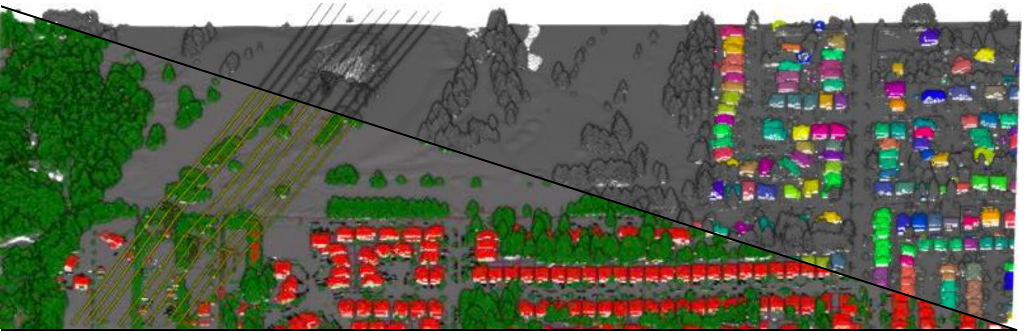
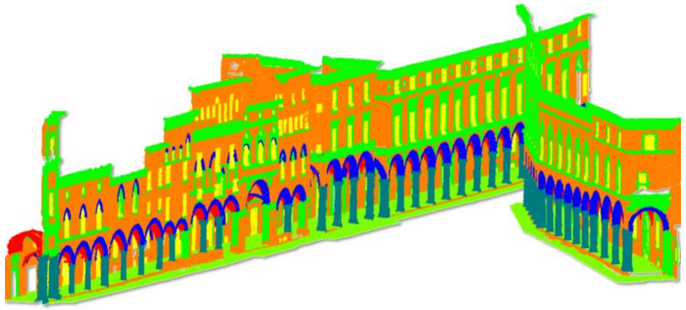
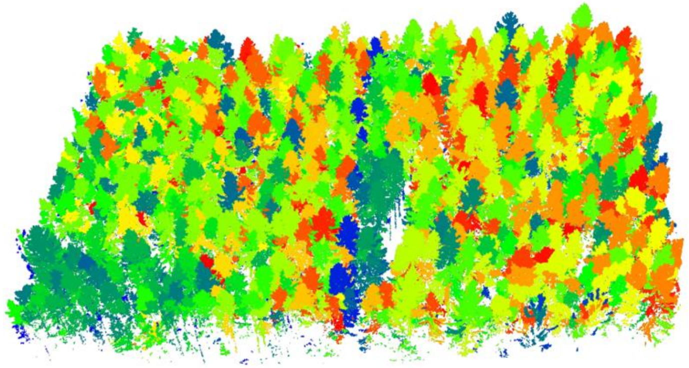
Giving a meaning to 3D point clouds

Fabio REMONDINO

3D Optical Metrology (3DOM) unit - Bruno Kessler Foundation (FBK)
Trento, Italy
Email: remondino@fbk.eu - http://3dom.fbk.eu

