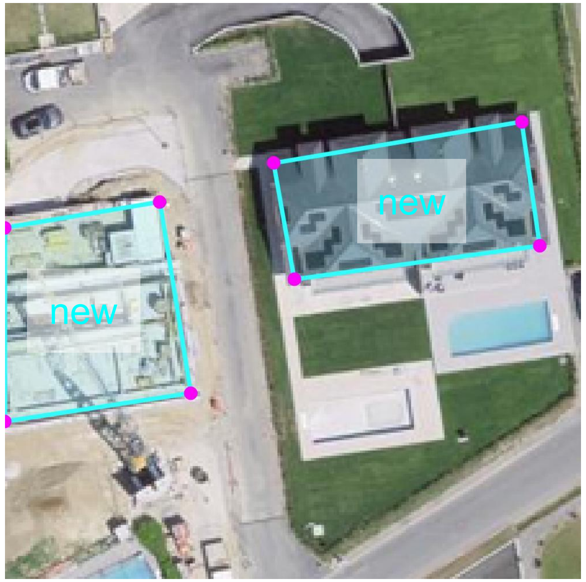
isprs 2026
TORONTO
XXV ISPRS CONGRESS
TORONTO, CANADA | 4-11 JULY 2026
FROM IMAGERY TO UNDERSTANDING

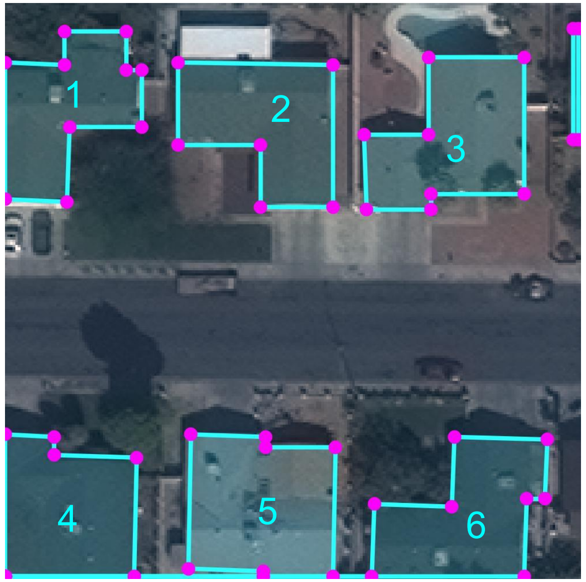
The P^3 Dataset: Pixels, Points and Polygons for Multimodal Building Vectorization

Raphael Sulzer[1,2], Liuyun Duan[1], Nicolas Girard[1], Florent Lafarge[2]
1 Luxcarta Technology, 2 Centre Inria d'Université Côte d'Azur

Introduction – Building outline vectorization

- Many geospatial applications require **vectorized 2D building outlines** e.g. for:
    - Urban planning, change detection, real estate analytics, disaster response, 3D building modelling, …
- Building outlines are commonly extracted from **satellite or aerial images** using computer vision and machine learning techniques





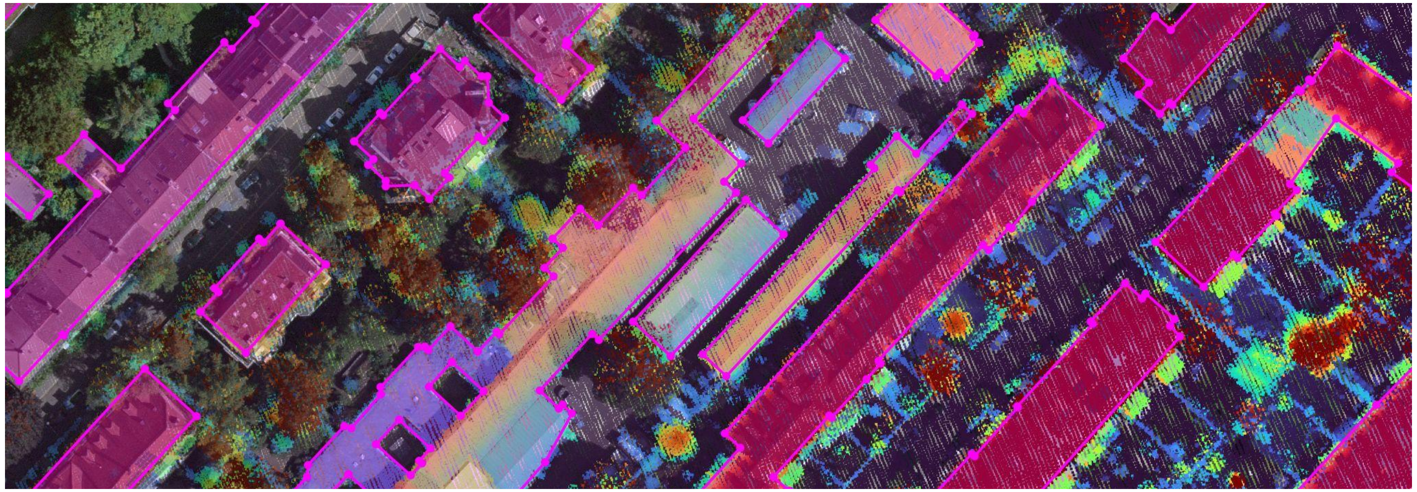
Motivation

- Over the past decade, more and more countries started to provide **aerial LiDAR point clouds** as open data, e.g.:
  - high density
    - typically >10 pts/m^2
  - high accuracy
    - 5–10 cm altimetric
    - 10–20 cm planimetric
  - low temporal sampling
    - every ~5 yrs new acquisition

Dataset	**Country**
LiDAR HD	France
AHN	Netherlands
USGS 3DEP	United States
swissSURFACE3D	Switzerland
Environment Agency LiDAR	United Kingdom (England)
…	…

