Data Augmentation for Multi-Temporal Rooftop Solar Panel Segmentation in Lithuanian Orthophotos

Justinas Lekavičius, Valentas Gružauskas

Artificial Intelligence Methods Lab
Institute of Computer Science, Vilnius University
Corresponding author: justinas.lekavicius@mif.vu.lt

Artificial Intelligence Methods Lab

Follow us: aiml.lt

Abstract

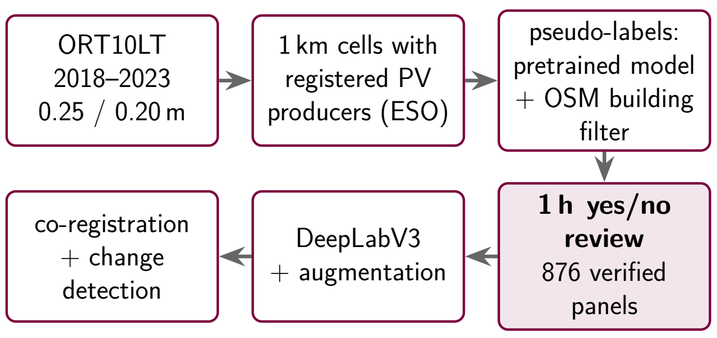
Building-level rooftop PV data are not openly available in Lithuania. Using national orthophotos (2018–2023), we test whether classic and generative (pix2pix) augmentation improves DeepLabV3 segmentation accuracy and consistency between years. A one-hour review of pseudo-labels raised IoU from 0.44 to **0.71**; GAN-refined copy-paste added **+0.019 IoU** (best inference), and hard negatives **+0.050 IoU** under dark shadows. Co-registration raised multi-year consistency to **0.874**; new installations were detected with **92.7 %** precision.

Problem

Monitoring rooftop PV growth needs building-level data, but Lithuanian open data report electricity producers only per 1×1km grid cell. Orthophotos can fill this gap: each year covers part of the country, so most places are imaged at least twice in 2018–2023. But labelled data exist mainly for other countries (France, China), and years differ in illumination, shadows, season and geometry.

Objective: test whether data augmentation improves segmentation accuracy and multi-year consistency, and detect new installations per grid cell.

Dataset construction



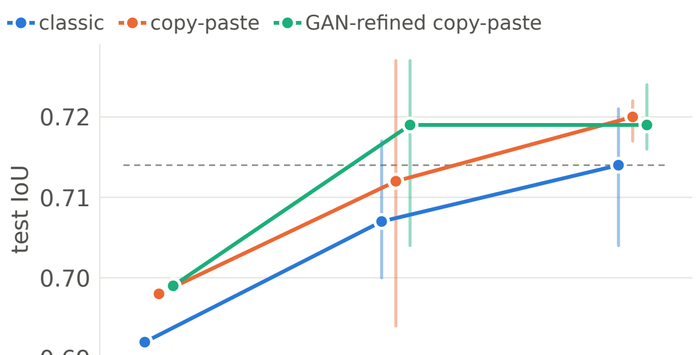
512×512 px patches at 0.2 m, year pairs 2018→2021, 2019→2022, 2020→2023, 2659 training patches.

Test set: 197 locations, 394 hand-labelled images from held-out cells.

Weak labels (pretrained-model masks kept on OpenStreetMap buildings) were only **7 %** correct; one hour of accepting or rejecting them gave the verified labels:

Training labels	IoU	Prec.	False m²/img
Pretrained model only	0.31	0.73	–
Weak labels	0.44	0.52	45.6
Verified labels	**0.71**	**0.91**	**1.1**

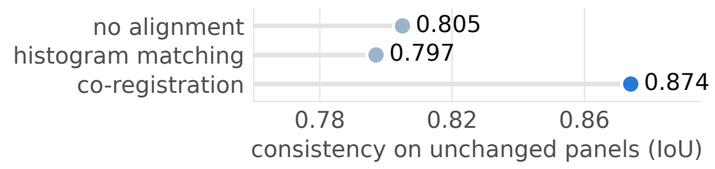
How many verified labels are needed?



