LBS 2023
18th Conference on Location Based Services
November 20-22, Ghent, Belgium

Segment Anything Model for Building Extraction with Few Samples in Settlements of Forcibly Displaced People from High-Resolution Satellite Imagery

Yunya Gao1, Hui Zhao2
1Christian Doppler Laboratory for GEOHUM, Department of Geoinformatics – Z_GIS,
Paris Lodron Universität Salzburg, Austria
2 School of Natural Resources and Environment, University of Florida, USA

This study was partly funded by the Christian Doppler Laboratory for Geospatial and EO- Based Humanitarian Technologies, Department of Geoinformatics – Z_GIS, Paris-Lodron-University of Salzburg, Austria

Content

- Social Background
- Technical Background
- Methodology
- Results and Discussion
- Conclusions

Social Background

Forcibly Displaced People (FDP)

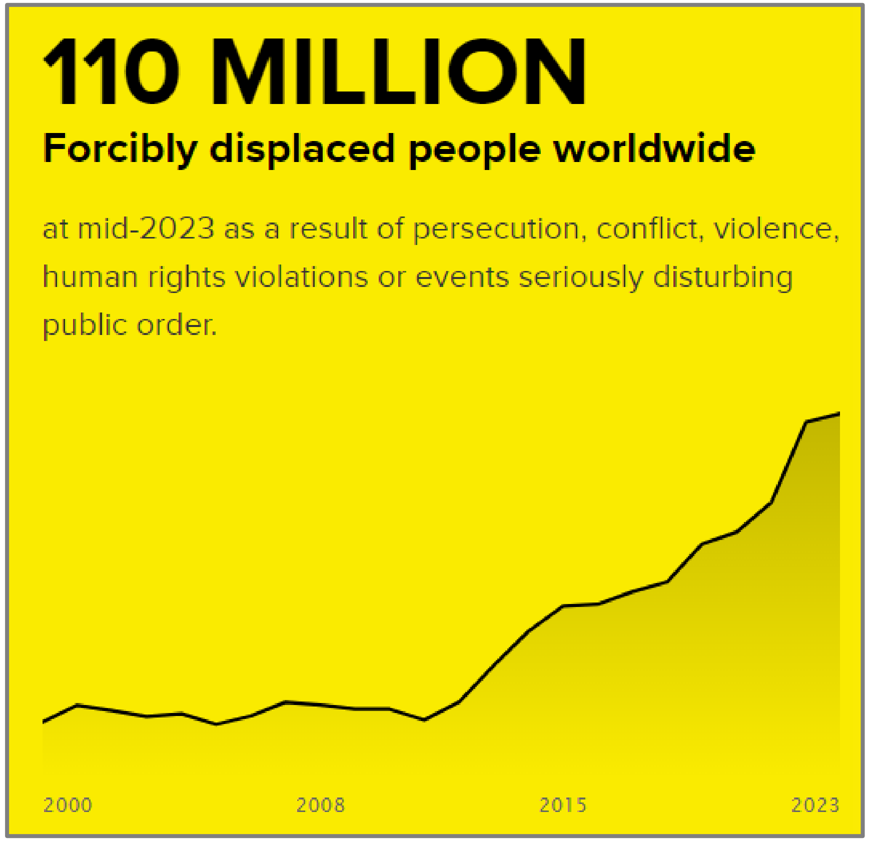
- Internally Displaced People (IDP) → Remain within their country
- Refugees → Cross country border, recognized and protected under international law
- Asylum-Seekers → Cross country border, applying for international protection

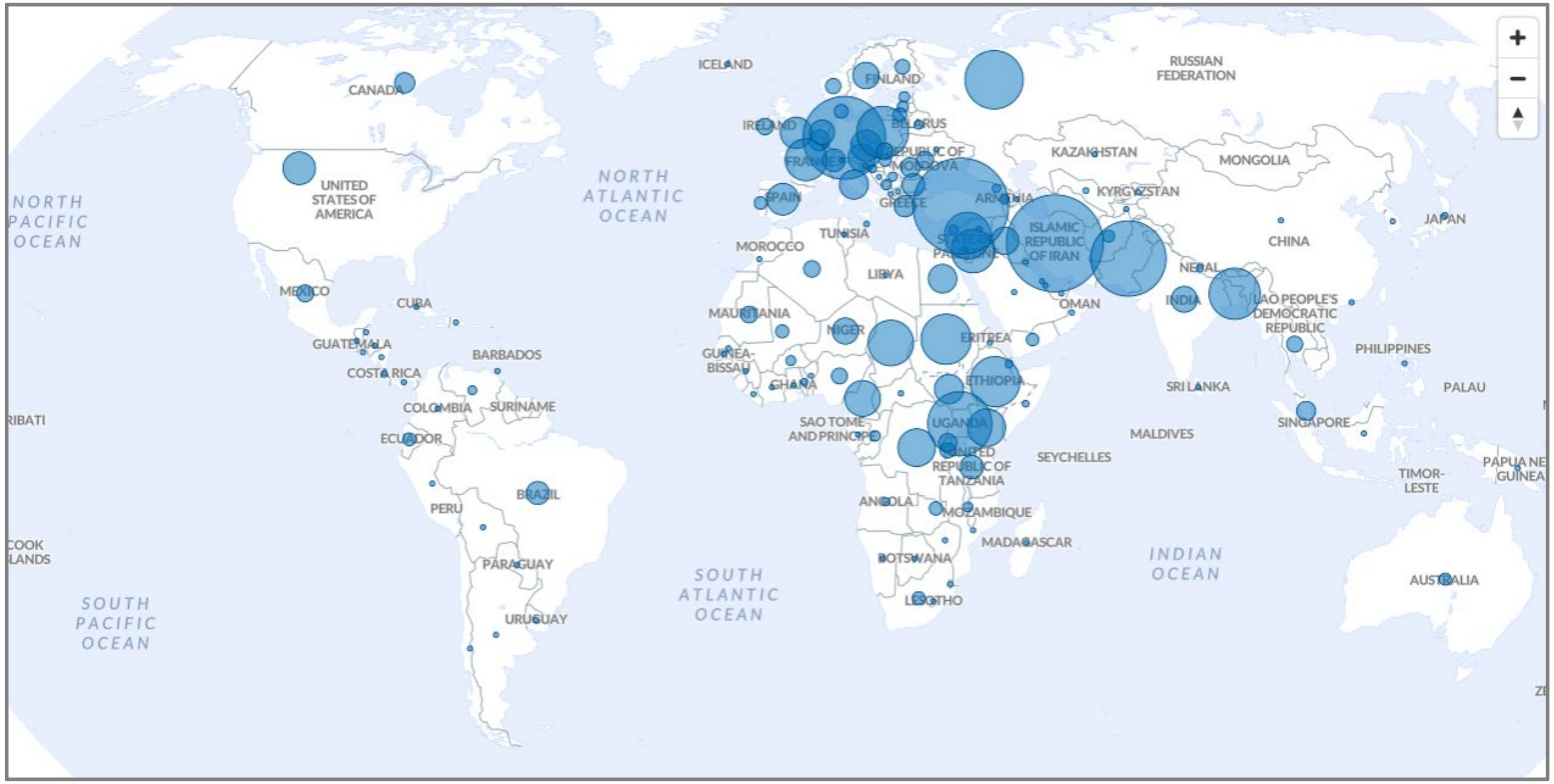
Reasons

- Persecution
- Conflict
- Generalized violence
- Human rights violations
- Natural disasters