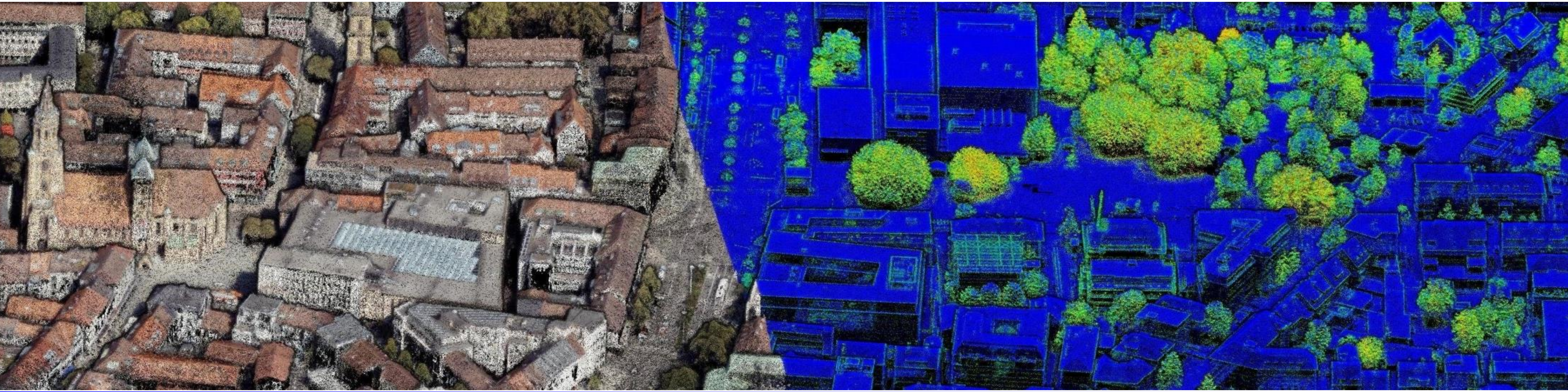






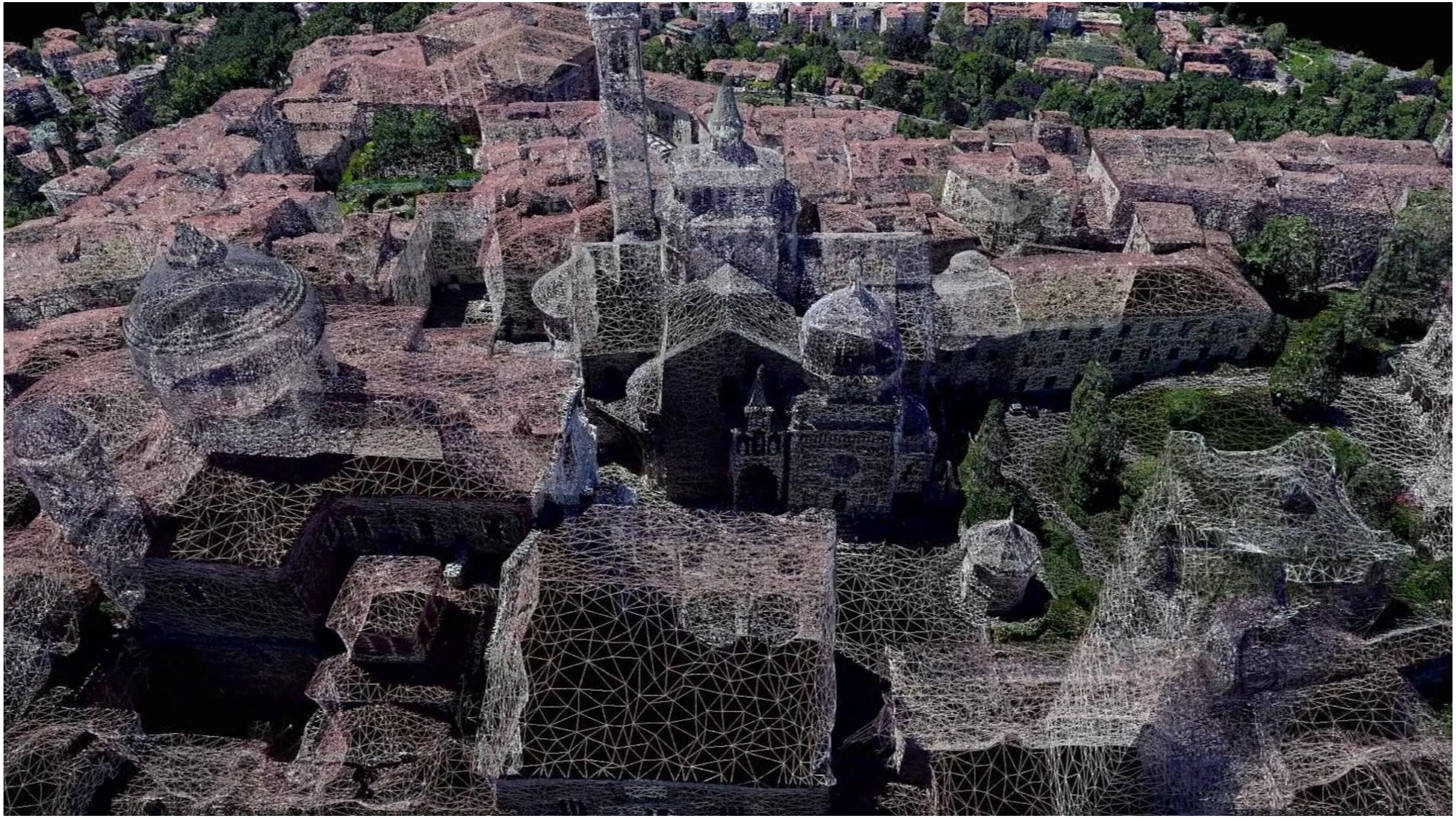
Point Cloud Generation

	PASSIVE (IMAGING)	HYBRID	ACTIVE (RANGING)
AERIAL & SATELLITE	• Linear sensors • Large frame cameras • Multi-view cameras (oblique)	• Single frame + LiDAR • Oblique + LiDAR	• Traditional linear Airborne Laser Scanning • SPL/Geiger-mode Airborne Laser Scanning
TERRESTRIAL	• DSLR cameras • Panoramic / spherical cameras • Smartphones	• Mobile Mapping systems • Hand-held / backpack system • RGB-D sensors	• TOF laser scanner (long-range) • Triangulation laser scanners (short-range) • Structured light systems (short-range)



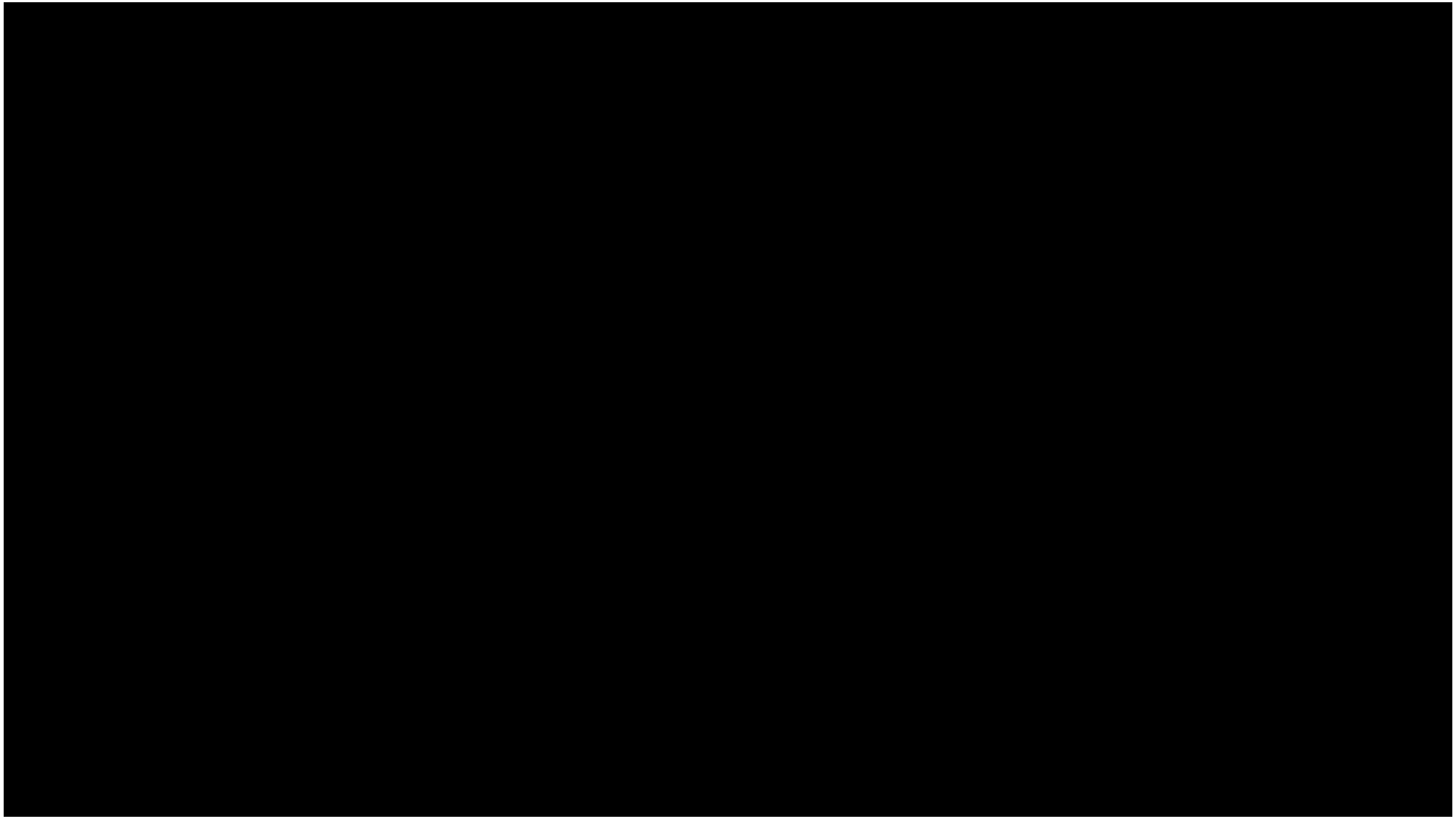
Point Cloud Generation

Vexcel Ultracam Osprey,
flown by AVT Airborne Sensing:
GSD (N) 10 cm, 80/60%, flight height 1800 m