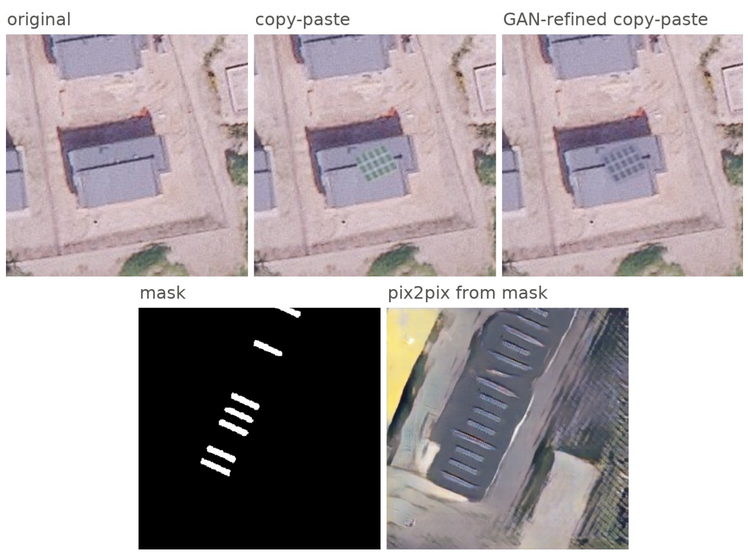
25 % of the labels already gives IoU 0.69. With 50 %, GAN-refined copy-paste (0.719) matches classic with all labels (0.714, dashed). Mean of 3 runs; bars: run range.

Registration and normalisation

The two years are offset by a median **1.05 m** (up to 4.1 m). Histogram matching did not help, but co-registration removed this offset (chart: classic model, 47 test locations).



Augmentation strategies



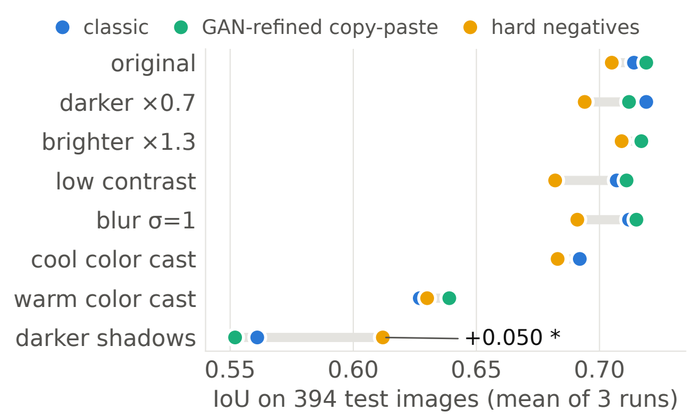
- **Classic:** flips, rotations, color jitter (all models).
- **Copy-paste:** verified panels pasted onto roofs.
- **GAN-refined copy-paste:** pix2pix re-renders the pasted panel to fit the roof.
- **pix2pix from masks:** whole image from a mask (earlier work: +2 % IoU).
- **Hard negatives:** rejected shadows and dark roof patches pasted as background.

Accuracy and multi-year consistency

	Single runs		Best inference	
Augmentation	**IoU**	**Cons.**	**IoU**	**Cons.**
classic	0.714	0.795	0.715	0.790
copy-paste	0.720	0.804*	0.728*	**0.805***
GAN-refined copy-paste	0.719	0.784	**0.734***	0.799*
hard negatives	0.705*	0.791	0.726*	0.801*
pix2pix from masks†	0.689	0.750*		

3 runs per setting. Cons. (consistency): IoU between the years on unchanged panels. Best inference: 3-run average + flips/rotations + tuned threshold. * differs from classic (95 % CI); † one run.

