

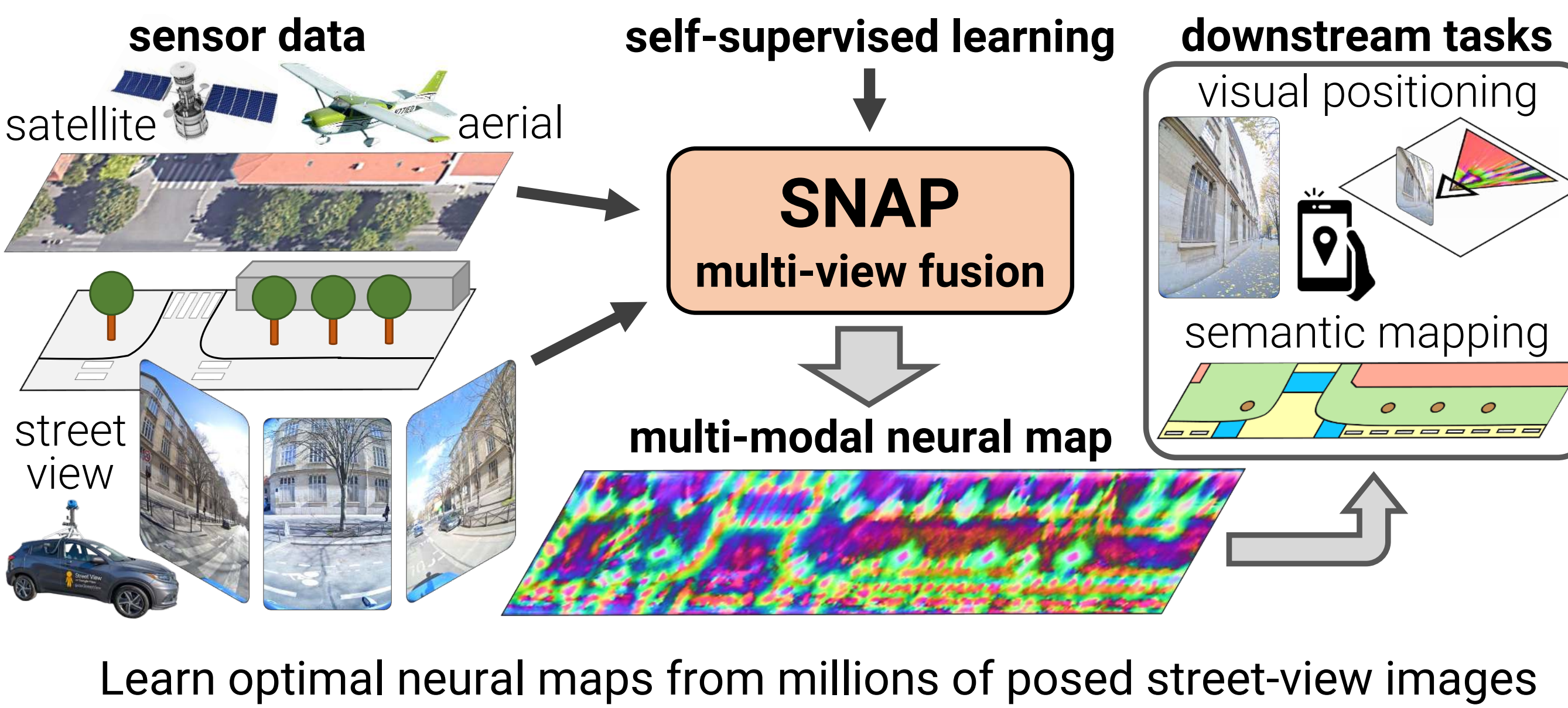
