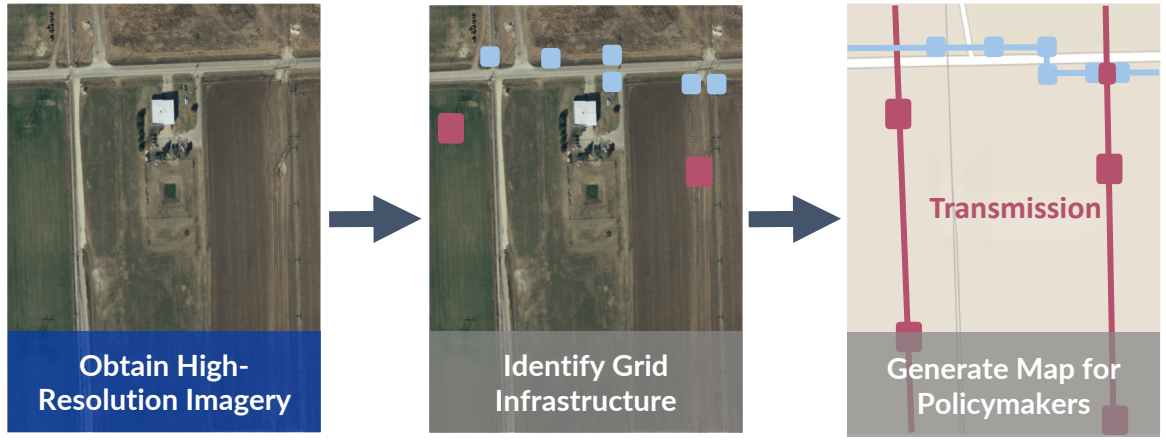
A Wider Lens on Energy: Adapting Deep Learning Techniques to Inform Energy Access Decisions

Varun Nair, Paul Rhee, Jichen Yang, Fanjie Kong, Dr. Kyle Bradbury, Dr. Jordan Malof

Problem

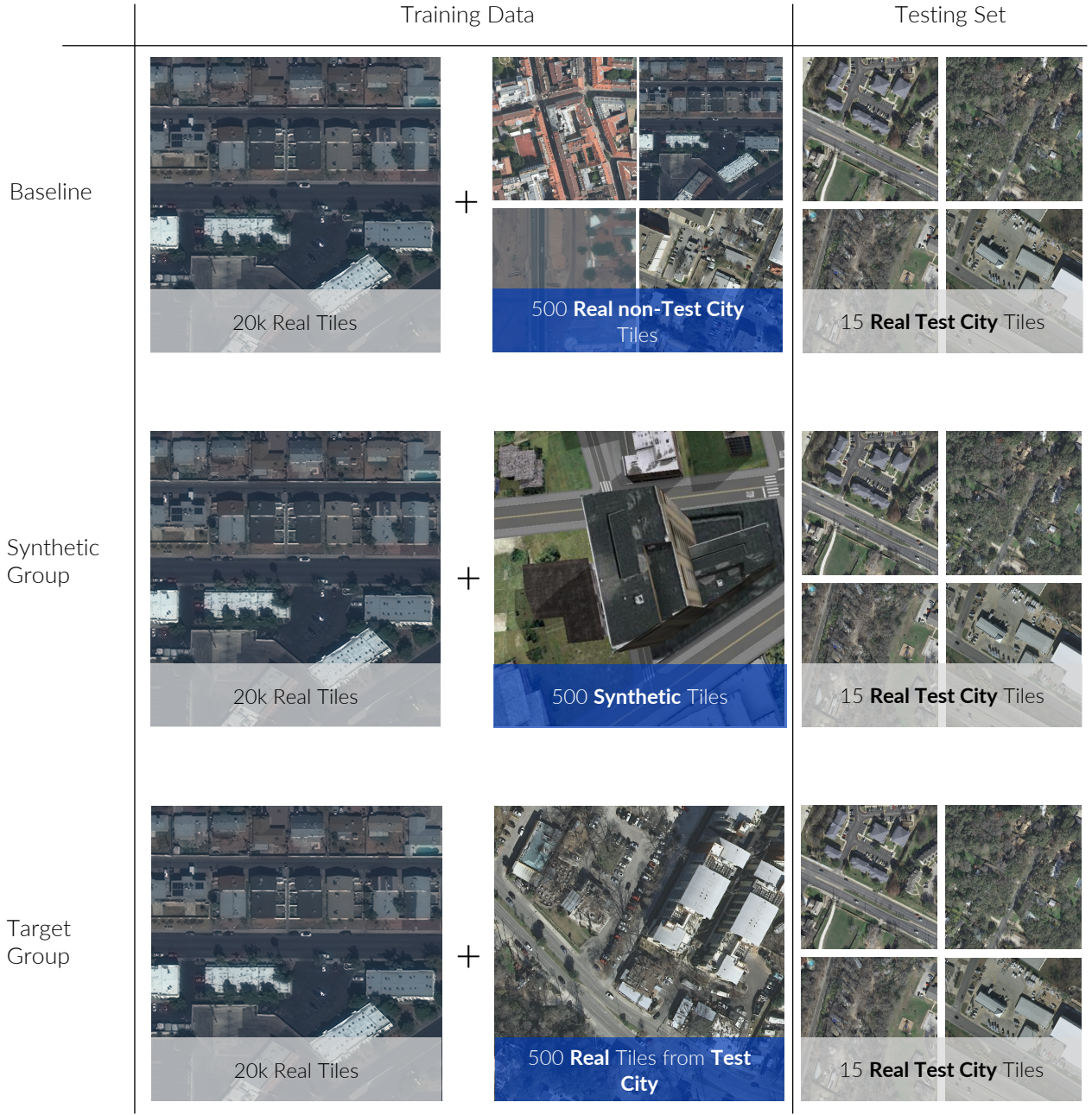
- We often deal with sparse information when mapping energy infrastructure
- These maps are important to determining the best way to electrify unelectrified regions
- The leading models at identifying this infrastructure **do best with lots of geographically diverse data**
- This data is not only **expensive**, but often doesn't exist for regions of interest