Social Background

UNHCR
The UN Refugee Agency





The map shows the number of refugees that UNHCR protects and/or assists.

GEOHUM
Christian Doppler Laboratory
for geospatial and EO-based humanitarian technologies

Social Background

- Lack living resources
- Lack access to mental health care
- Family separation
- Dangerous travels
- Exposure to violence and abuse
- Struggle with securing sufficient water and accessing toilets, hygiene
- 80% of FDP depend on wood for cooking and heating, result in deforestation and elevate risks for women and girls

2 ZERO HUNGER

End hunger, achieve food security and improved nutrition and promote sustainable agriculture

7 AFFORDABLE AND CLEAN ENERGY

Ensure access to affordable, reliable, sustainable and modern energy for all

3 GOOD HEALTH AND WELL-BEING

Ensure healthy lives and promote well-being for all at all ages

10 REDUCED INEQUALITIES

Reduce inequality within and among countries

6 CLEAN WATER AND SANITATION

Ensure availability and sustainable management of water and sanitation for all

SDG Indicator 10.7.4: "Proportion of the population who are refugees, by country of origin"

SDGs advocate providing adequate living resources to FDP and their host communities

Social Background

Benefits of Updated Building Footprints within Refugee/IDP settlements from RS:

- Estimating population
- Tracking the demographic dynamics
- Facilitating better management, logistics planning
- Preparedness and prevention of conflicts

MEDECINS SANS FRONTIERES

UNHCR The UN Refugee Agency

GEOHUM Christian Doppler Laboratory

Technical Background

2018

Polygonal Mapping by HiSup (Xu et al. 2023)

Image Segmentation

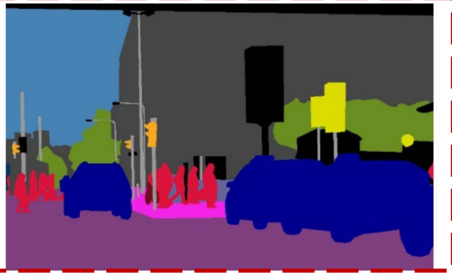
Things
- Countable objects
- e.g., persons, dogs, cars

Stuff
- Amorphous regions of similar material, which is uncountable
- e.g., sky, ocean, grass

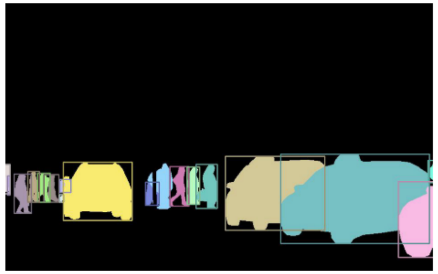
Kirillov et al. (2018)

How to deal with Things and Stuff

Semantic Segmentation



Instance Segmentation



Panoptic Segmentation

(1) Binary semantic segmentation of buildings within FDP settlements

(2) Convert to Shapefile polygons for humanitarian operations

GEOHUM
Christian Doppler Laboratory
for geospatial and EO-based humanitarian technologies

Technical Background

Label Efficient Approaches

Weakly supervised learning
Self-supervised learning
Zero-shot learning
One-shot learning
Few-shot learning
Meta learning
Transfer learning
……

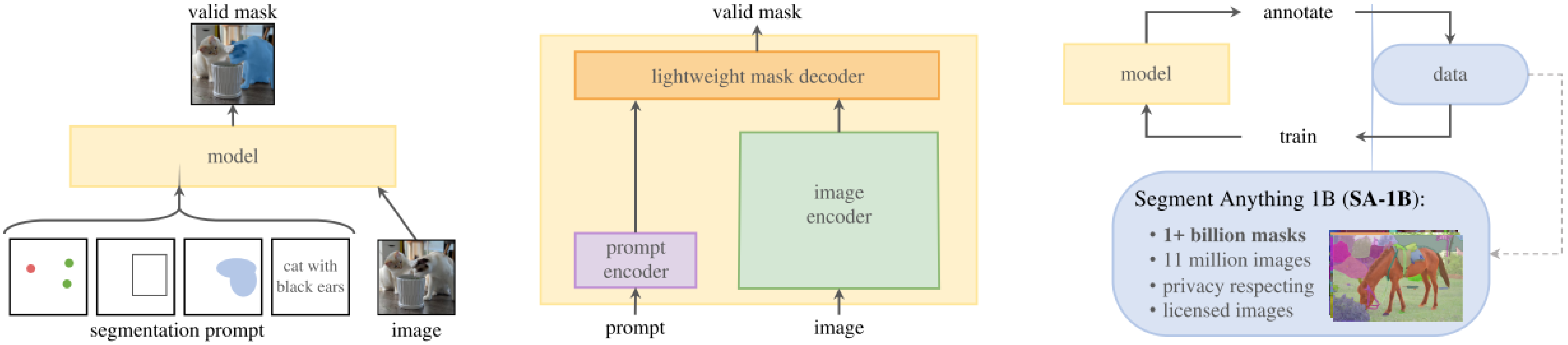
Fine-tuning Foundation Models
Segment Anything Model

Meta

Sint-Baafshuis 9000,
Biezekapelstraat 2, 9000 Gent
Image source: Google Map

Technical Background

Kirillov et al. (2023)



(a) **Task**: promptable segmentation (b) **Model**: Segment Anything Model (**SAM**) (c) **Data**: data engine (top) & dataset (bottom)

SAM executes promptable segmentation, which is different from semantic segmentation:

1) the masks produced by SAM are unlabeled
2) SAM operates based on prompts

GitHub

Here are 140 public repositories matching this topic...