[Toschi, I., Ramos, M.M., Nocerino, E., Menna, F., Remondino, F., Moe, K., Poli, D., Legat, K., Fassi, F., 2017: Oblique photogrammetry supporting 3D urban reconstruction of complex scenarios. ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., Vol. XLII-1-W1, pp. 519-526]

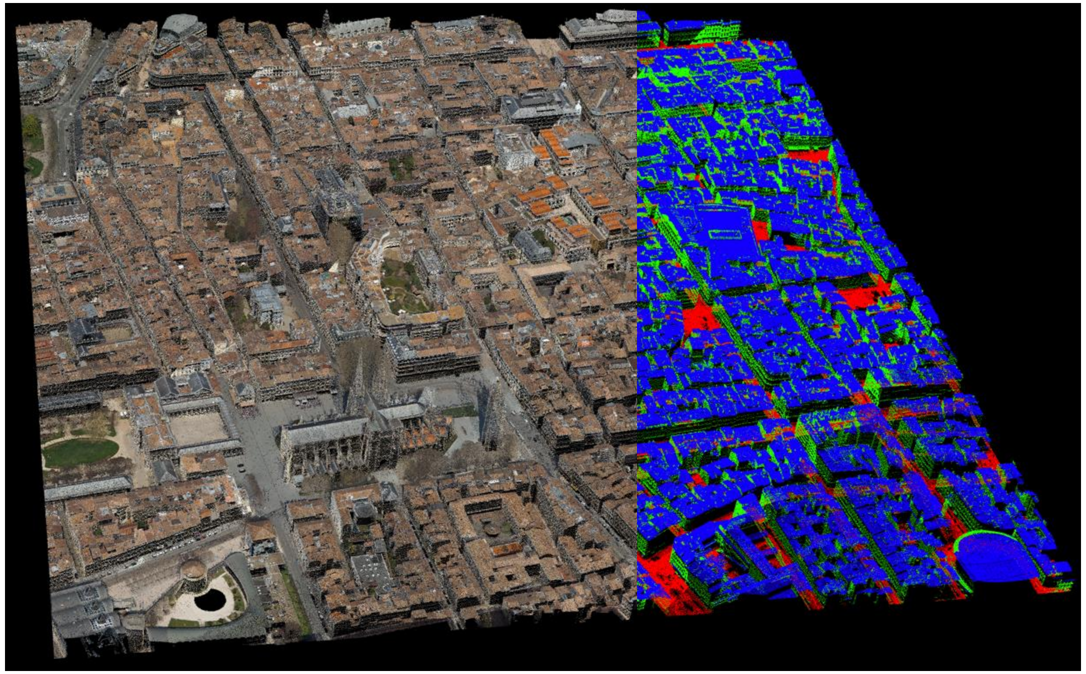
Semantic segmentation / 3D Classification