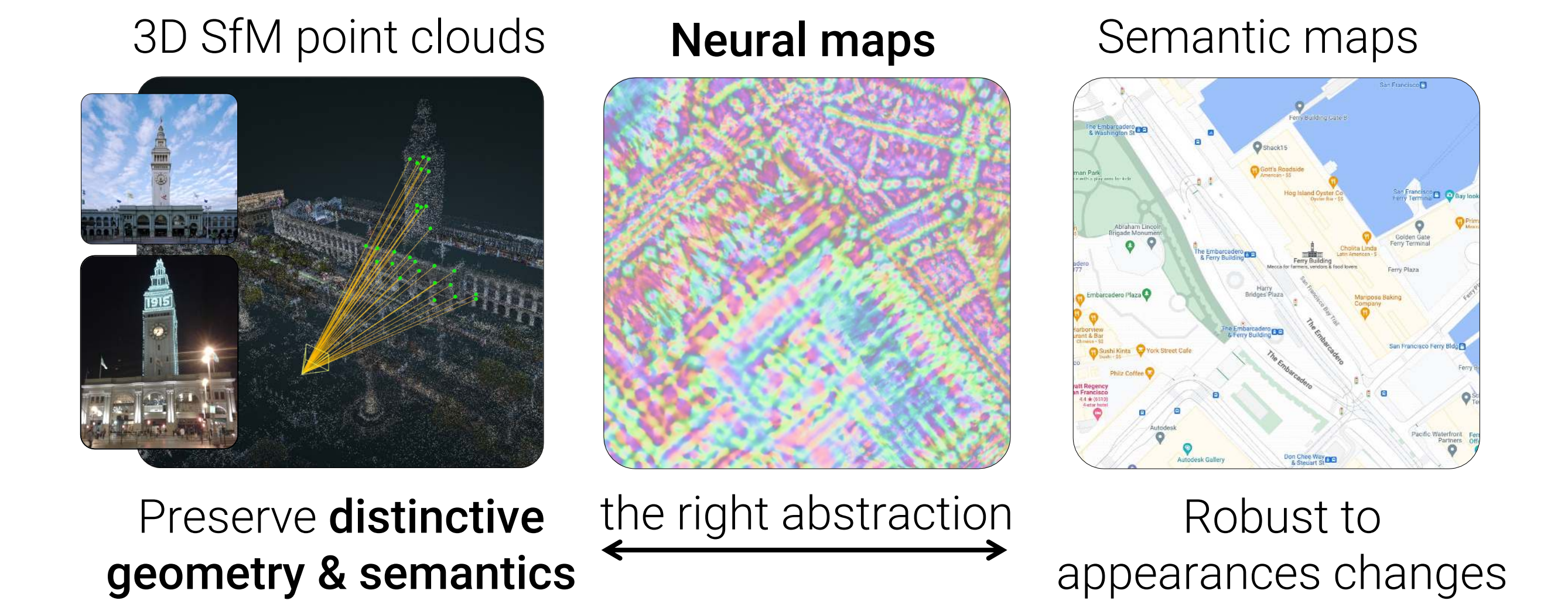