Technical Background

One potential reason for SAM's poor performance in satellite image segmentation could be the unequal distribution of image data present in SAM's training dataset

Text2Seg

Zhang *et al.* (2023)

- Tackle semantic segmentation tasks in satellite imagery
- Integrate Grounding DINO and CLIP
- **Demand minimal effort yet result in a lot of errors in building extraction**

SAM for remote sensing: From Zero to One Shot

Osco *et al.* (2023)

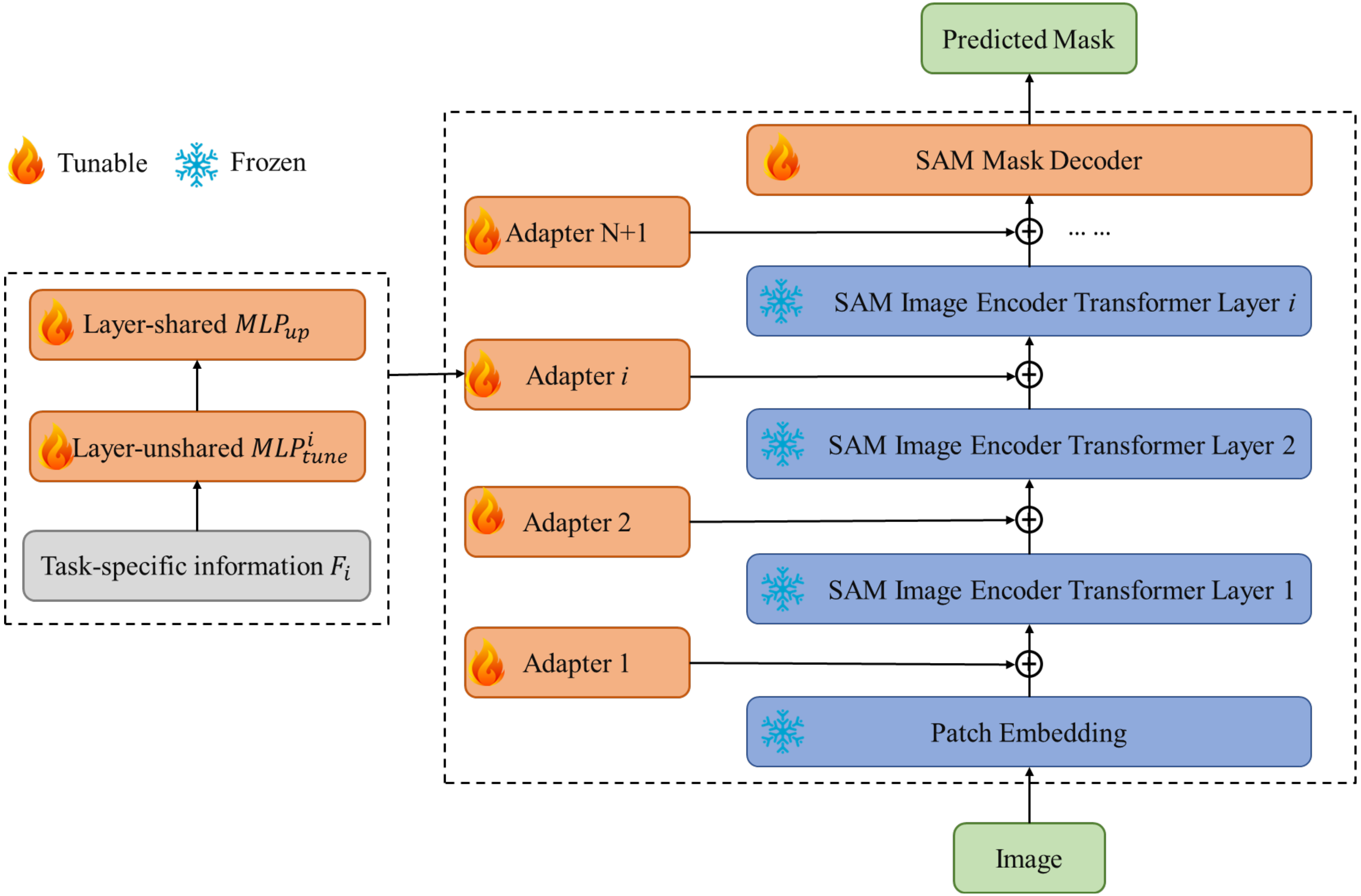
- One-shot training using PerSAM
- **The building class accuracy improves significantly when compared to SAM which is only adjusted using bounding box prompts**
- **SAM's performance changes based on the input imagery's spatial resolution, with more errors at lower resolutions**

RSPrompter

Chen *et al.* (2023)

- Automated instance segmentation method tailored for remote sensing images
- Propose a prompt generator designed to learn to create suitable prompts for SAM input
- Diminish semantic disparities and avoid the overfitting of the head
- **Merely fine-tuning the SAM decoder with minimal data might not always work**

Technical Background



Adapter:

1) two multilayer perceptron (MLP)
2) an activate function within two MLP

SAM-Adapter
Image Encoder

Feature Pyramid Network (FPN) with

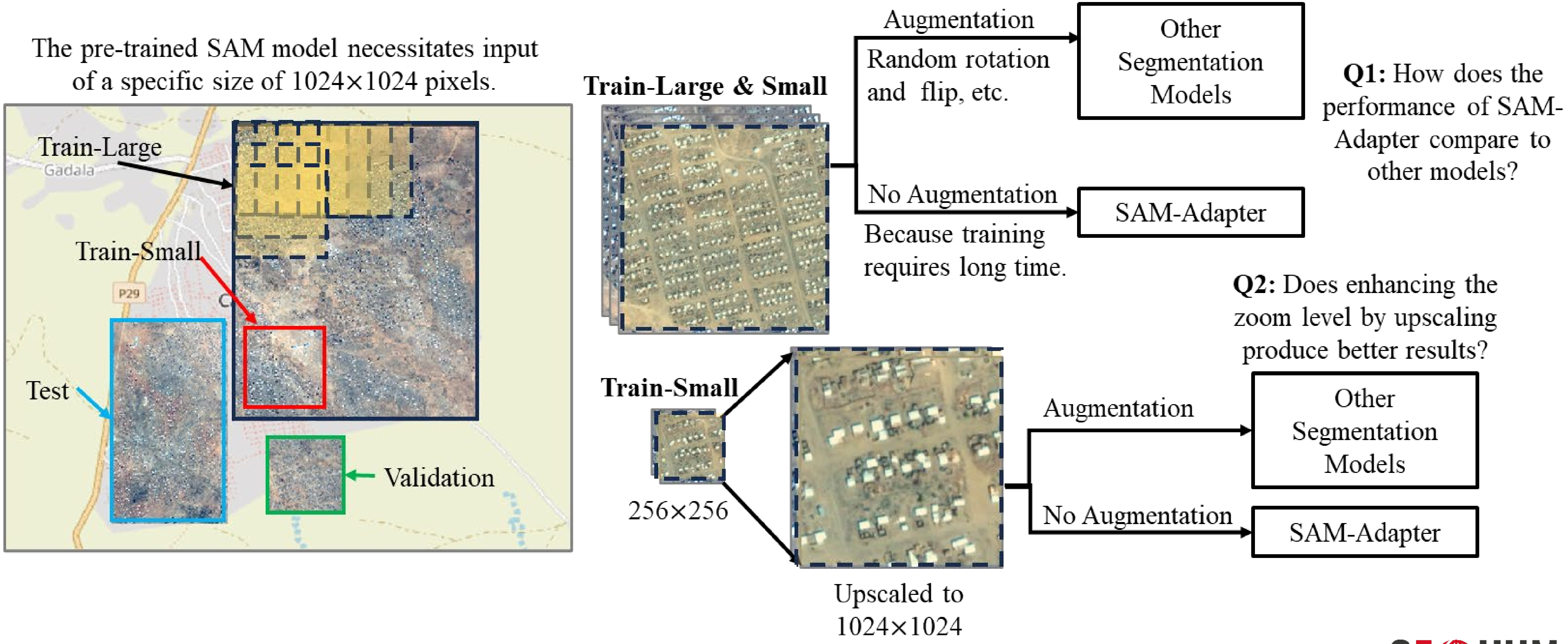
1) Mix Transformer-B0 (MiT)
2) MobileNet-v3-Large (MobileV3L)
3) ResNet34 [59]

Unet with

4) MobileNet- v3
5) ResNet34
6) ResNet101

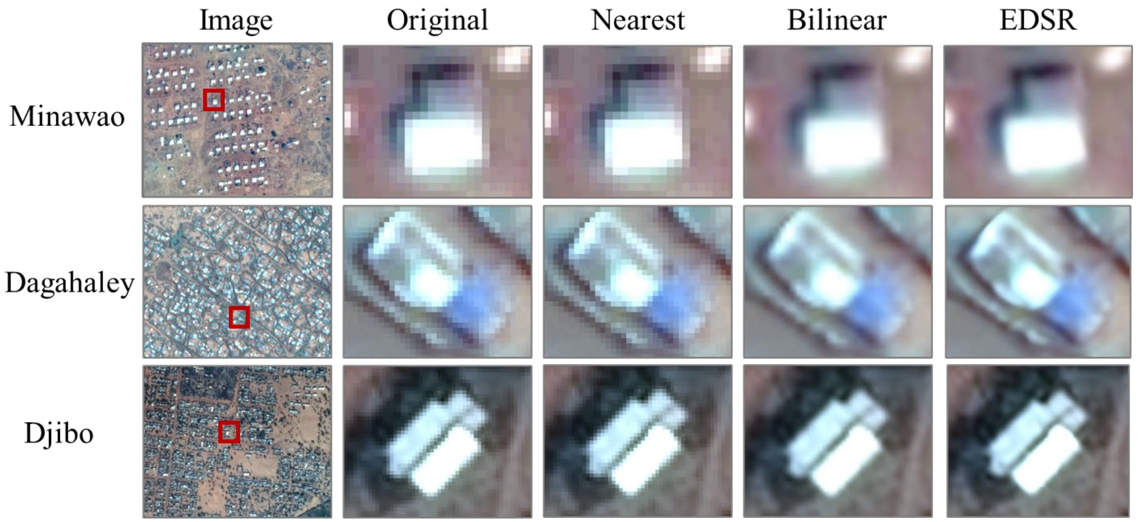
GEOHUM
Christian Doppler Laboratory

Methodology



Methodology

Refugee camp	Retrieved date	Sensor	Resolution (m)	Data type	Extent/pixel	Nr. of patches
Dagahaley	08/04/2017	WorldView-3	0.3	Train_Large	5754, 5074	350
				Train_Small	2010, 1944	56* / 350**
				Validation	1389, 1373	7
				Test	4783, 3101	
Djibo	12/12/2019	Pleiades-1A	0.5	Train_Large	4062, 5594	280
				Train_Small	2204, 2044	56 / 280
				Validation	1315, 1276	7
				Test	3851, 2924	
Minawao	12/02/2017	WorldView-2	0.5	Train_Large	3832, 4625	224
				Train_Small	1682, 1744	56 / 224
				Validation	1188, 1187	7
				Test	1817, 3165	