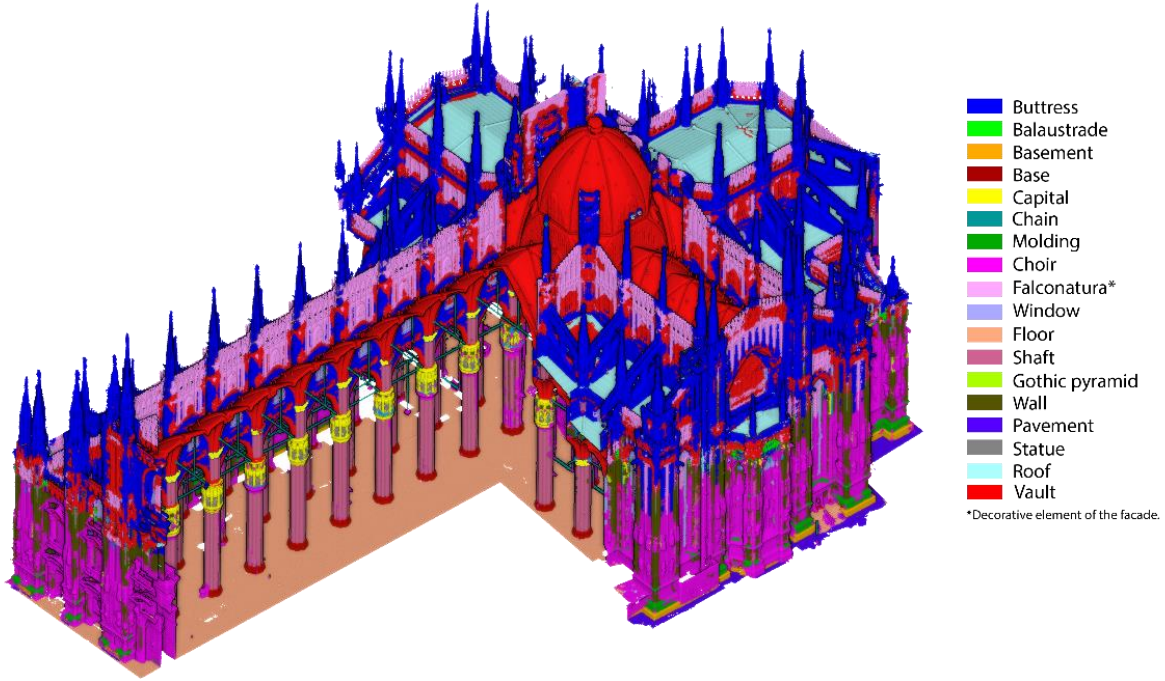


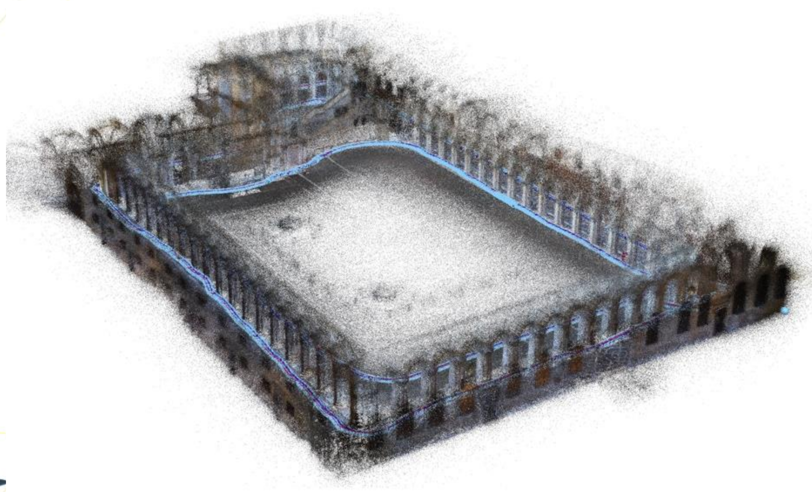
https://www.youtube.com/watch?v=h4m-Hu-7rfs

[Özdemir, E., Remondino, F., Golkar, A., 2021. An Efficient and General Framework for Aerial Point Cloud Classification in Urban Scenarios. Remote Sensing, Vol.13, 1985]

Point Clouds

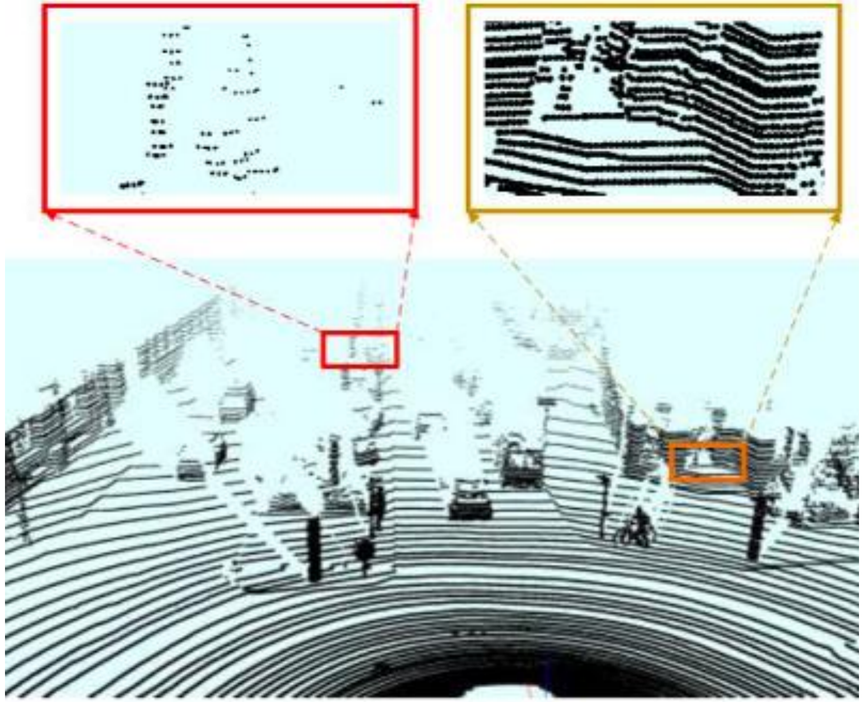


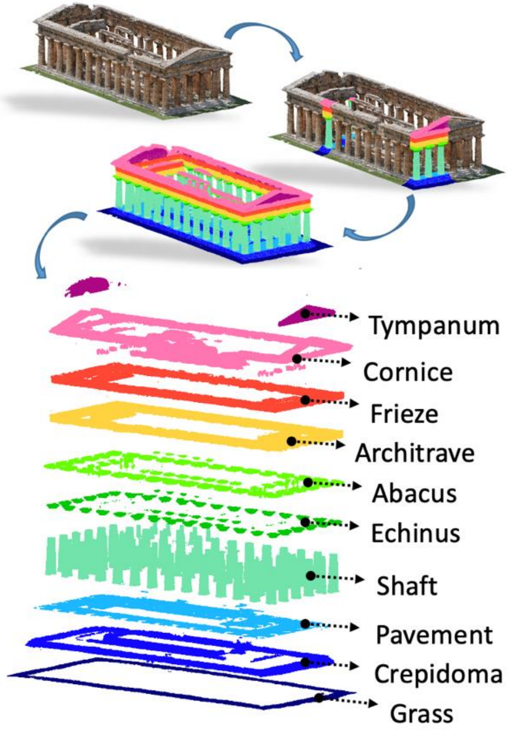




Challenge

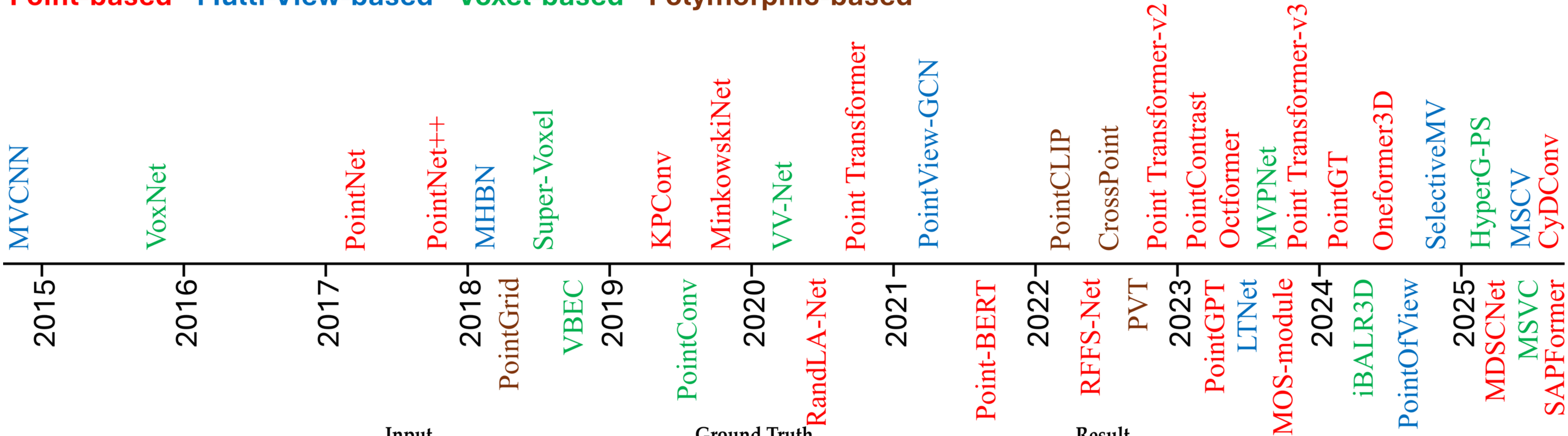
- Every **scenario** is different (shape, size, objects, etc.)
- Every **dataset** is different (resolution / density, sensor features/attributes, holes, etc.)
- **Needs** are different (LULC, change detection, forestry management, conservation / restoration, etc.)
- Point cloud are **difficult by definition** (orderless, irregularity, noise, non-uniform, uneven density, completeness, etc.)