# SNAP! Self-Supervised Neural Maps for Visual Positioning and Semantic Understanding

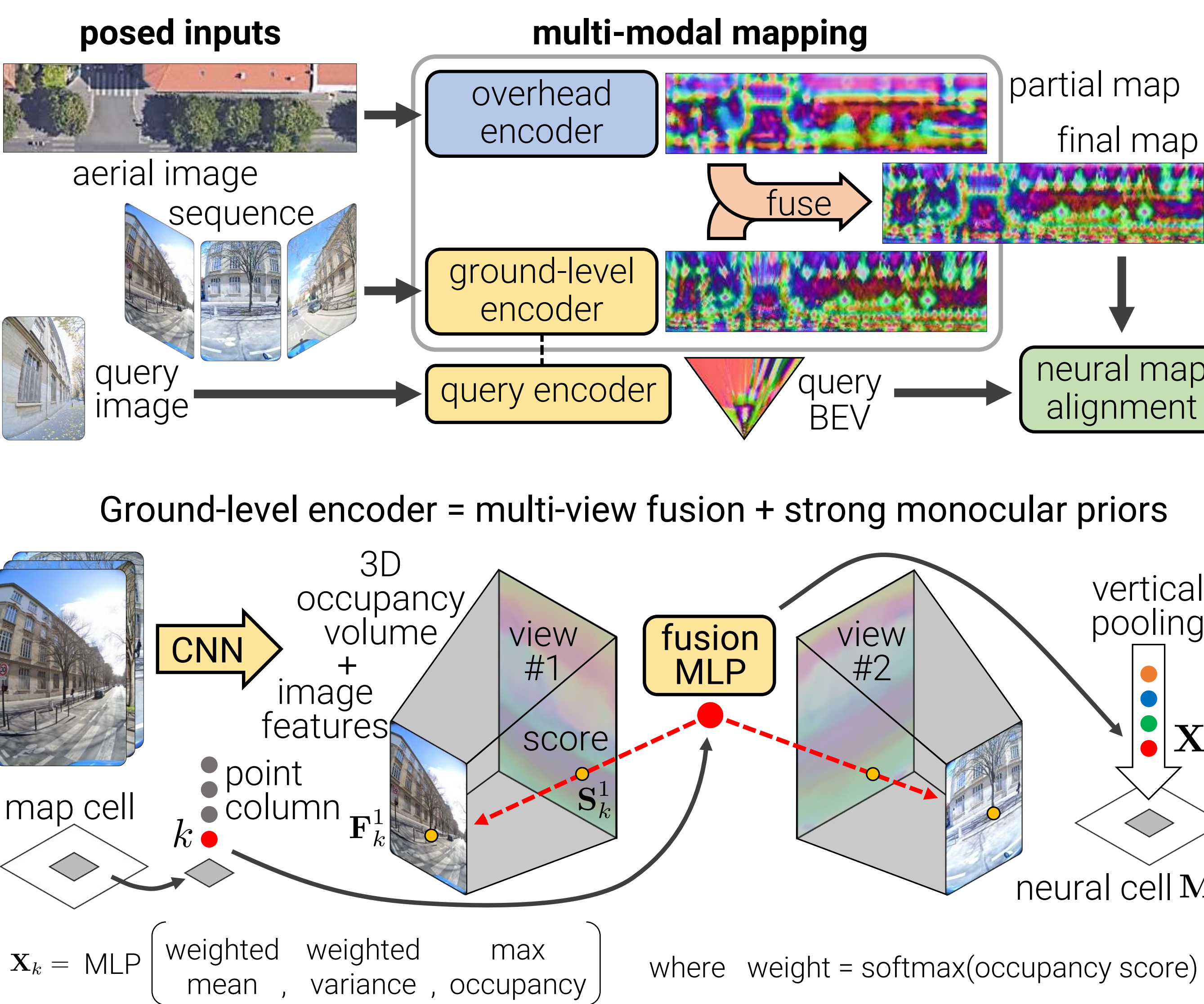
Paul-Edouard Sarlin<sup>1</sup> Eduard Trulls<sup>2</sup> Marc