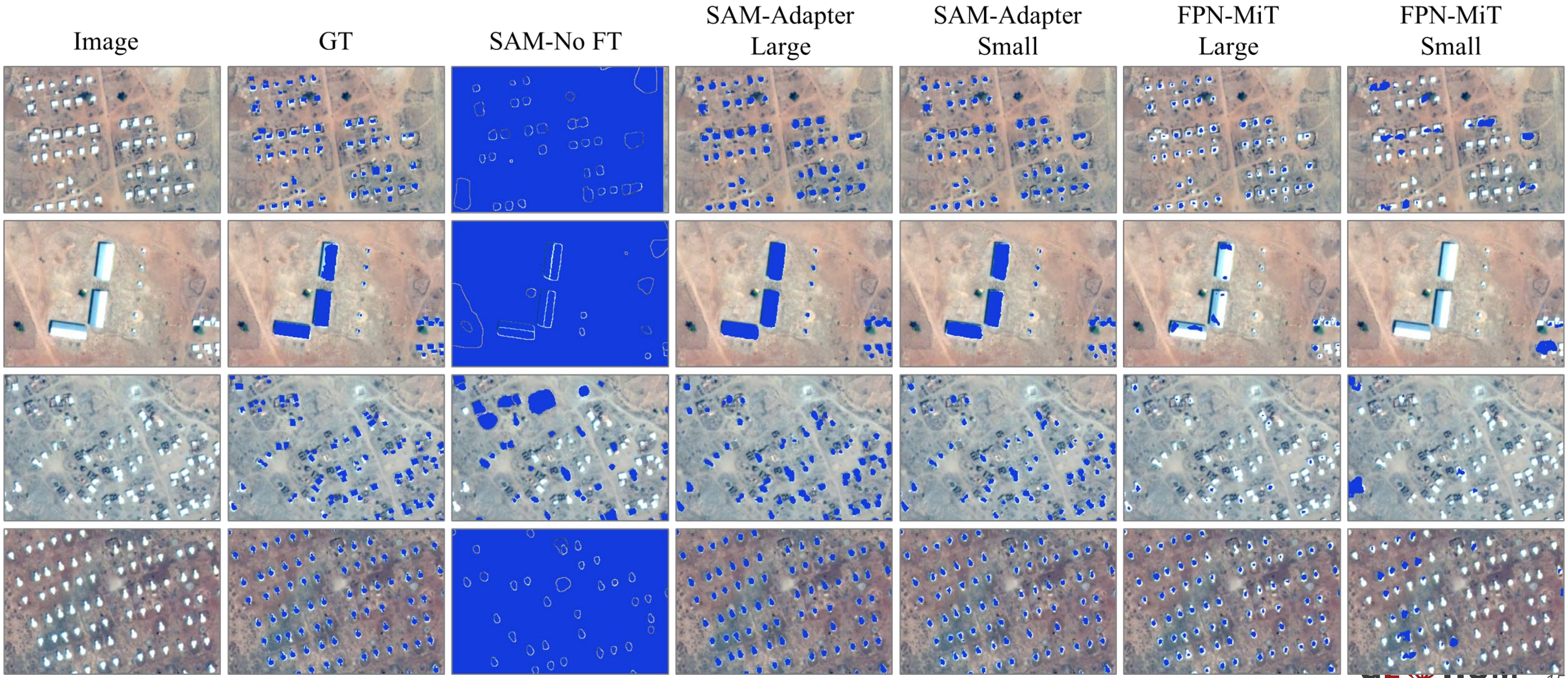
Results and Discussion

Model	Data size	Dagahaley				Djibo				Minawao			
		IoU	F1	Precision	Recall	IoU	F1	Precision	Recall	IoU	F1	Precision	Reca
FPN-MiT	Large	**0.523**	0.687	0.784	0.612	**0.546**	0.706	0.762	0.658	**0.515**	0.680	0.648	0.71
	Small	**0.297**	0.458	0.429	0.490	0.284	0.443	0.365	0.563	0.194	0.326	0.621	0.22
FPN-MobileV3L	Large	0.465	0.635	0.835	0.513	0.461	0.631	0.844	0.504	0.351	0.519	0.821	0.38
	Small	0.251	0.402	0.564	0.312	**0.304**	0.466	0.610	0.377	**0.195**	0.326	0.668	0.21
FPN-ResNet34	Large	0.351	0.519	0.803	0.384	0.293	0.453	0.857	0.308	0.158	0.273	0.769	0.16
	Small	0.138	0.239	0.721	0.143	0.107	0.267	0.170	0.614	0.139	0.180	0.114	0.42
Unet-ResNet101	Large	0.505	0.671	0.670	0.672	0.455	0.626	0.749	0.537	0.261	0.414	0.413	0.41
	Small	0.140	0.245	0.146	0.758	0.118	0.301	0.215	0.502	0.121	0.245	0.153	0.62
Unet-MobileV3L	Large	0.557	0.715	0.657	0.785	0.453	0.623	0.705	0.559	0.278	0.435	0.857	0.29
	Small	0.159	0.274	0.265	0.284	0.128	0.314	0.203	0.702	0.129	0.228	0.424	0.15
Unet-ResNet34	Large	0.432	0.604	0.793	0.488	0.382	0.553	0.855	0.408	0.300	0.461	0.631	0.36
	Small	0.129	0.229	0.141	0.612	0.158	0.272	0.270	0.275	0.144	0.252	0.164	0.54
SAM-Adapter	Large	**0.619**	0.765	0.793	0.738	**0.657**	0.793	0.796	0.790	**0.583**	0.736	0.779	0.69
	Small	**0.560**	0.718	0.626	0.842	**0.588**	0.741	0.769	0.714	**0.571**	0.727	0.699	0.75
SAM	No FT	0.150	0.261	0.156	0.795	0.093	0.170	0.094	0.903	0.038	0.074	0.039	0.73
	SR-No FT	0.219	0.360	0.231	0.809	0.104	0.189	0.107	0.833	0.067	0.125	0.068	0.81(

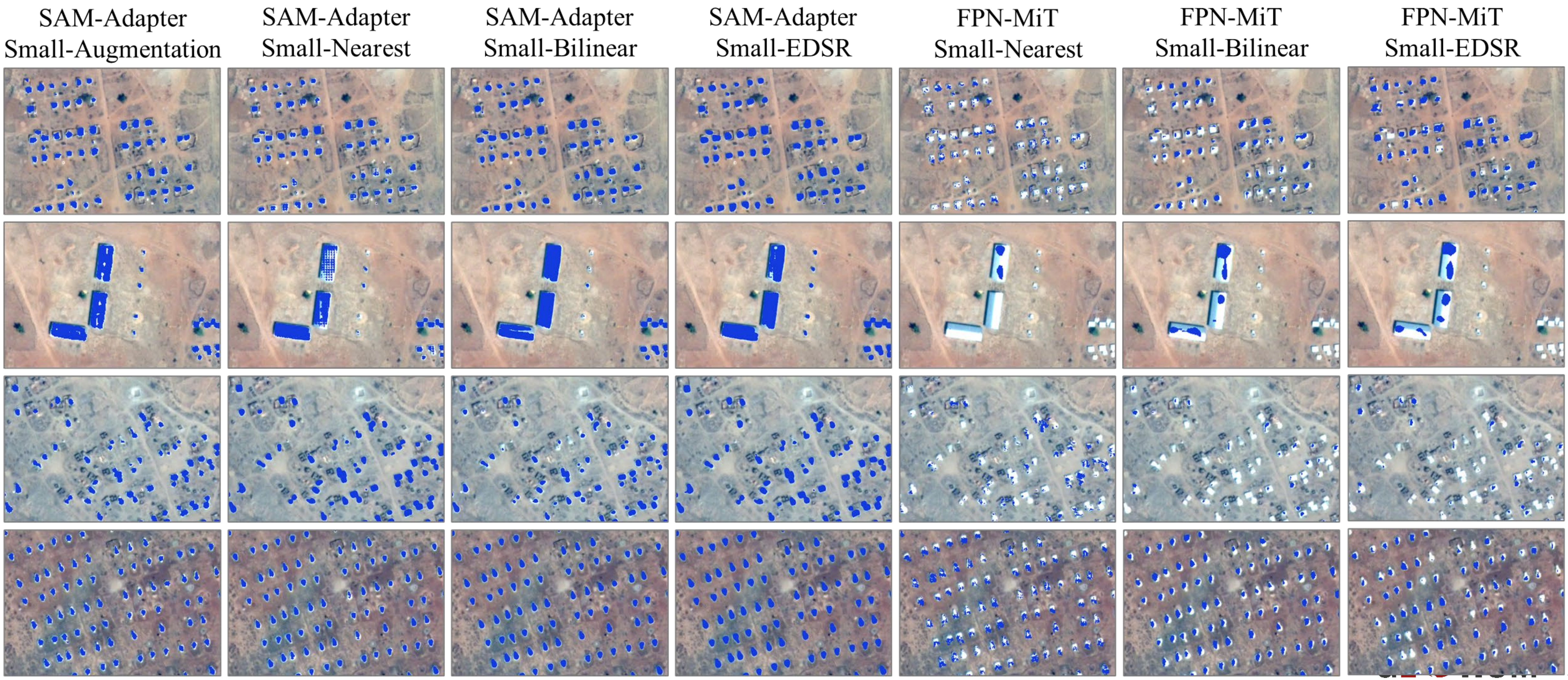
- SAM-Adapter outperforms other six selected semantic segmentation models.
- When using smaller training data, the improvement is more significant.
- Among the six selected other segmentation models, FPN-MiT model performs best.