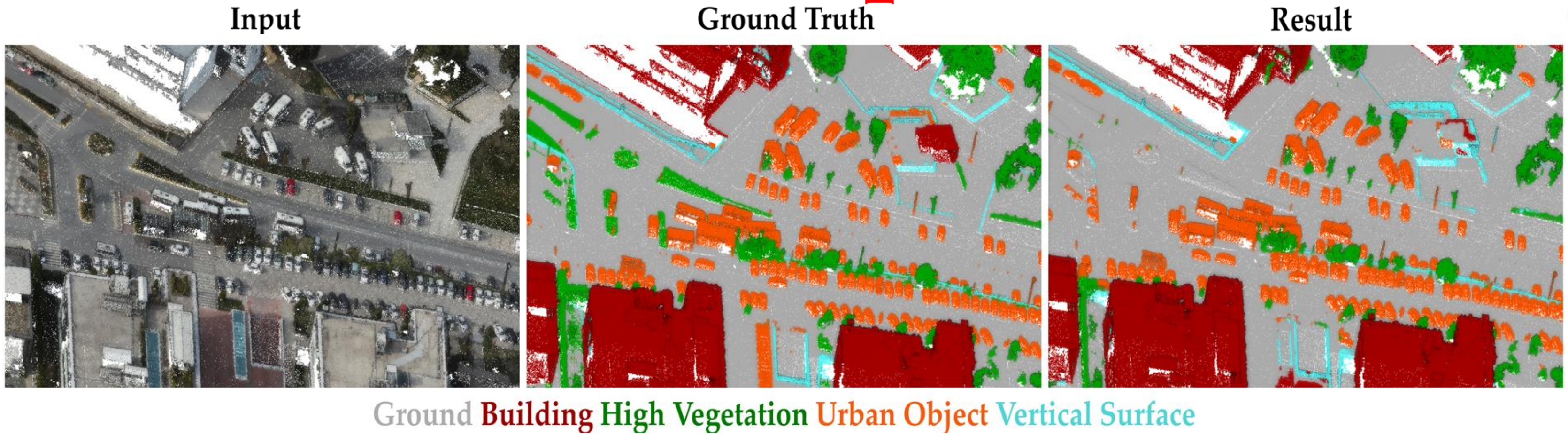




Point Cloud Classification - Methods

Point-based **Multi-View-based** **Voxel-based** **Polymorphic-based**

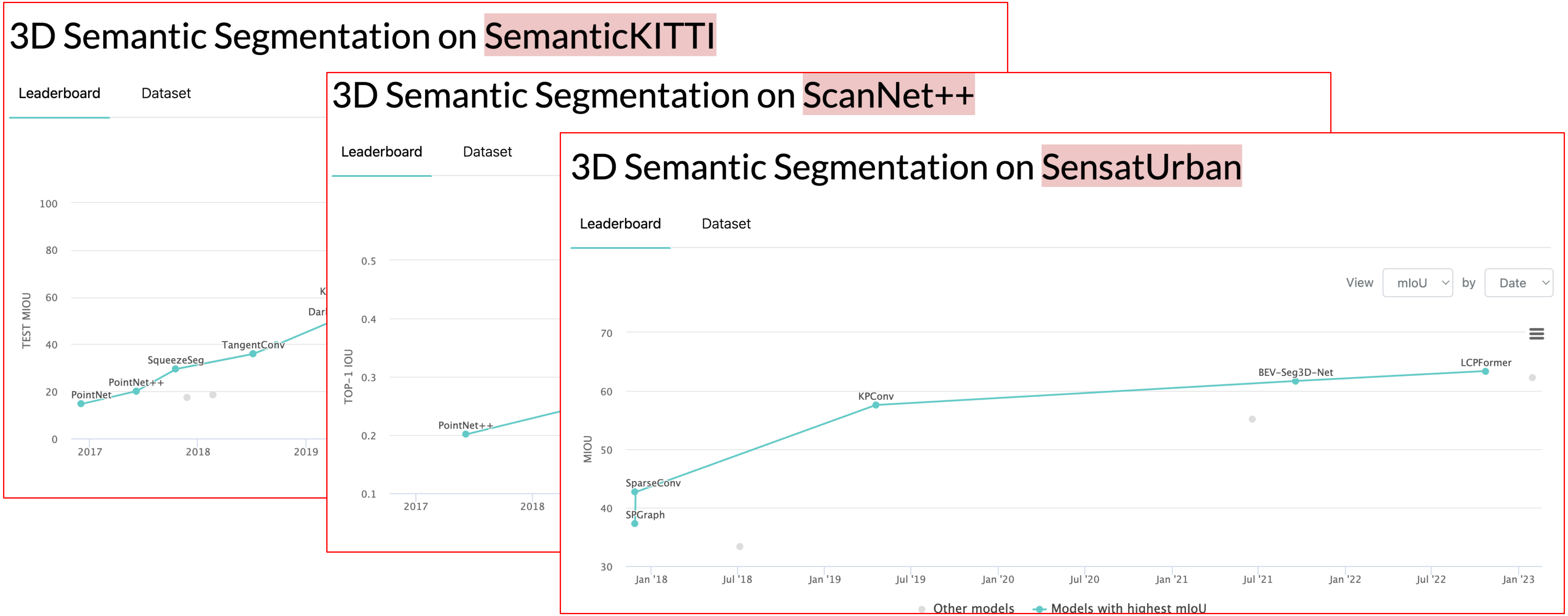




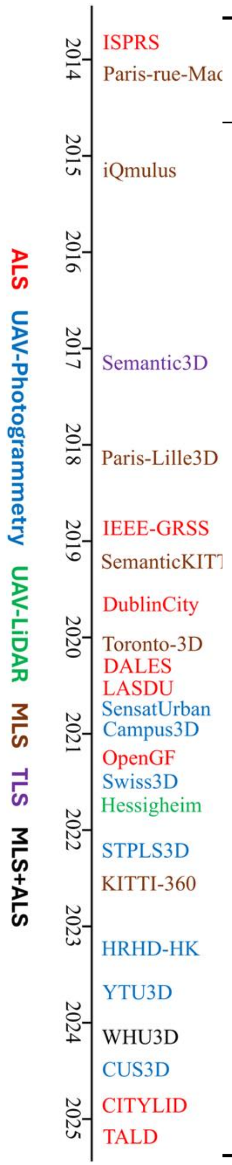
[Bayrak, O., Farella, E.M., Ma, Z., Remondino, F., Uzar, M., 2025: **Combining 3D Urban Objects From All Around the World to Improve Object Classification and Semantic Segmentation**. Remote Sensing, in press]

Point Cloud Classification - Methods

- one-size-fits-all? nooooooooo ☹
- no unique winner, methods are fit for specific datasets, lack of generalization / replicability
- https://paperswithcode.com/task/3d-semantic-segmentation

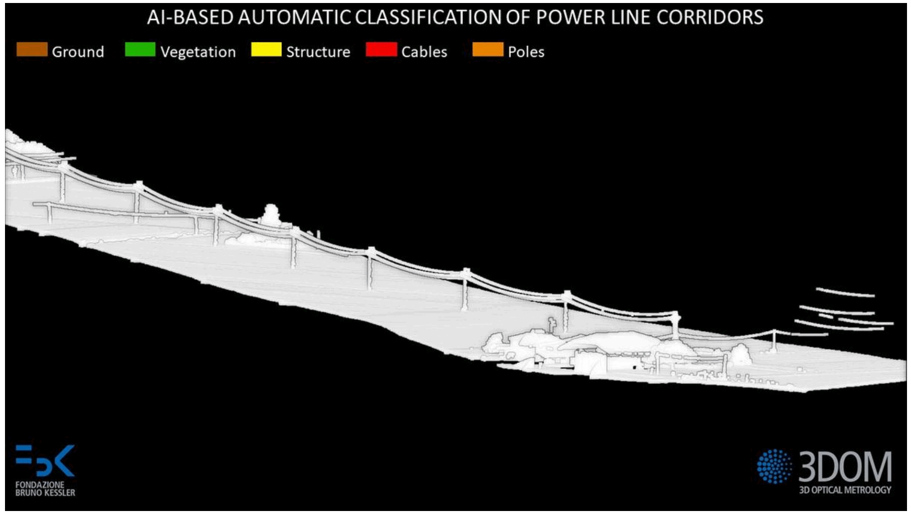


Point Cloud Classification - Datasets



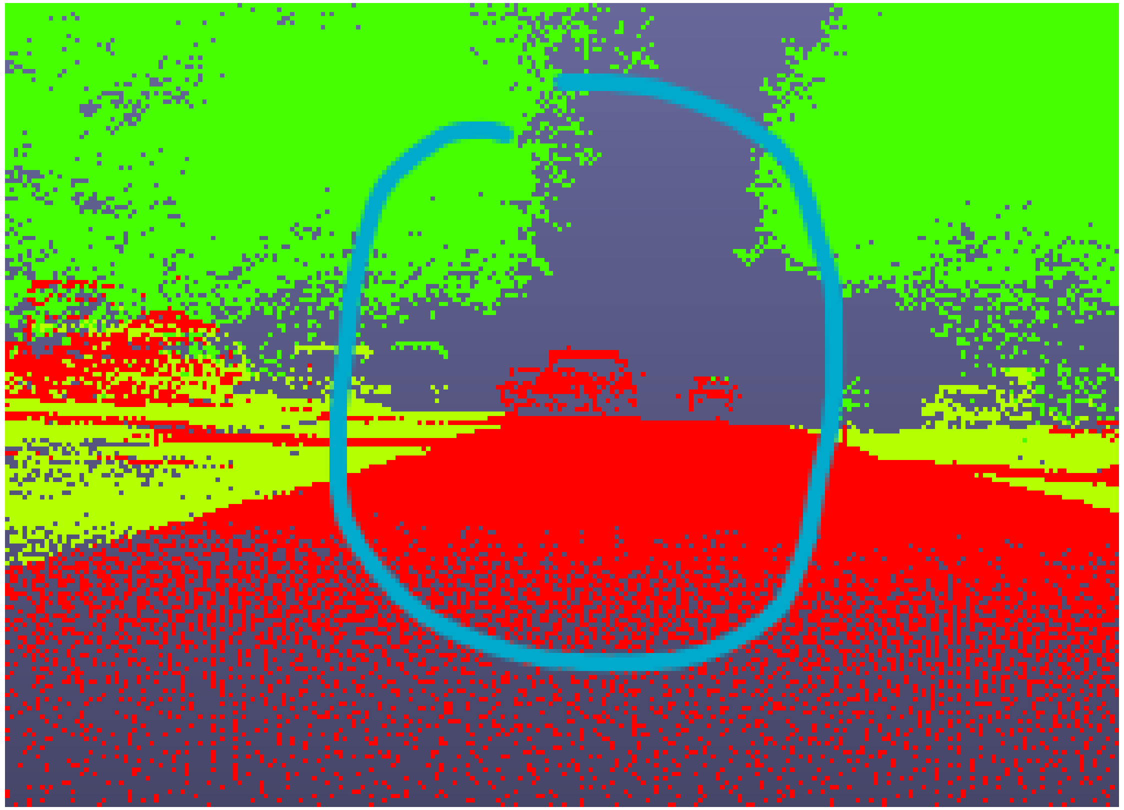
Name	Classes	Points (mil)	Spatial Size (m^2)	RGB	Sensor
ISPRS [33]	9	1.2	1.6 x 10^6	No	ALS
Paris-rue-Madame [34]	17	20	0.16 * 10^3	No	MLS