Pollefeys<sup>1</sup> Jan Hosang<sup>2</sup> Simon Lynen<sup>2</sup>

## Overview

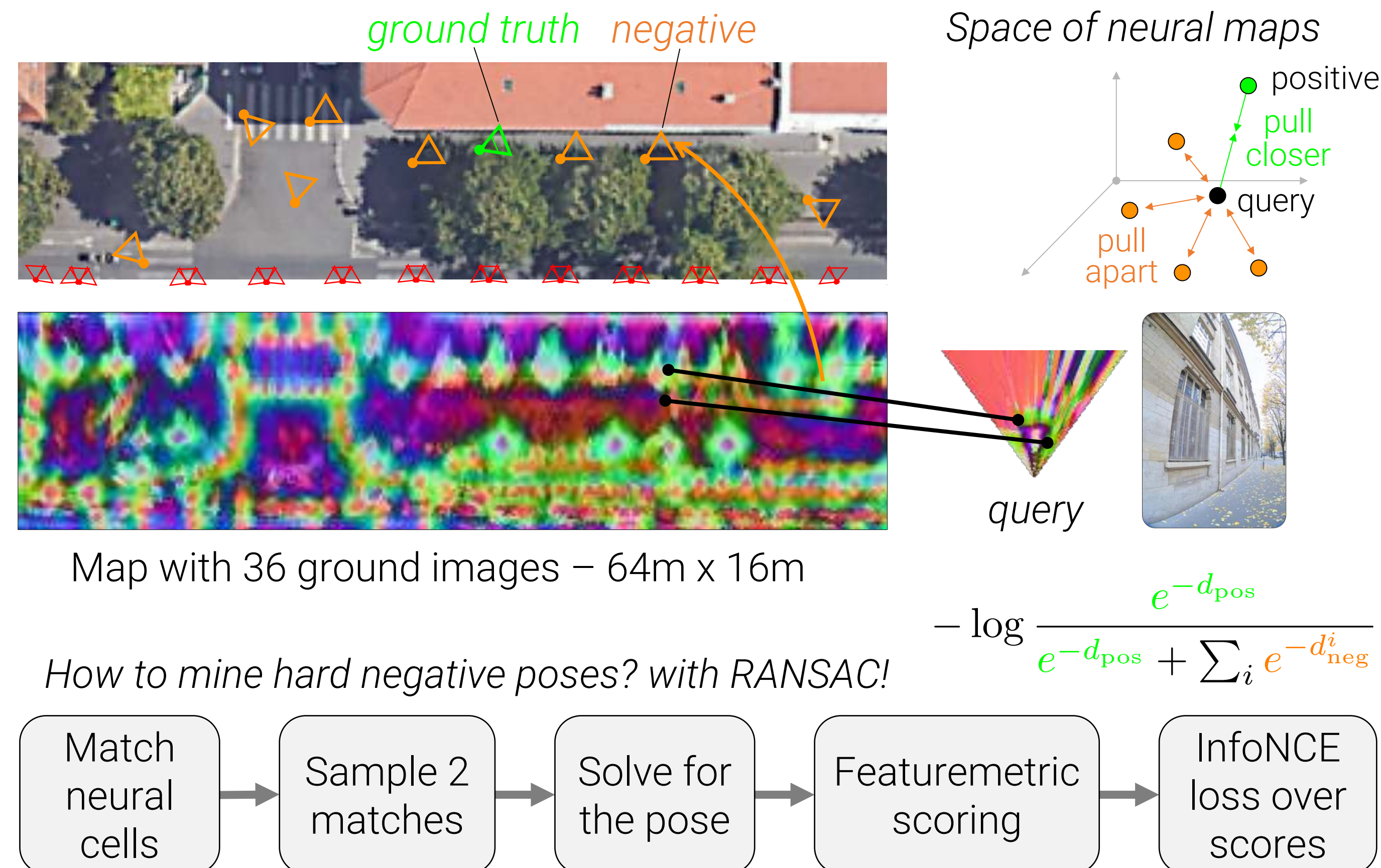
What makes a good map for humans and machines to navigate the world?