GEOHUM Christian Doppler Laboratory for geospatial and EO-based humanitarian technologies

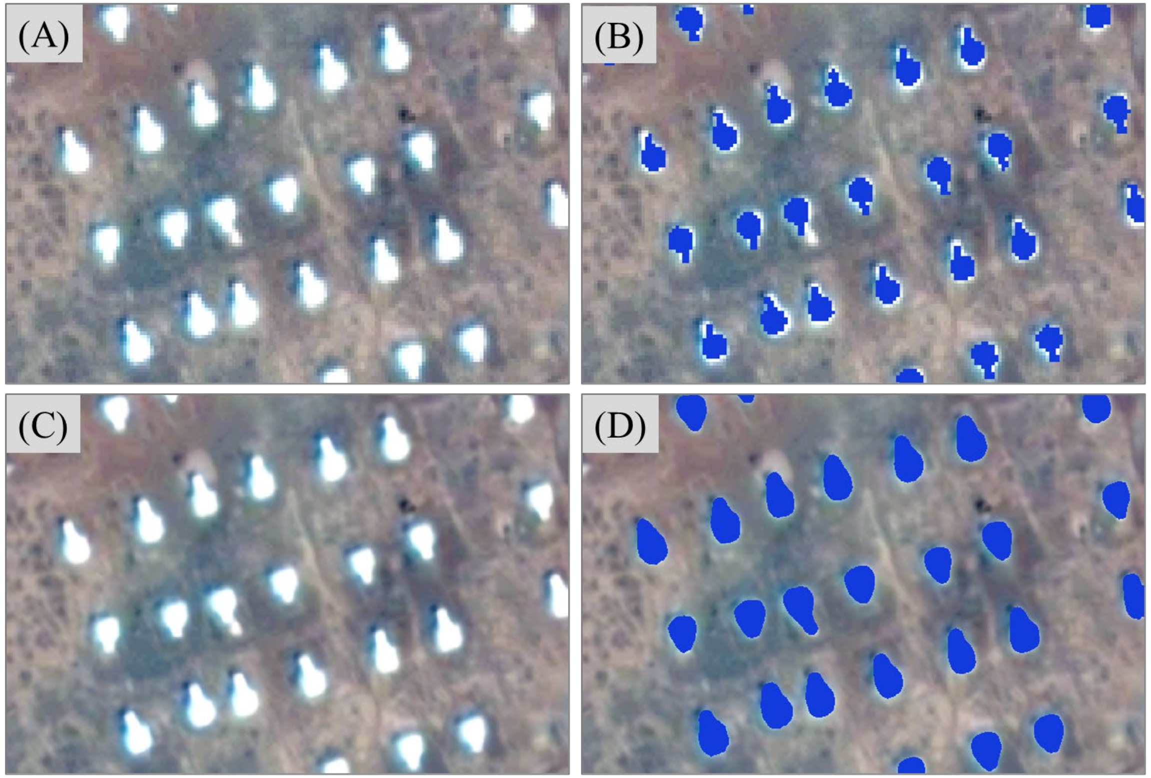
Results and Discussion



Results and Discussion



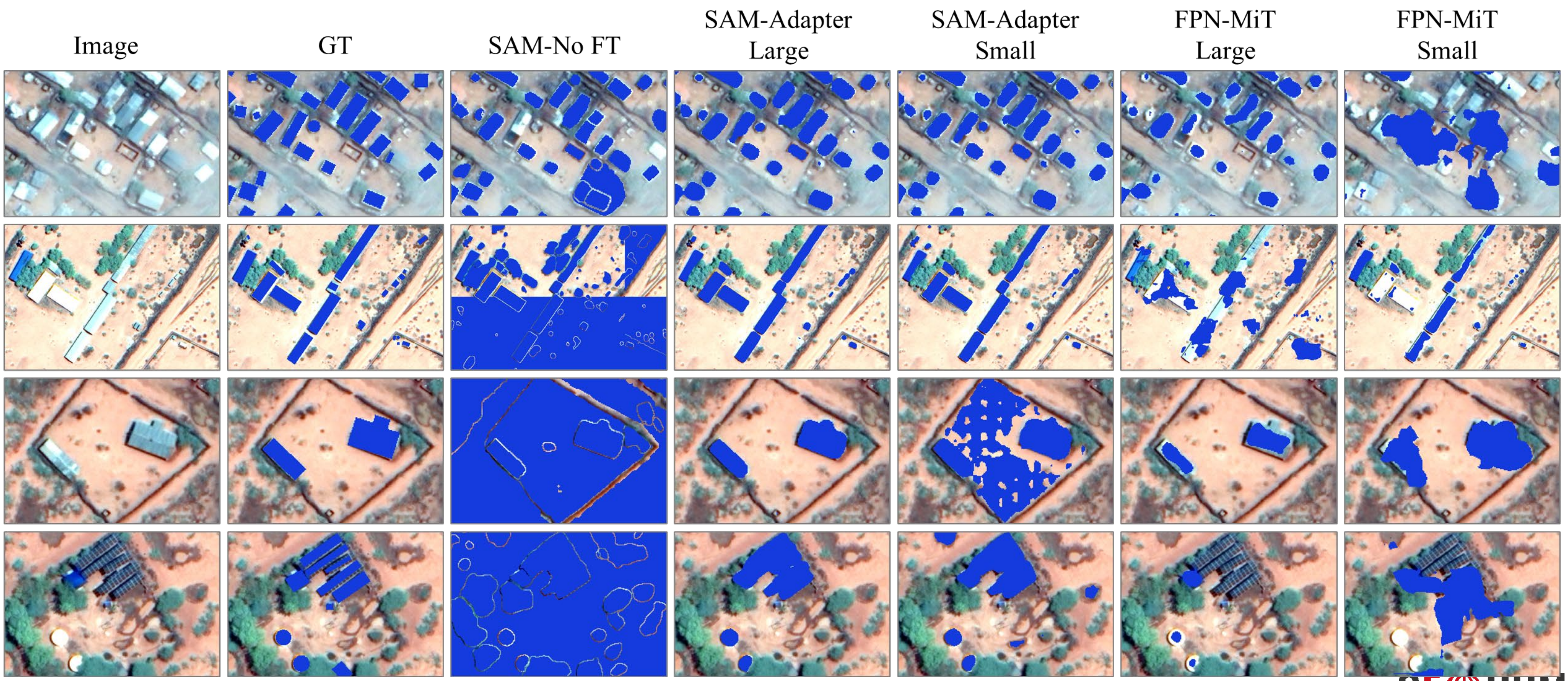
Results and Discussion



An example in the Minawao refugee camp showcases the influence of upscaling