iQmulus [35]	8 (22)	300	10 x 10^3	No	MLS
Semantic3D [36]	8	4009	-	No	TLS
Paris-Lille3D [19]	9 (50)	143	1.94 x 10^3	No	MLS
IEEE-GRSS [37]	5	102	34 x 10^6	No	ALS
SemanticKITTI [38]	22 (28)	4549	39.2 x 10^3	No	MLS
DublinCity [7]	13	260	2 x 10^6	No	ALS
Toronto3D [39]	8	78.3	1 x 10^3	Yes	MLS
DALES [40]	8	505.3	10 x 10^6	No	ALS
LASDU [41]	5	3.12	1.02 x 10^6	No	ALS
SensatUrban [15]	13	2847.1	7.64 x 10^6	Yes	UAV-Photo
Swiss3DCities [42]	5	226	2.7 x 10^6	Yes	UAV-Photo
Campus3D [43]	24	937.1	1.58 x 10^6	Yes	UAV-Photo
OpenGF [44]	2	500	47 x 10^6	No	ALS
Hessigheim [20]	11	125.7	8 x 10^4	Yes	UAV-LiDAR
STPLS3D [21]	6	-	6 x 10^6	Yes	UAV-Photo
KITTI-360 [45]	37	1000	*73.7km	No	MLS
HRHD-HK [46]	7	273	9 x 10^6	Yes	UAV-Photo
YTU3D [47]	45	1700	2 x 10^6	Yes	UAV-Photo
WHU3D [24]	37	393	6.5 x 10^3	No	MLS + ALS
CUS3D [48]	10	152.3	2.85 x 10^6	Yes	UAV-Photo
CITYLID [49]	9	15000	1060 x 10^6	No	ALS
TALD [50]	4	121	9 x 10^6	No	ALS