Solution

- It's **simple and cheap** to create synthetic diverse imagery for various terrains and geographies
- Can then **train on real datasets augmented with synthetic data** to improve performance
- These models should be **more robust to context changes** such as groundcover, vegetation, and urban density that infrastructure is seen in

Experimental Design

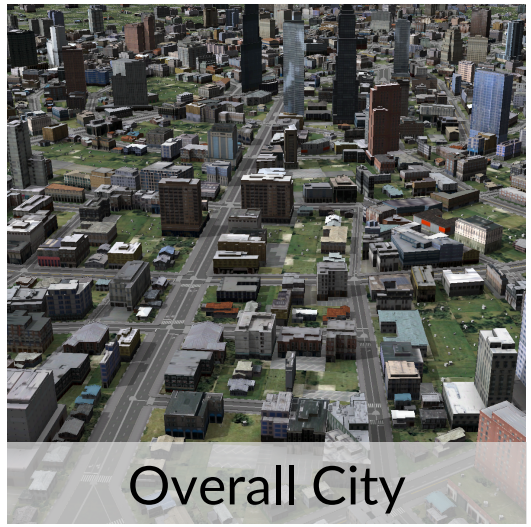


Methodology

1 Designing the City

Façade Textures

Roof Textures



Overall City

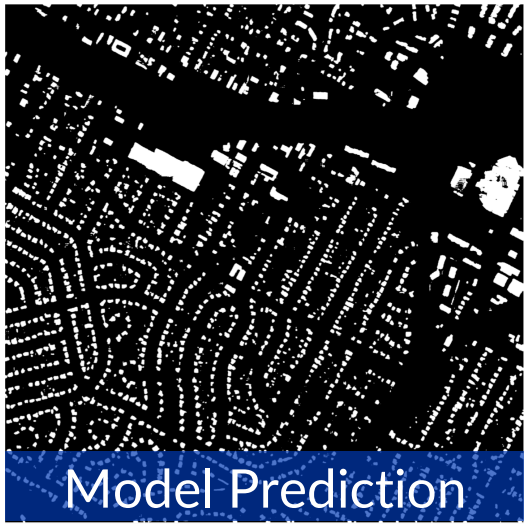
1. Used OpenStreetMap to **mimic the street layout** of cities
2. Generated various random layouts
3. Pulled roofing and façade textures from Google Maps

2 Testing the Models

1. **Tested** the trained models on real satellite imagery
2. The model **predicts where buildings are** in testing imagery
3. **Compared predictions with ground truth** to measure accuracy using IoU

Test City

Ground Truth



Model Prediction

Results & Conclusions

Trained On	Vienna Testing Results	
	Mean IoU	Standard Deviation
Baseline (Real Non-Target Imagery)	0.655	0.0096
Synthetic Group (Baseline + Synthetic Data)	0.663	0.0032
Target Group (Baseline + Real Target Imagery	0.738	0.0015

- We were able to design synthetic cities to resemble their real counterparts (left figure)
- But, can this improve the **identification of energy infrastructure in satellite imagery**?
- Training on real non-target city and synthetic data improved performance relative to training on the real data (baseline) alone
- Adding synthetic imagery designed to match that city boosted its performance on Vienna (above table)
- In data-poor environments or when applying to new terrains, it is possible that augmenting available training data that includes synthetic data could extend models' generalizability in new domains