by SR models on the performance of SAM-Adapter. (A) Original image; (B) Ground truth; (C) Upscaled image; (D) Predicted masks from SAM-Adapter, which are smoother than ground truth labels.

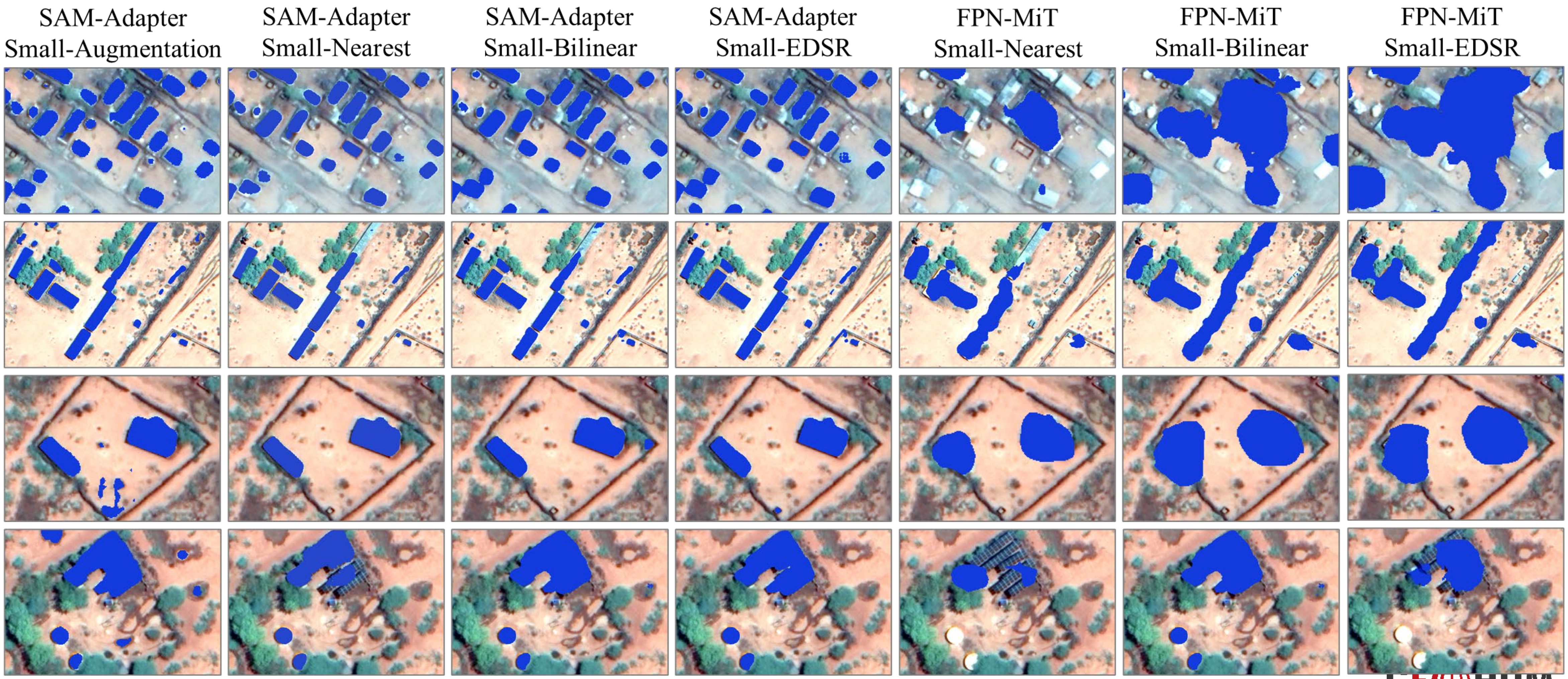
GEOHUM
Christian Doppler Laboratory
for geospatial and EO-based humanitarian technologies