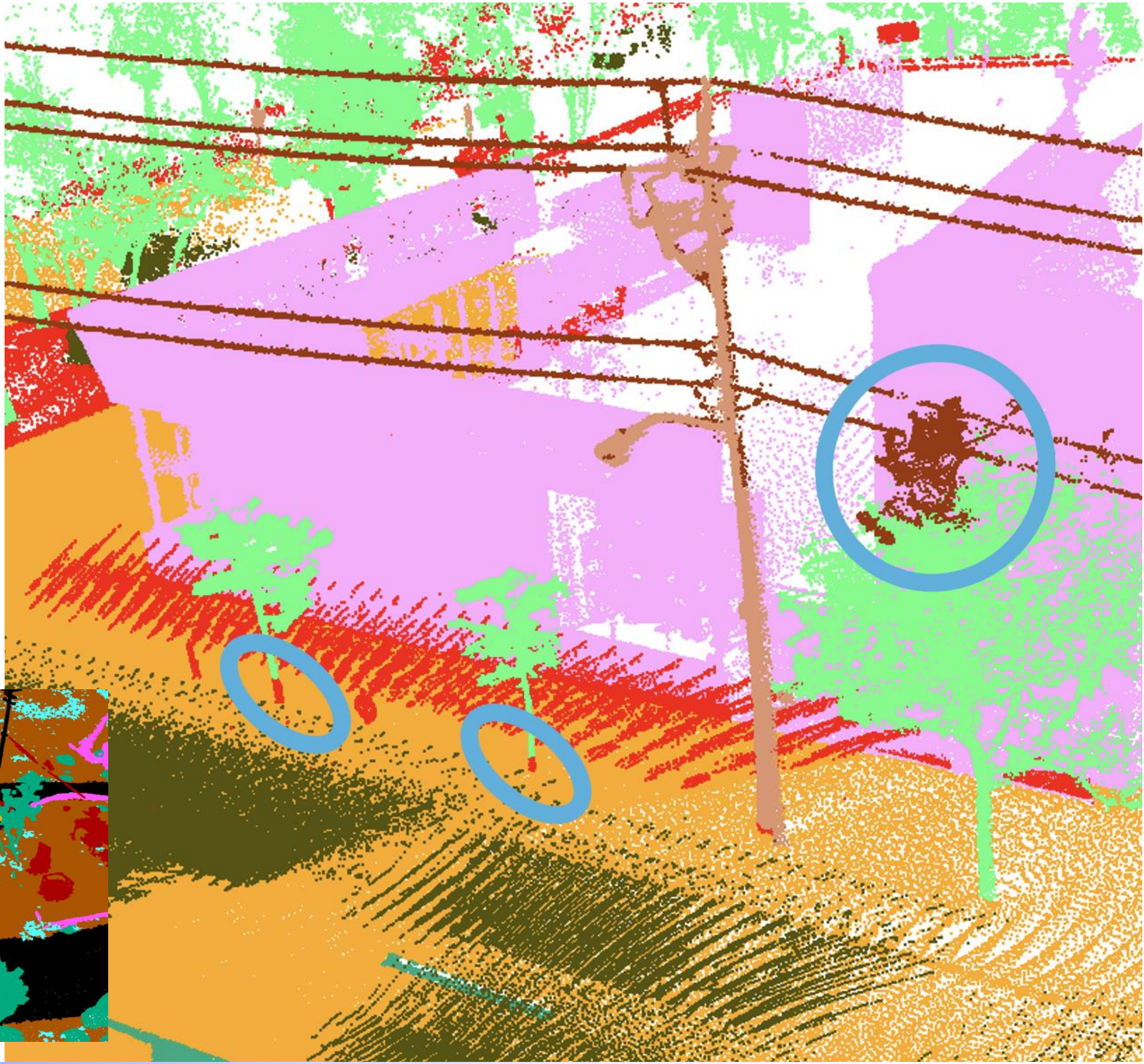
- ❑ no standards in class definition
- ❑ general lack class diversity
- ❑ imbalance (under-represented) classes
- ❑ multiple sensors (cameras, LiDARs)
- ❑ different attributes (RGB, I, etc.)
- ❑ different densities / resolutions
- ❑ annotation errors



[Bayrak, O. Ma, Z., Farella, E.M., Remondino, F., Uzar, M., 2024: **ESTATE: A Large Dataset of Under-Represented Urban Objects for 3D Point Cloud Classification**. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2-2024, 25–32]

Annotations





3D Classification - Some works

- A large dataset of under-represented urban objects for 3D point cloud classification (ESTATE)

(Onur Bayrak, Zhenyu Ma, Elisa Farella)

- Knowledge Extended Neural Network (KENN)

(Eleonora Grilli, Marteen Bassier)

- Semantic segmentation in 3D forestry

(Narges Takhtkeshha, Gottfried Mandlburger, Juha Hyyppä)