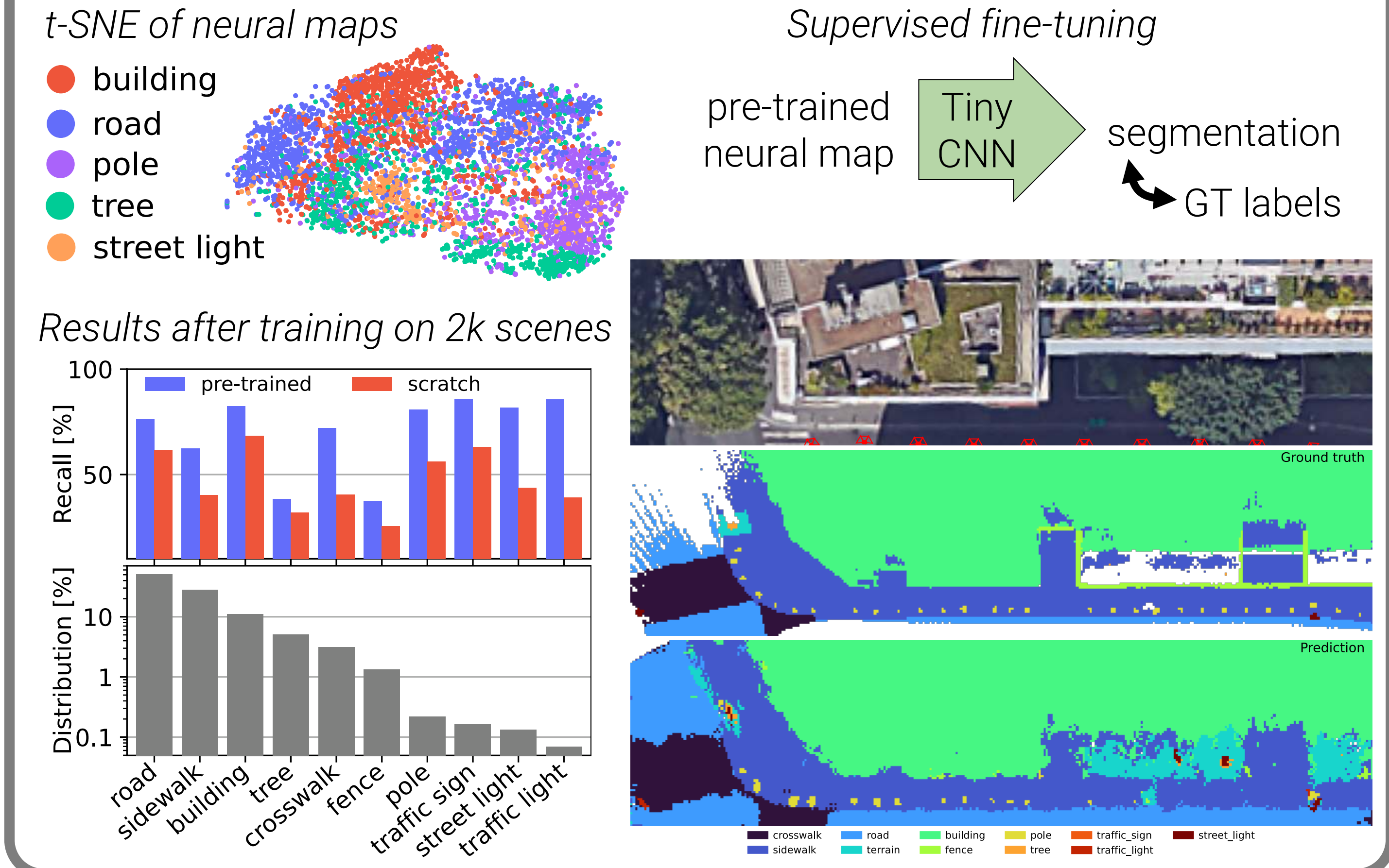
## Architecture



## Self-supervised training

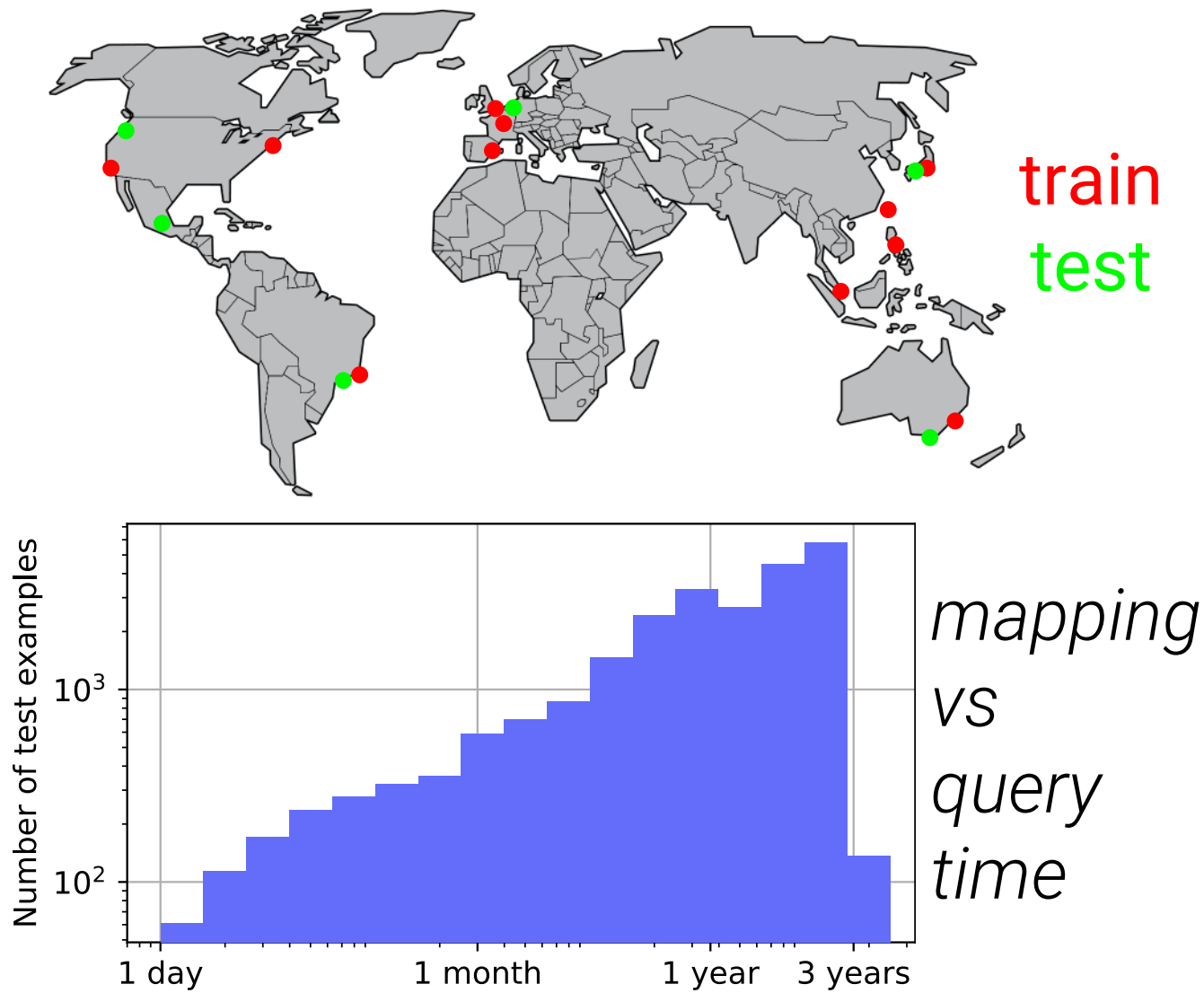