Problem statement

- Can we predict **building outlines from aerial LiDAR point clouds?**
  - machine learning can probably do it, but…
  - state-of-the-art machine learning models require large and diverse training data
    - i.e. vectorized building outlines + corresponding LiDAR point clouds

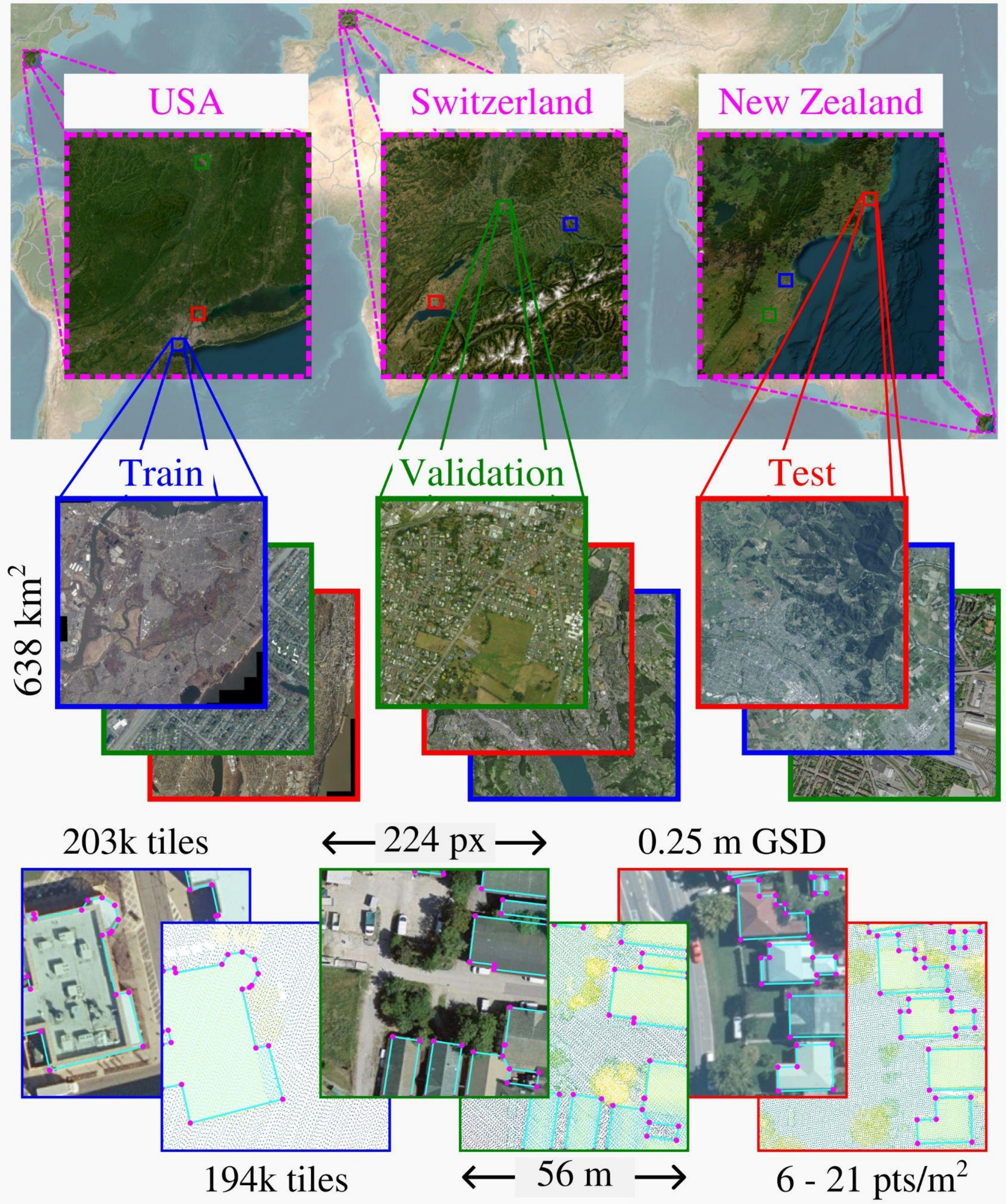


Related datasets

	Dataset	Modality	Ground sampling	Area (km^2)	Tiling	Diversity	Pixel label	2D outline polygon	2D roof wireframe	3D building wireframe	Origin
2D	Inria (Maggiori et al., 2017)	AI	0.3m	810	T+	+	✓				M
	⋮										
	RoofSat (Boyer et al., 2024)	SI	0.3m	14.9	T	-			✓		H
	RoofVect (Hensel et al., 2021)	AI	0.1m	2	SB	-			✓		H
3D	Building3D (Wang et al., 2023)	ALS	30pts/m^2	998	SB	+				✓	M
	City3D (Huang et al., 2022a)	ALS	4-50 pts/m^2	N/A	SB	+				✓	A
	RoofN3D (Wichmann et al., 2018)	ALS	4.7 pts/m^2	1010	SB	-				✓	A
2D+3D	Roof3D (Schuegraf et al., 2023)	AI+DSM	0.3m	22.4	T	-	✓				M
	Potsdam (Rottensteiner et al., 2012)	AI+DSM	0.05m	3.4	T	-	✓				H
	Vaihingen (Rottensteiner et al., 2012)	AI+ALS	0.08m, 4pts/m^2	1.5	T	-	✓				H
	Map2ImLas (Anjanappa et al., 2026)	AI+ALS	0.075m, 10-14pts/m^2	217	T+	-	✓	✓			A

