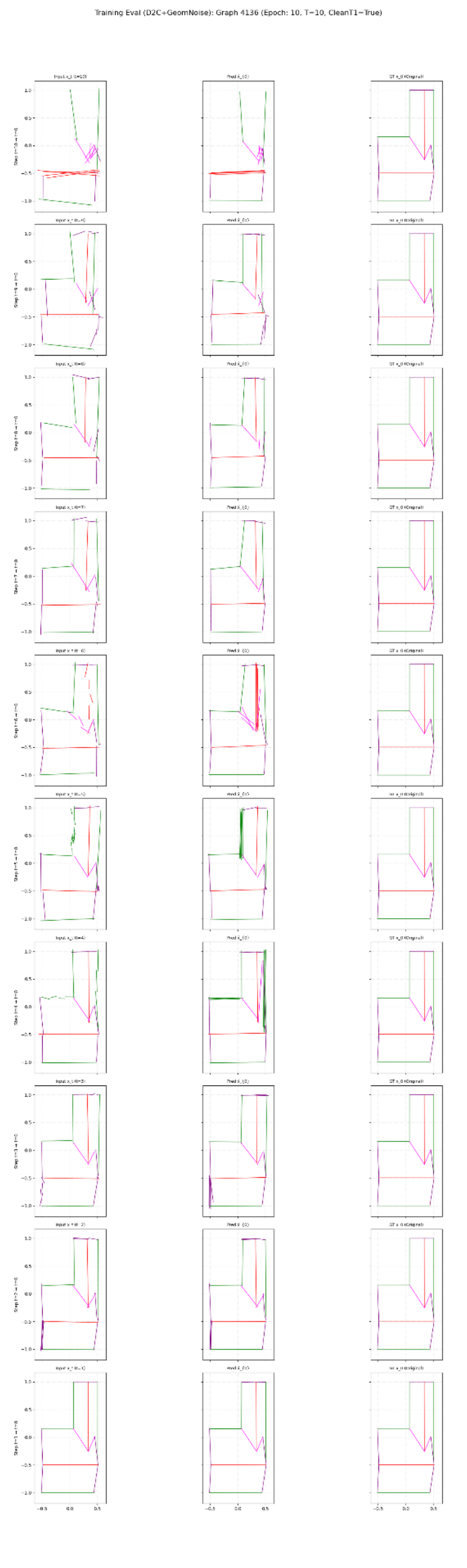
Examples of Final Model Output on Satellite Imagery




Input x_t (t=6)
Pred x̂_{0}
GT x_0 (Original)
Step t=6 → t=0
Input x_t (t=5)
Pred x̂_{0}
GT x_0 (Original)
Step t=5 → t=0
Input x_t (t=4)
Pred x̂_{0}
GT x_0 (Original)
Step t=4 → t=0
1.0
0.5
0.0
−0.5
−1.0


Input x_t (t=3)
Pred x̂_{0}
GT x_0 (Original)
Step t=3 → t=0
Input x_t (t=2)
Pred x̂_{0}
GT x_0 (Original)
Step t=2 → t=0
Input x_t (t=1)
Pred x̂_{0}
GT x_0 (Original)
Step t=1 → t=0
1.0
0.5
0.0
−0.5
−1.0
−1.00 −0.75 −0.50 −0.25 0.00 0.25 0.50 0.75 1.00


Training Eval (D2C+GeomNoise): Graph 11578 (Epoch: 10, T=10, CleanT1=True)
Ridge
Hip
Valley


Training Eval (D2C+GeomNoise): Graph 4136 (Epoch: 10, T=10, CleanT1=True)
Ridge
Valley

6. Future Work

The next phase of work will focus on realizing the model's full potential and enhancing its real-world applicability.

- **Complete Model Training:** The highest priority is to continue training the current model configuration for the full 100 epochs, or until the validation loss fully plateaus.
- **Scale to Larger Dataset:** Transition training to a new, significantly larger dataset (estimated 1.2 million graphs vs. the current ~30,000). This will expose the model to a much wider variety of roof structures and improve generalization.
- **Enhance Geometric Augmentation:** Refine the existing augmentation strategies to more accurately simulate the types of errors produced by upstream line detectors. This includes developing methods to intelligently **add** new line segments (to fix false negatives) and **delete** spurious ones (to fix false positives) from the input graph, making the model a true end-to-end cleaning tool.
- **Hyperparameter Optimization:** Systematically explore different model configurations, loss types and learning rate schedules to further optimize performance, especially once training on the large dataset begins.

Conclusion

Progress has been made in developing a robust roof graph reconstruction system. The implementation of a connectivity-aware loss function and an expanded diffusion schedule has effectively solved prior convergence issues and improved the model's ability to perform fine-grained adjustments. The current model demonstrates excellent learning potential, achieving a validation loss of **3.81e-02** after only 10 epochs. While performance is already promising, the steadily decreasing loss curve indicates that the model is far from its peak performance. Future work will focus on completing the training cycle and scaling up to a much larger dataset to fully develop this powerful reconstruction capability.