Results and Discussion

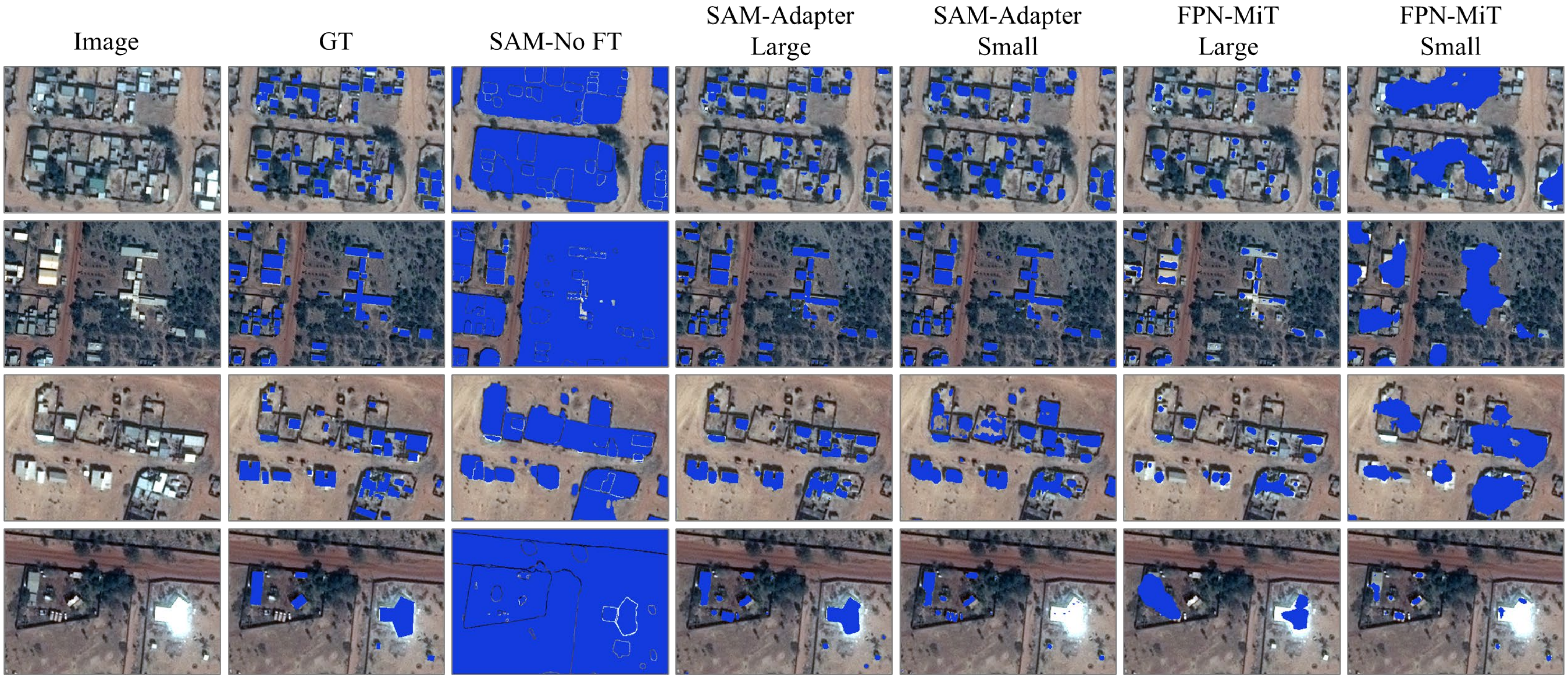


Christian Doppler Laboratory

Results and Discussion

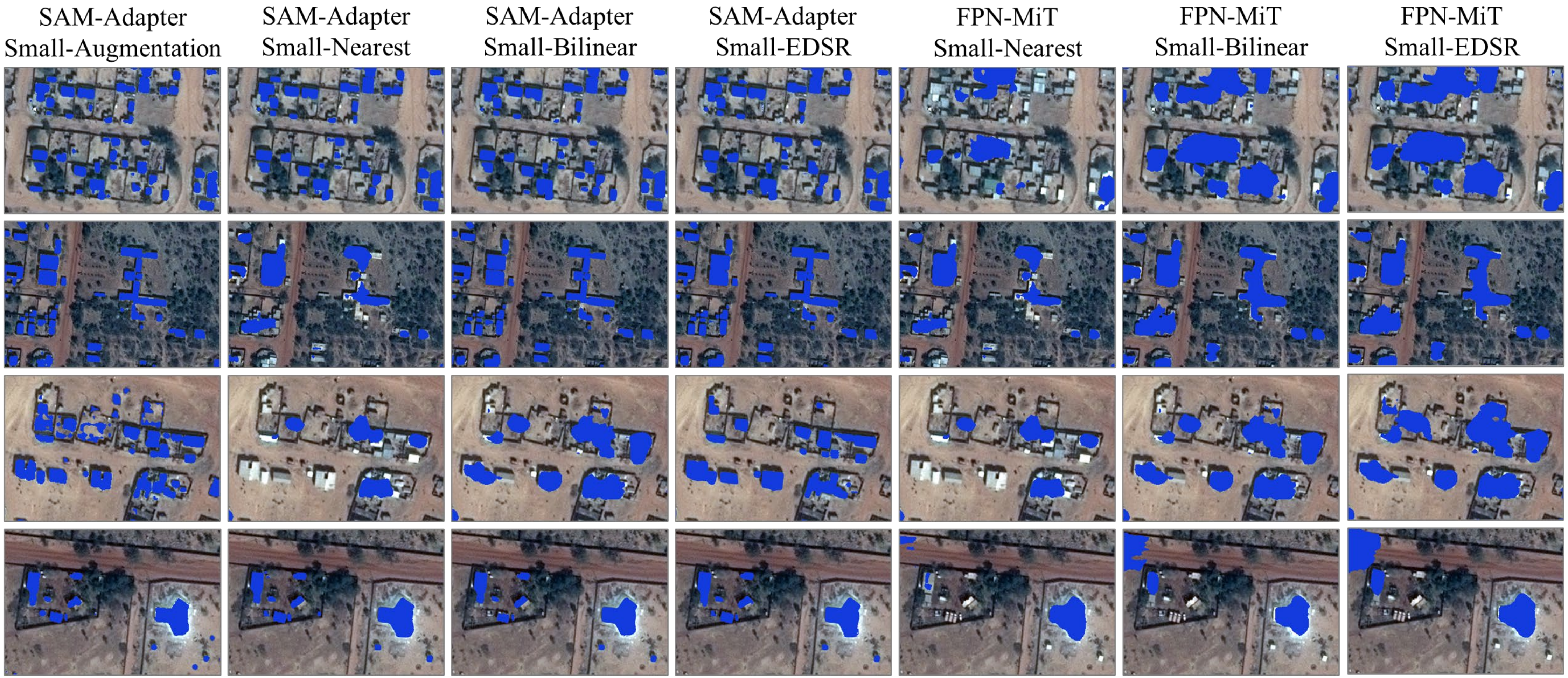


Results and Discussion



GEOHUM
Christian Doppler Laboratory
for geospatial and EO-based humanitarian technologies

Results and Discussion



GEOHUM Christian Doppler Laboratory for geospatial and EO-based humanitarian technologies

Conclusions

- In a broader context, this work pioneers the implementation of SAM-Adapter for refugee dwelling extraction tasks and offers a comprehensive workflow.
- SAM-Adapter stands out as a highly effective tool for semantic segmentation tasks, especially when confronted with limited pixelwise label data.
- Its adaptability and remarkable performance across diverse refugee camps highlight its potential in refugee-dwelling extraction for humanitarian operations.
- The significance of upscaling methods (especially Super Resolution) on model performance is profoundly essential.