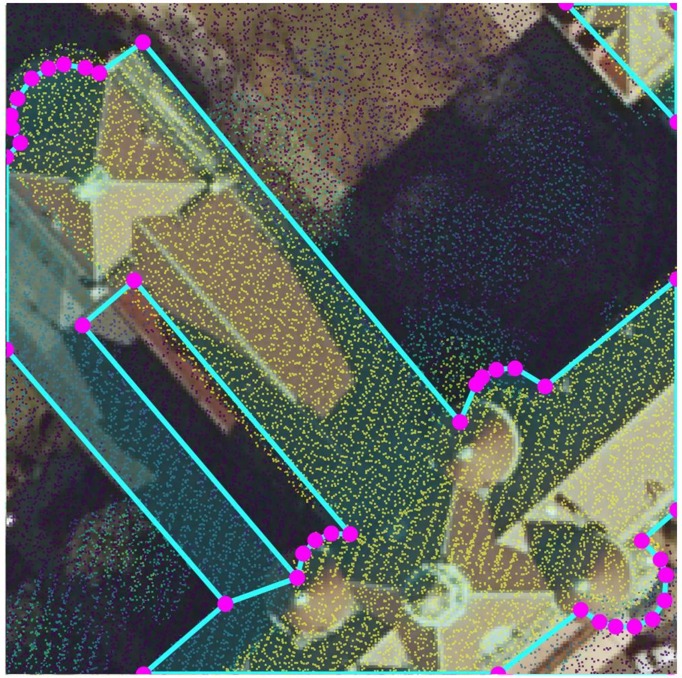
The P³ dataset

- Data from 3 countries
- over 200 km² per country split in train/val/test sets
- Spatially aligned LiDAR and image tiles
- Building vector annotations in standard MS-COCO format
- Ready to use for deep learning models

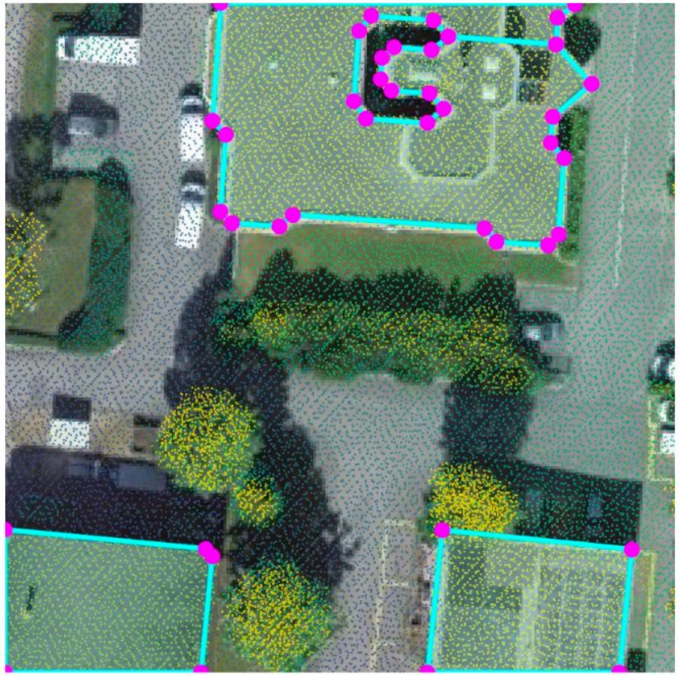


The P^3 dataset — Diverse

- Radiometrically diverse images
- LiDAR point clouds with varying density
- Vectorized building polygons with complex geometries