## Effective pre-training for semantic mapping

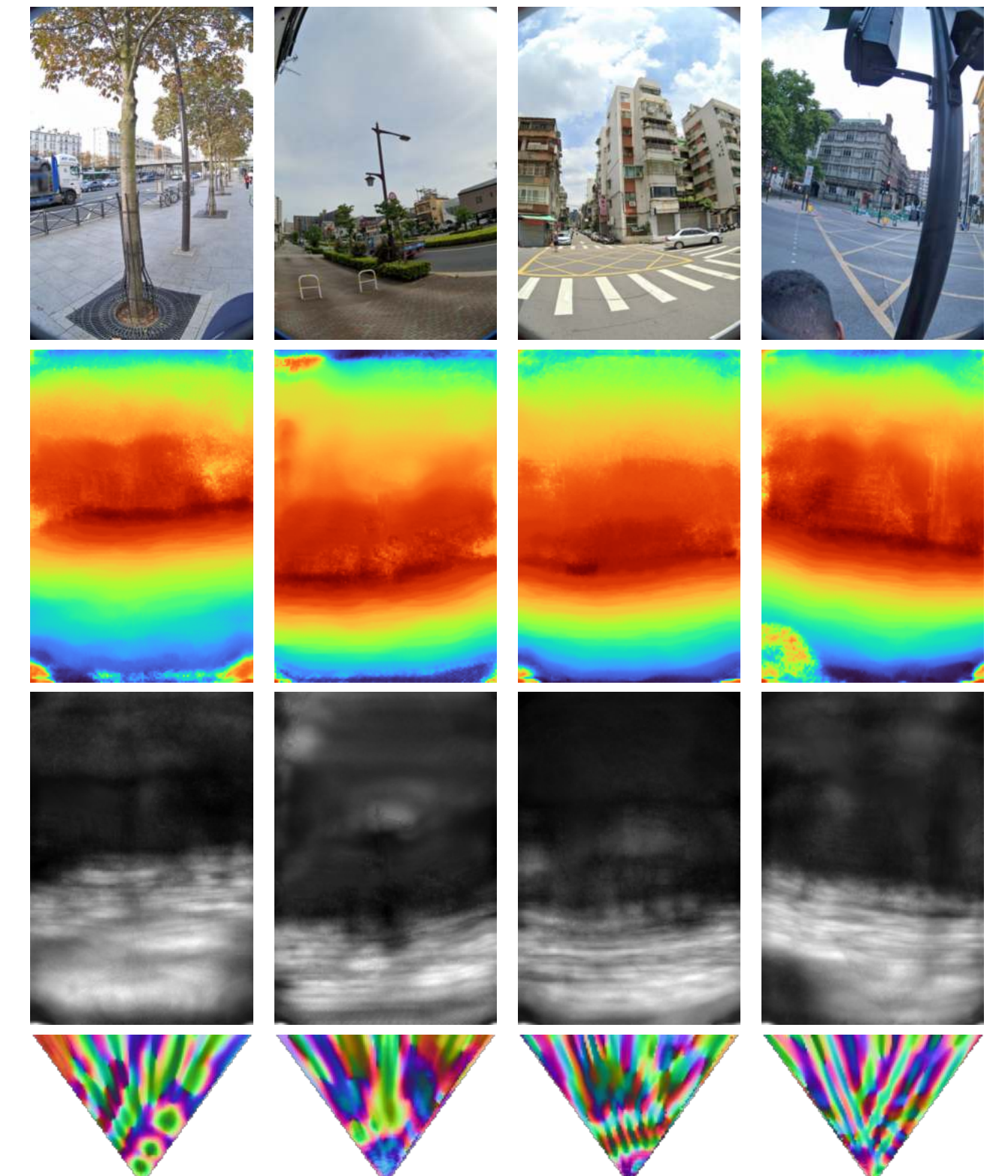


## Visual positioning with challenging