AutoCompass

Accurate Visual Localization on Public Maps by Learning from Weak Labels

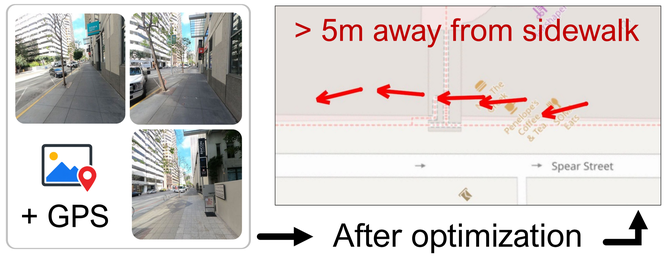
Javier Tirado-Garín Alan Savio Paul Shuai Chen Axel Barroso-Laguna Tommaso Cavallari Daniyar Turmukhambetov Victor Adrian Prisacariu Eric Brachmann

Task: Visual localization on 2D maps

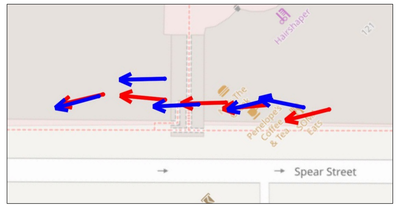
Camera position and orientation ?

Problem: Large-scale datasets are noisy

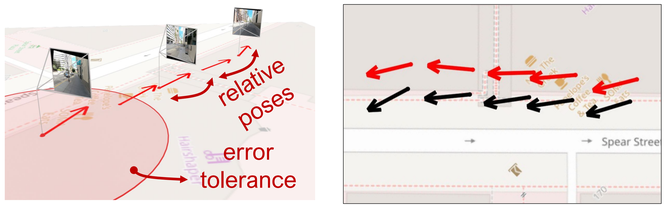
as pseudo ground-truth must be automated



Current models **learn and replicate** these errors



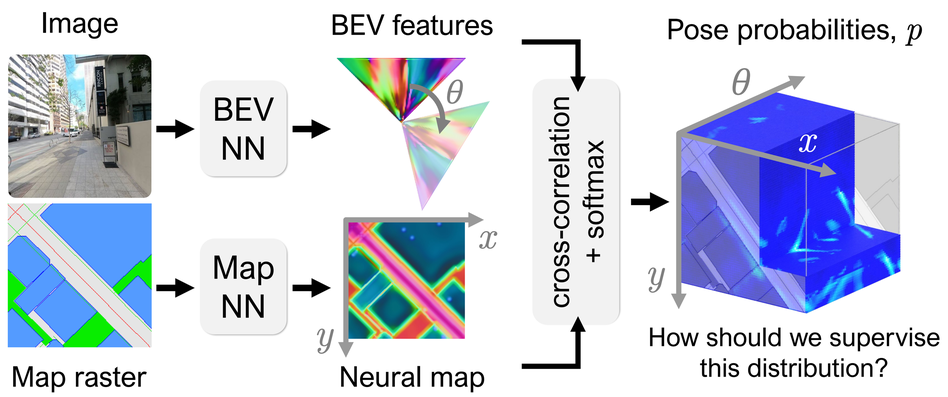
AutoCompass is robust to them



by using tolerances & relative poses

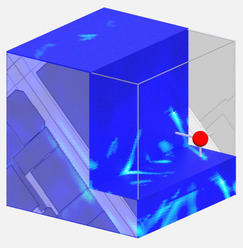
AutoCompass: Supervision with noisy pseudo-GT

Standard neural-map matcher



Standard 3-DoF supervision

$\max\ p(\text{GT pose})$



Not robust
Treats GT as exact

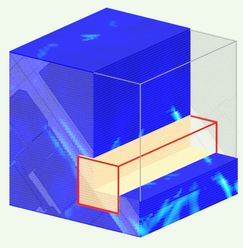
Needs 3-DoF (x, y, θ) labels

AutoCompass 2-DoF supervision

$\max\ p(\text{chunk})$

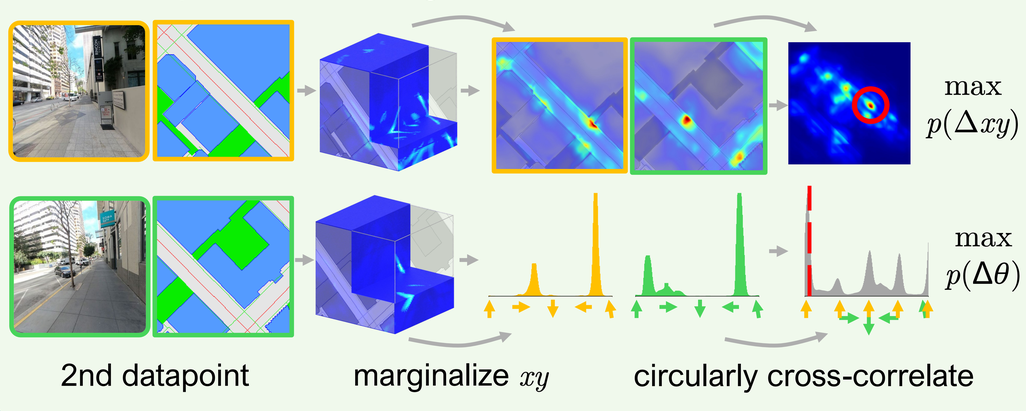
Robust
up-to the size of the chunk

Needs 2-DoF (x, y) labels

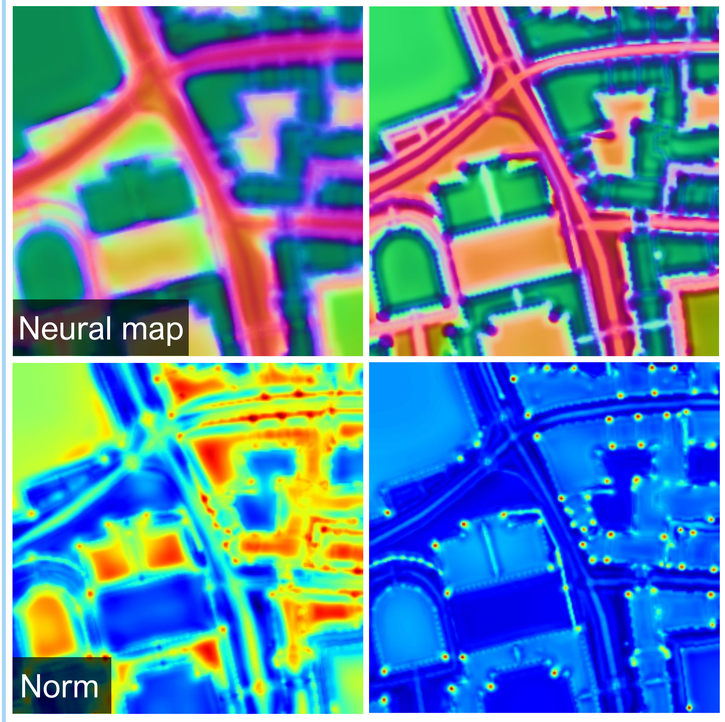


AutoCompass Relative poses supervision

Why?: Relative poses are invariant to shift errors in the GT. **How**:



Sharper features, finer localization



AutoCompass learns to focus on keypoints, such as building corners, rather than on broad regions

AutoCompass sets a new state-of-the-art

2-DoF labels are enough! Rel. Pose labels work best