USA

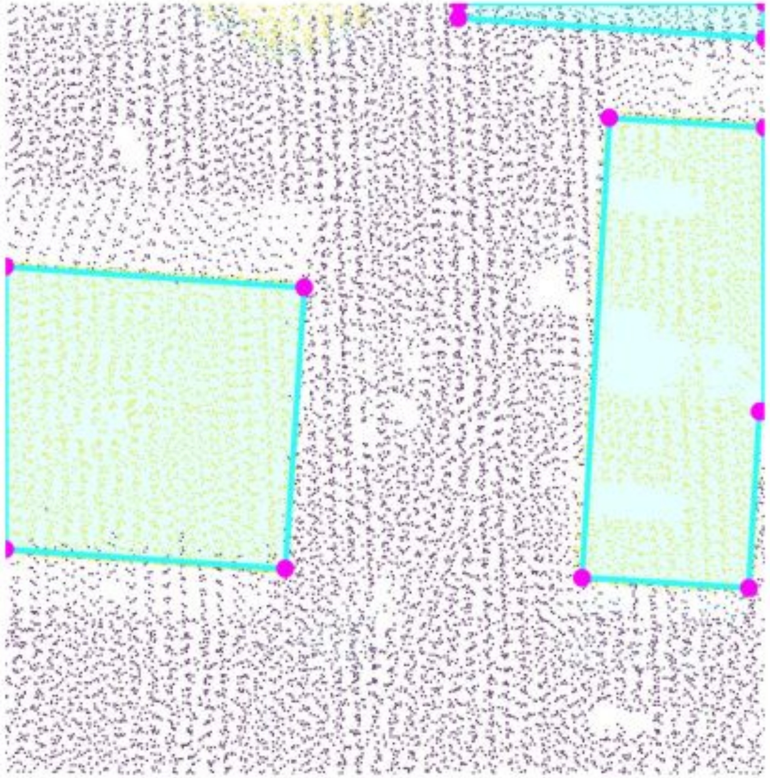

New Zealand


Switzerland

The P³ dataset — Accurate

- Orthogonal projections of building perimeter onto the ground

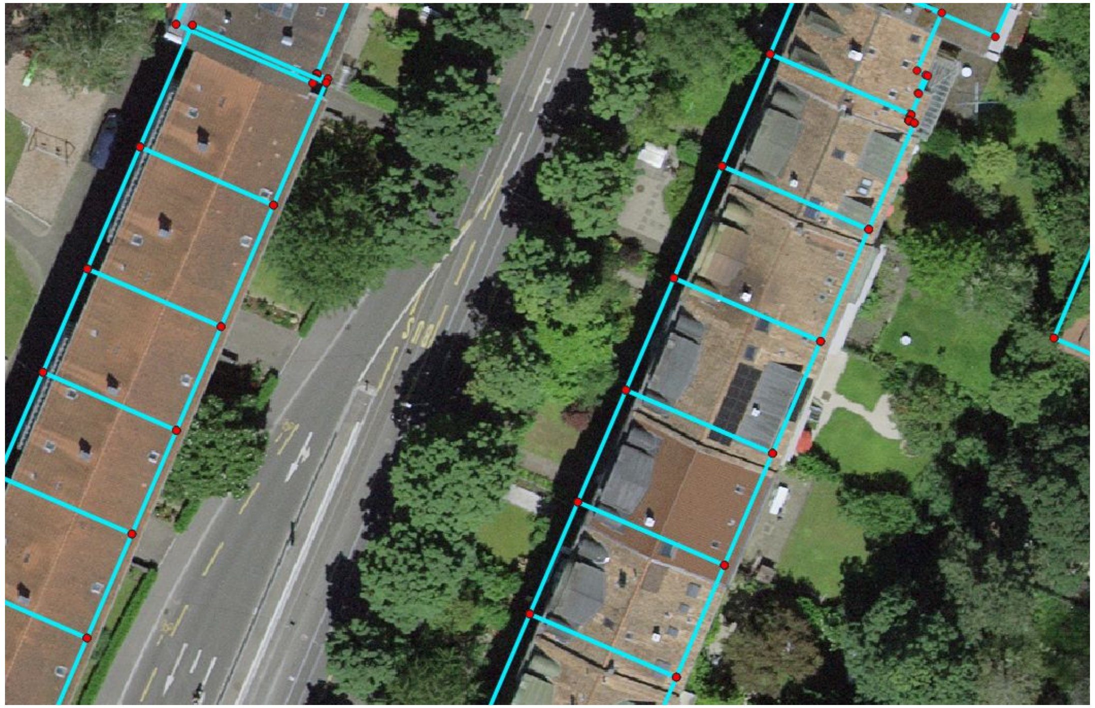
Image



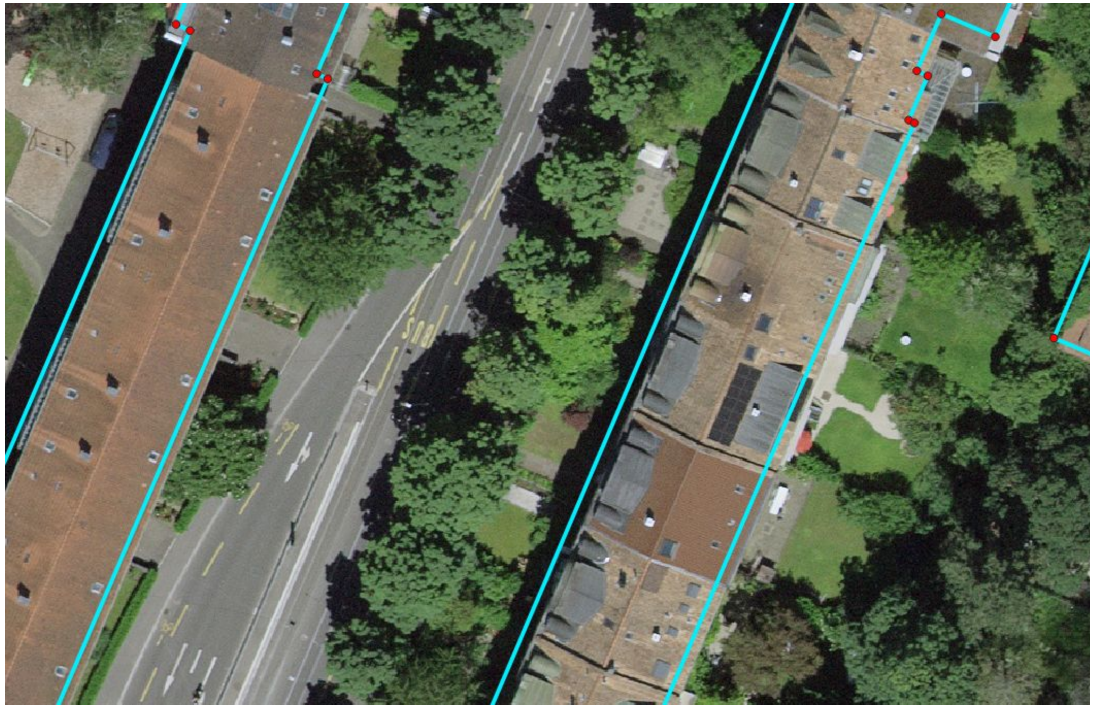
LiDAR

The P^3 dataset – Limitations

- No interior building separation



With separation

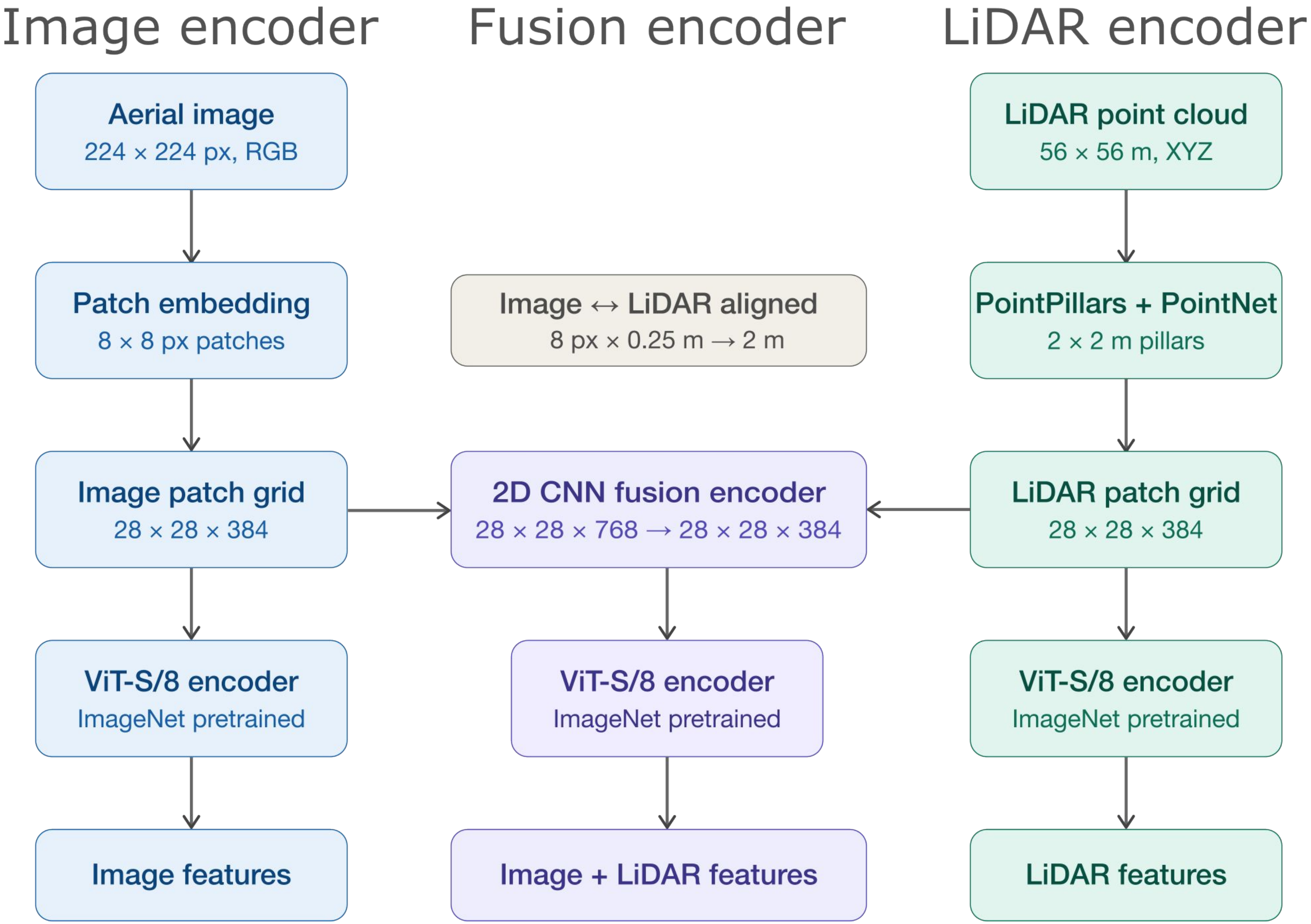


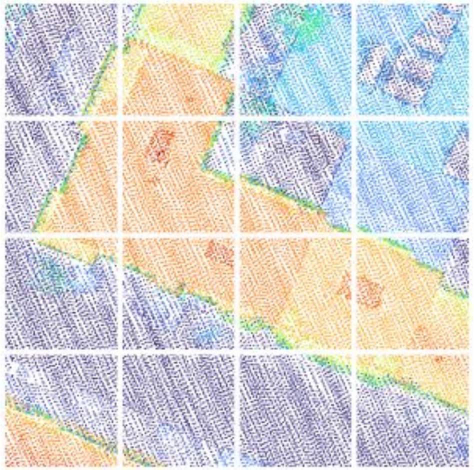