views

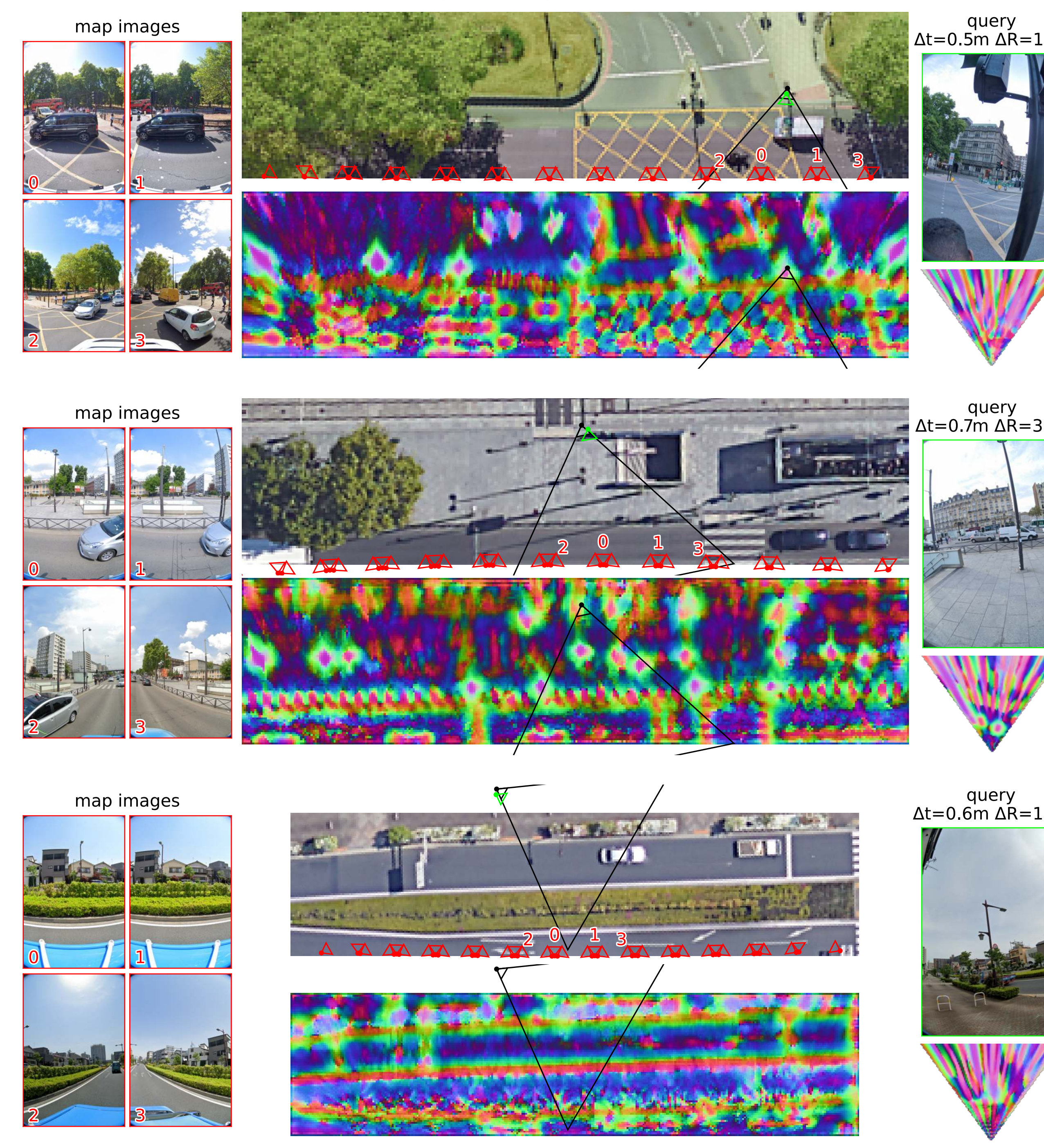
Large scale: 11 cities, 5 continents



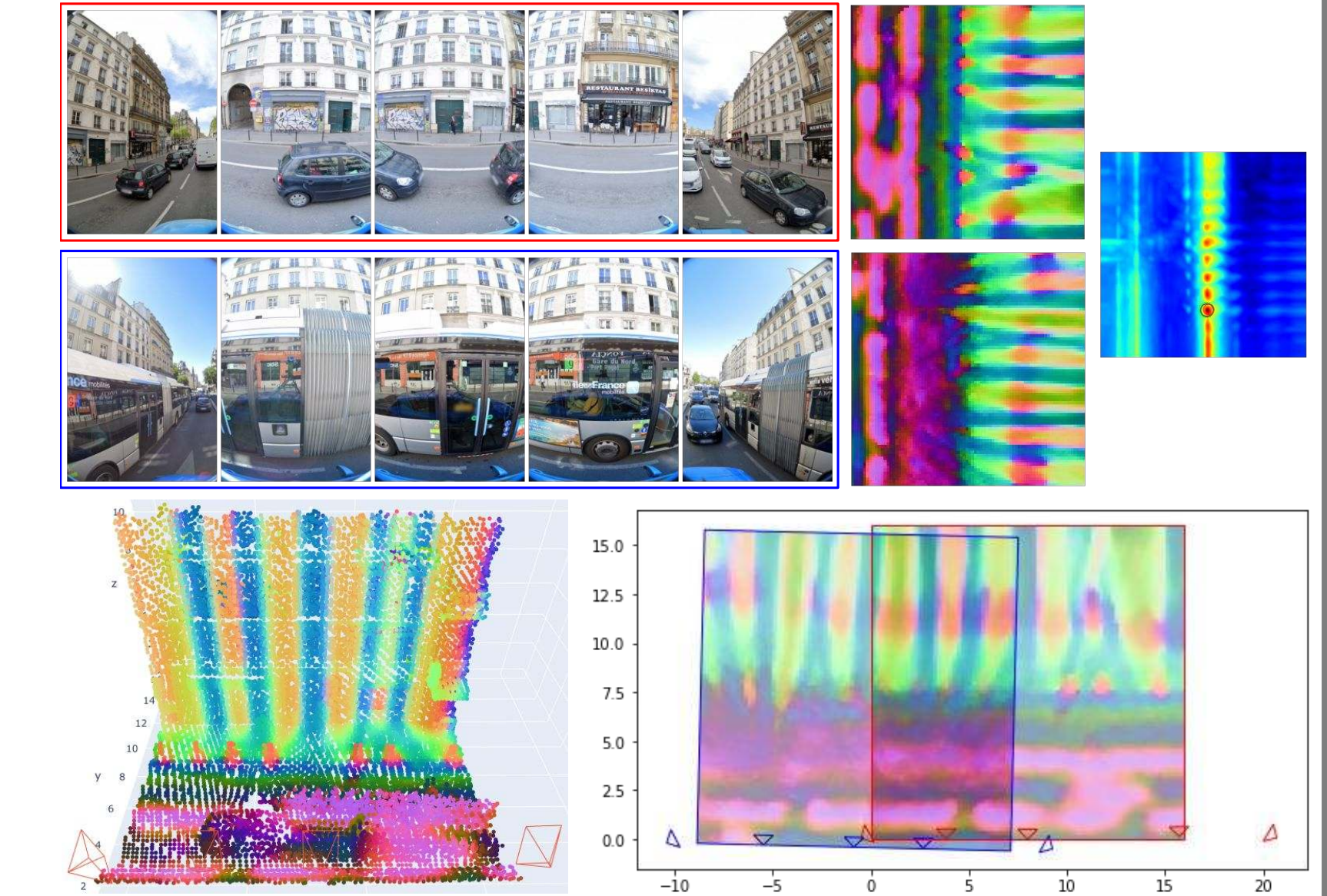
depth from pose supervision



Single-image localization



Sequence-to-sequence alignment



Single-image localization recall [%]