Robustness to illumination

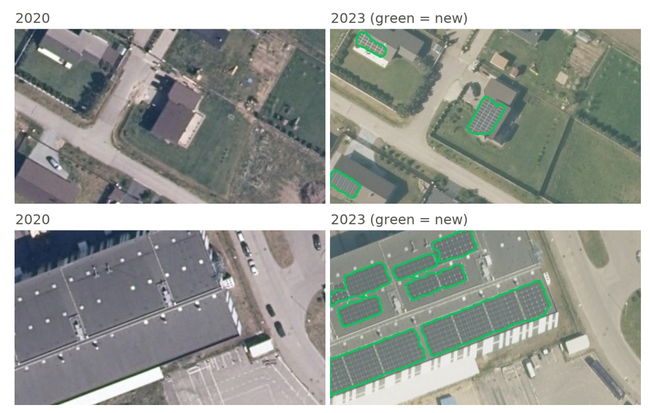


Test images were darkened, brightened, recolored, and blurred. A warm color cast and **darker shadows** lowered IoU the most. Generative augmentation did not make the model less sensitive to them, but **hard negatives recovered +0.050 IoU** [0.028, 0.069] under darker shadows. Year differences matter less than expected: trained on 2021–23 only, it scores 0.687 on 2018–20 (all years: 0.703).

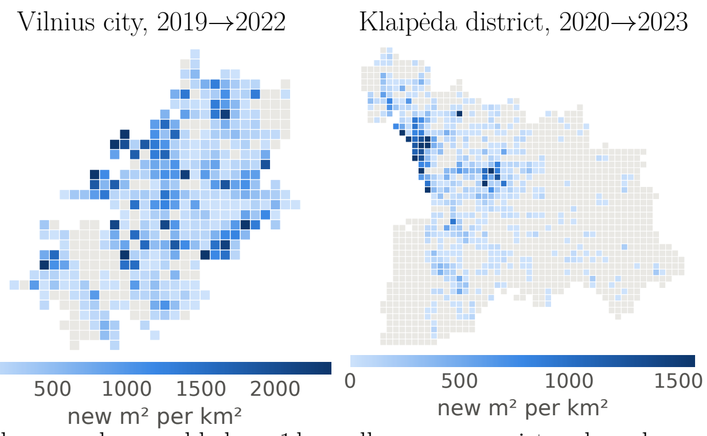
Change detection

Each 1 km cell is segmented in both years and co-registered. A **new installation** must be:

- absent in the earlier year within 1.6 m (roof displacement),
- not even weakly detected there (model score > 0.3; season, light),
- ≥ 8 m² and mostly on an OpenStreetMap building.



Grid maps: new rooftop PV



Color: panel area added per 1 km cell; grey: no registered producers or imaged once.

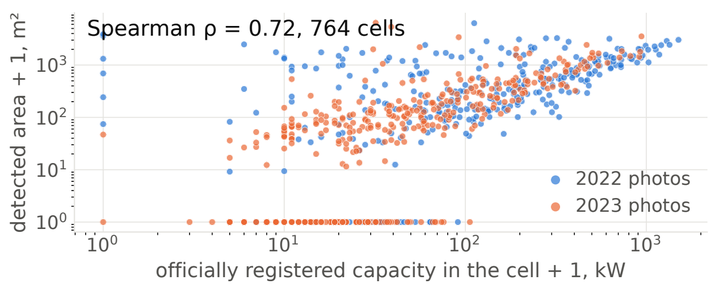
PV area tripled in Vilnius and quadrupled in Klaipėda district in three years:

Region (cells)	Area, m²	Growth	New
Vilnius (316)	71.9k → 215.3k	+199 %	3144
Klaipėda d. (448)	24.6k → 108.6k	+342 %	1642

Validation

Are detections real? Of 96 random detected new installations checked on before/after images, **89 (92.7 %)** were real (95 % confidence interval 86–96 %).

Do they match official data? Detected panel area vs registered capacity (kW) per cell: Spearman correlation **0.72** (764 cells).



Outliers: company roofs not assigned to cells (above the trend); panels installed later, missed or ground-mounted (bottom row).

Conclusions

- Clean labels matter most; generative augmentation helps when it works on **objects**, not on **whole-image** style.
- Co-registration and tolerance checks make yearly masks comparable for change maps.
- **Future work:** Open dataset, inference on all 10827 cells; fusion with meteorological and in-situ data (multimodal).

This project has received funding from the Research Council of Lithuania (LMTLT), agreement No S-ITP-25-3.