Without separation

Building polygon prediction

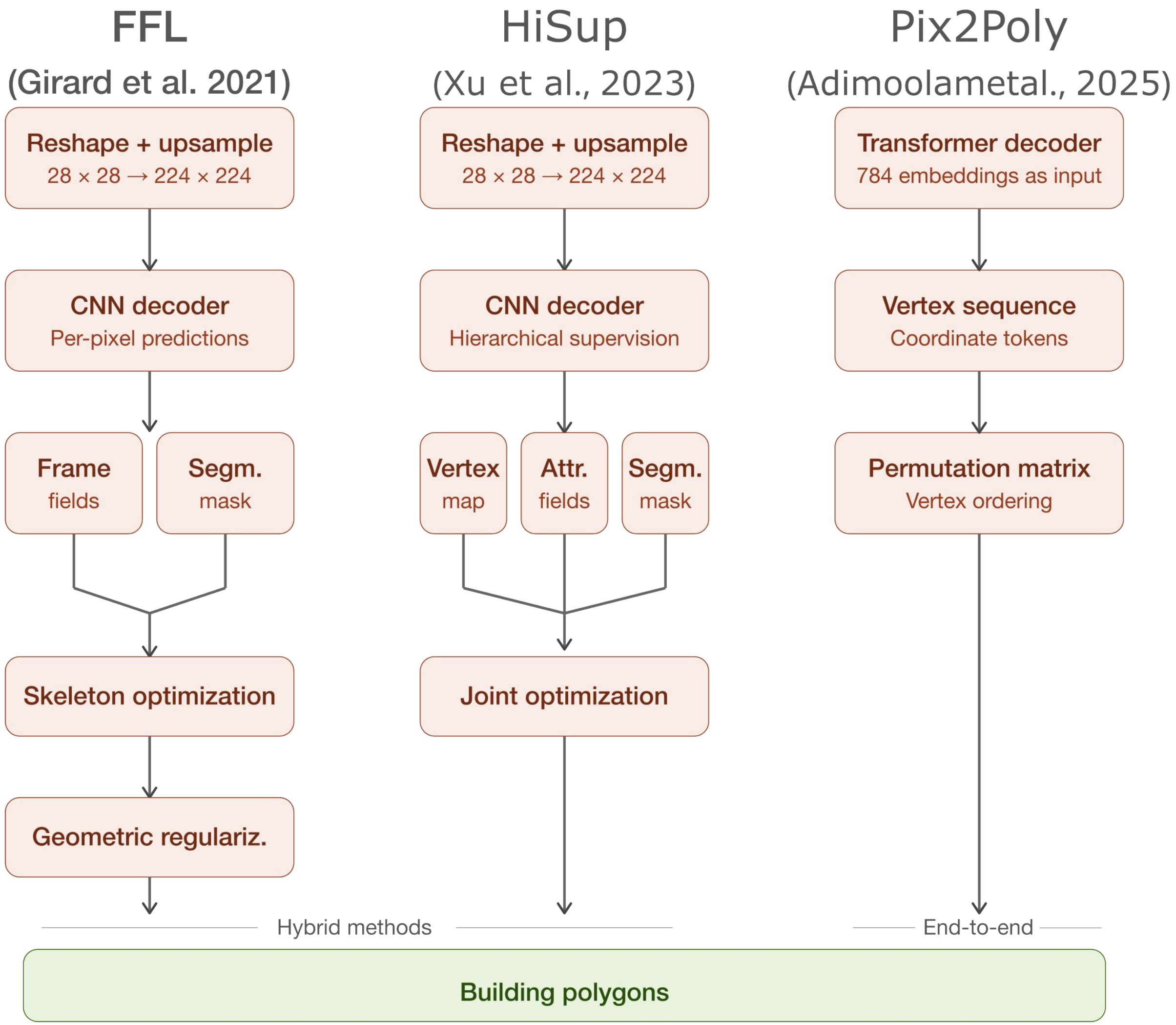
- Train state-of-the-art neural network architectures on the P^3 dataset
  - Frame Field Learning (Girard et al., 2021)
  - HiSup (Xu et al., 2023)
  - Pix2Poly (Adimoolametal et al. 2025)
- Evaluate and compare building polygons predicted from
  - Images only
  - LiDAR point clouds only
  - Both modalities combined

Building polygon prediction (Encoder)



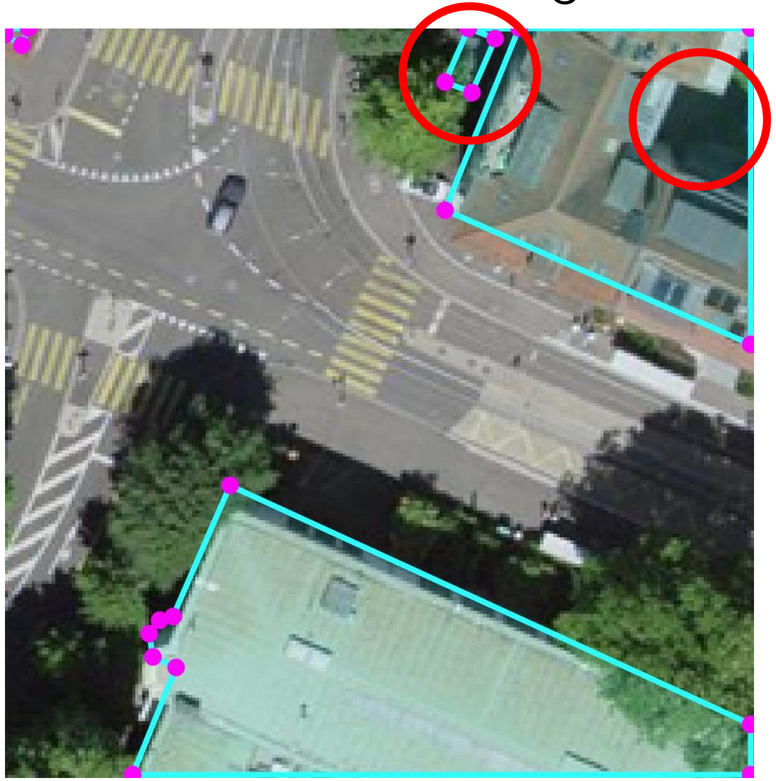


Building polygon prediction (Decoder)

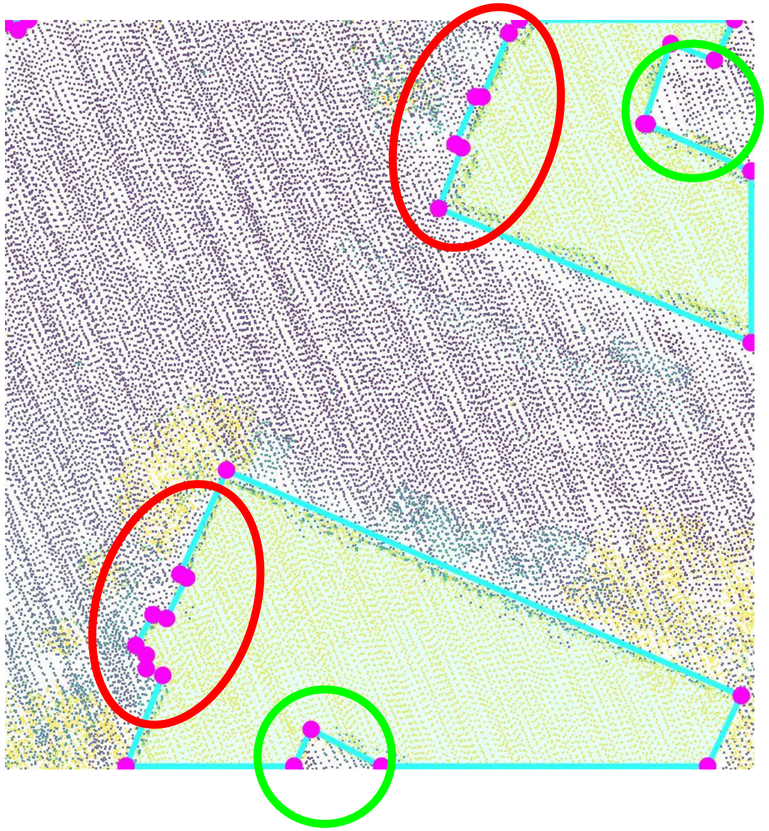


Building polygon prediction (Modality comparison)

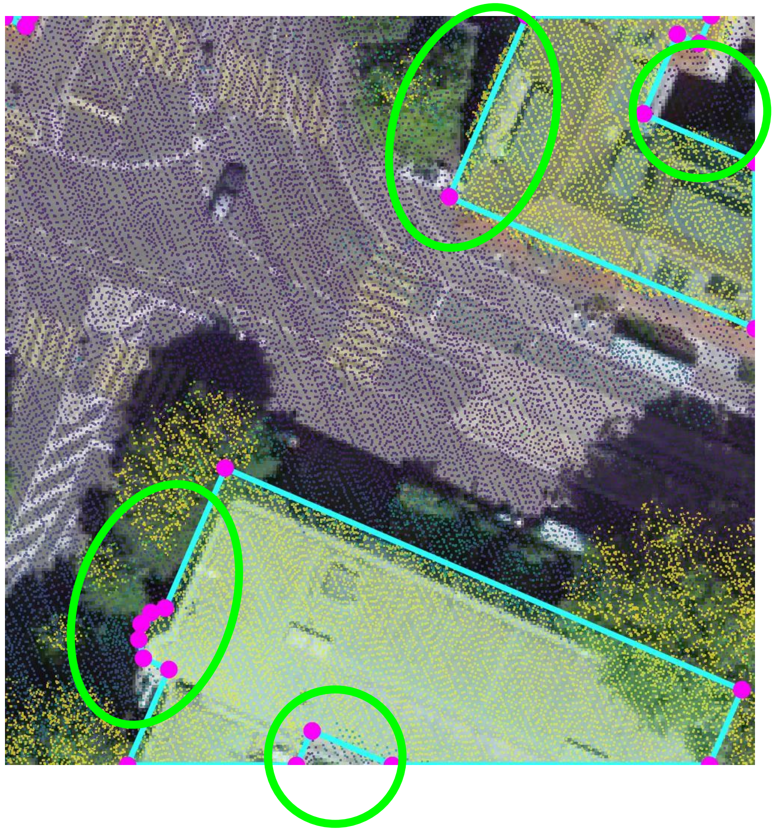
Prediction Image



Prediction Image



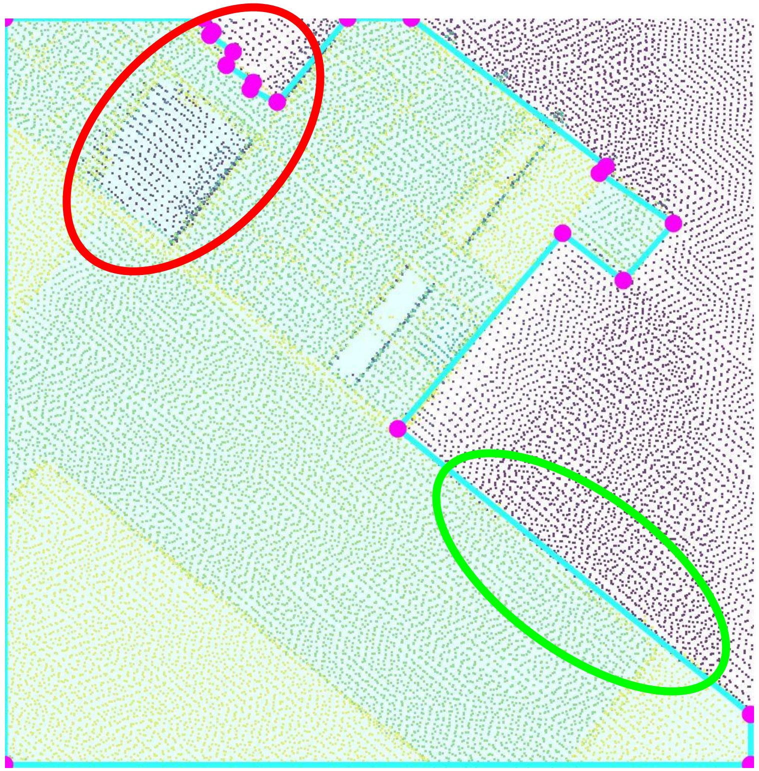
Prediction Fusion



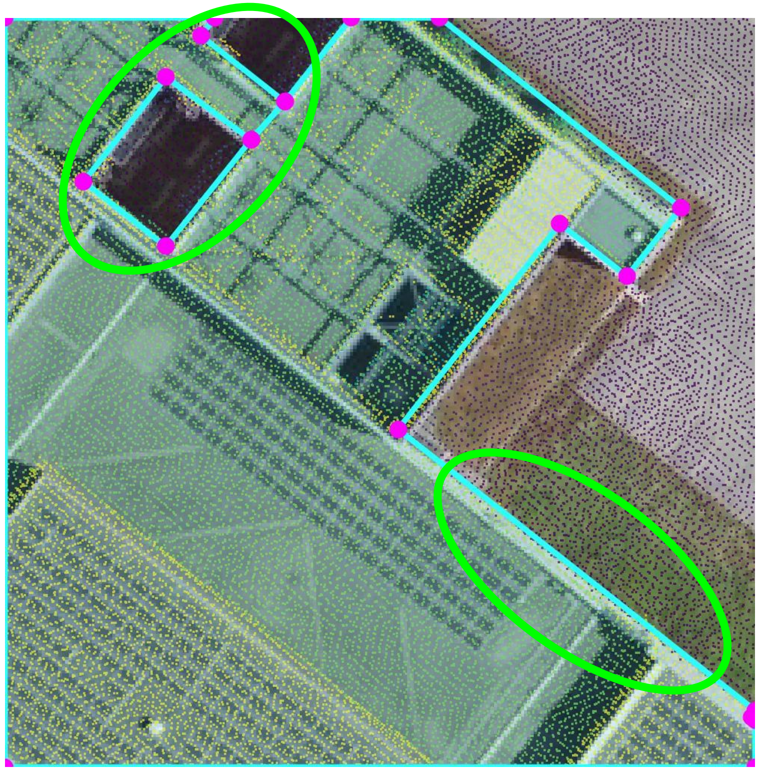
Building polygon prediction (Modality comparison)

Prediction Image

Prediction Image



Prediction Fusion



Building polygon prediction (Robustness)

- LiDAR data helps the model to become more robust to varying radiometry

USA New Zealand Switzerland

Building polygon prediction (LiDAR density)

- Even low density LiDAR allows to recover most building outlines

Ground truth | 1 pt/m^2 | 16 pt/m^2 | 64 pt/m^2 | 128 pt/m^2

Conclusion

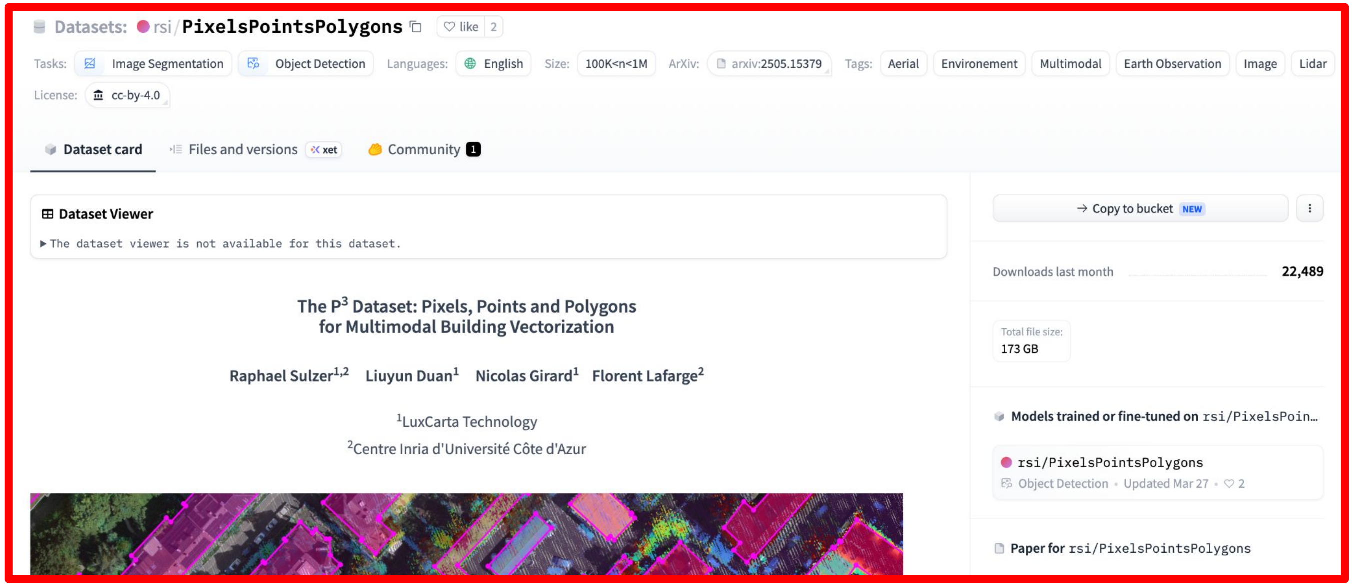
- LiDAR data leads to better building polygon predictions compared to images
- Image and LiDAR data combined leads to the best predictions

		Boundary				*Area*			*Complexity*	*Efficiency*	
Model	**Modality**	**POLIS [m] ↓**	**CD [m] ↓**	**HD [m] ↓**	**MTA [°] ↓**	**AP ↑**	**AR ↑**	**IoU ↑**	**NR=1**	**Time [s] ↓**	**Params ↓**
ViT [11] + **FFL** [14]	**Image**	3	2.7	12	41	0.275	0.46	0.839	0.847	0.532	23.7M
	LiDAR	2.35	1.94	9.66	44.8	0.359	0.545	0.87	0.829	0.582	23.7M
	Fusion	2.14	1.88	8.9	39.7	0.376	0.569	0.877	0.874	0.58	26.4M
ViT [11] + **HiSup** [41]	**Image**	2.46	2.48	11.4	35.2	0.287	0.493	0.85	0.885	0.124	30.8M
	LiDAR	2.05	2.01	9.58	37	0.347	0.54	0.87	0.882	0.165	30.8M
	Fusion	1.91	1.9	8.94	35.2	0.355	0.568	0.872	0.89	0.129	33.5M
ViT [11] + **Pix2Poly** [2]	**Image**	2.46	2.5	10.8	34.3	0.317	0.492	0.845	0.906	1.44	31.9M
	LiDAR	1.88	1.9	8.5	34.1	0.379	0.552	0.869	0.913	1.15	31.9M
	Fusion	1.8	1.82	8.16	33.4	0.398	0.578	0.87	0.915	1.15	34.6M

Table 2: **Modality ablation**. We compare baseline models trained and tested on different modalities of the Switzerland subset. For each metric, we highlight the best and second best scores.

Try it yourself

- The dataset, code and pretrained models are available on github and huggingface
  - https://huggingface.co/datasets/rsi/PixelsPointsPolygons
  - https://github.com/raphaelsulzer/PixelsPointsPolygons



3D version of the dataset (with LOD2 building models) coming soon…