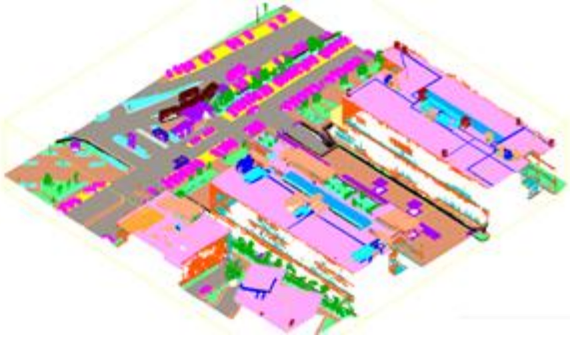
- Visual Language Models to support 3D classification

(Ashkan Alami)

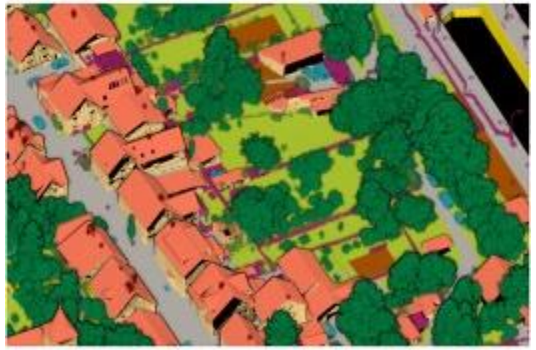
Point Cloud Classification - ESTATE

- Point-based vs object-based dataset

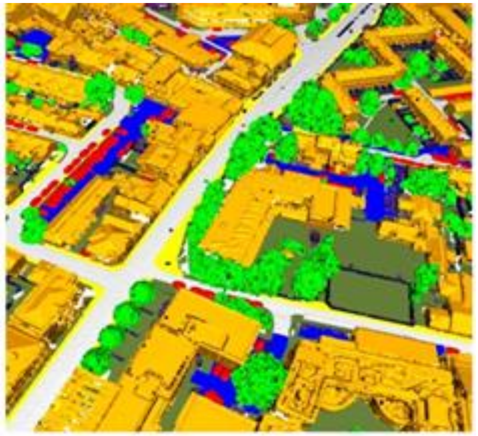
Point-based class ID assignment



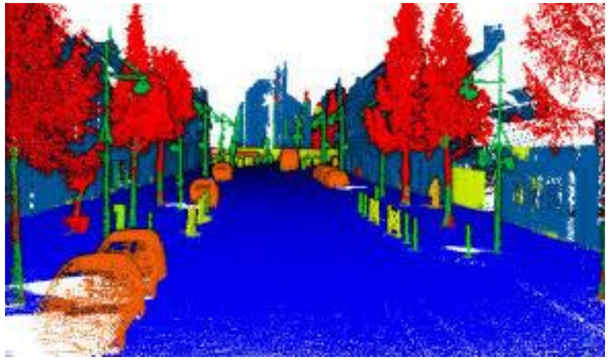
YTU3D



Hessigheim3D



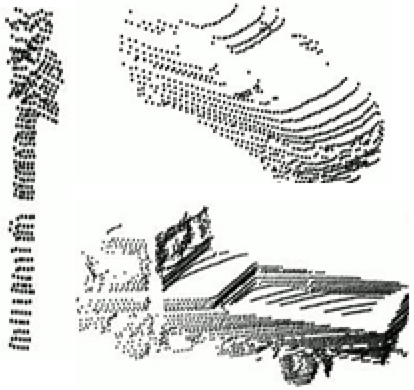
SensatUrban



Paris-Lille 3D

Object-based class ID assignment

Objaverse



Sydney Urban Objects

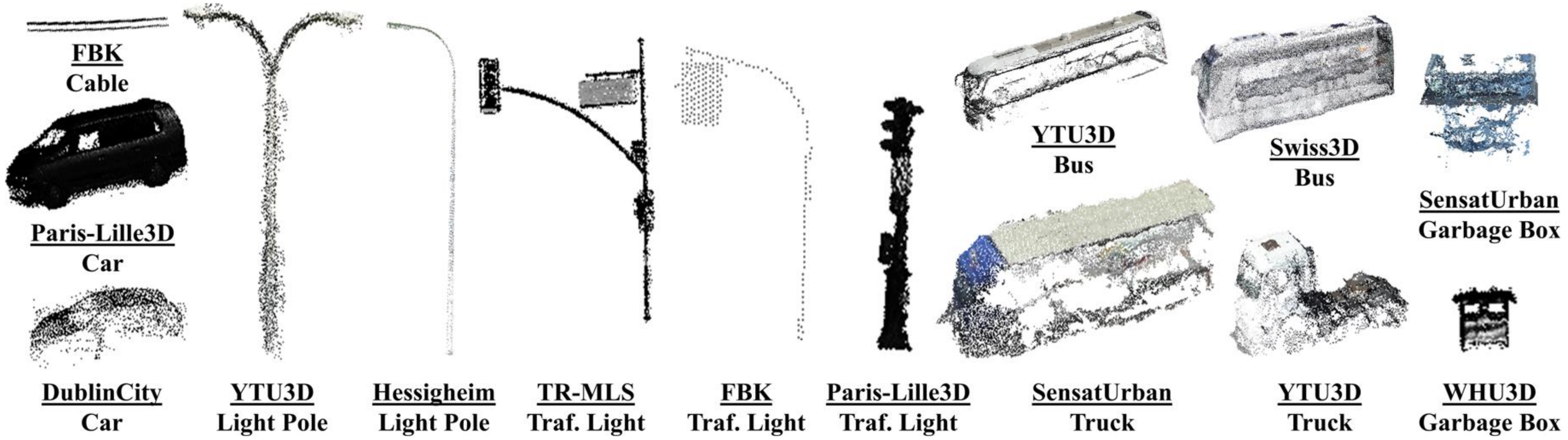
ShapeNet

ModelNet40

[Bayrak, O. Ma, Z., Farella, E.M., Remondino, F., Uzar, M., 2024: ESTATE: A Large Dataset of Under-Represented Urban Objects for 3D Point Cloud Classification. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2-2024, 25–32]

Point Cloud Classification - ESTATE

- ESTATE dataset - a large dataset of under-represented urban objects for 3D point cloud classification



Name	Reference	Classes	Total Numb. of Objects	Target	Scene Type
Sydney Urban Objects	De Deuge et al. (2013)	14	588	Outdoor scenes	Real World
ModelNet10	Wu et al. (2015)	10	4596	Indoor objects	Synthetic
ModelNet40	Wu et al. (2015)	40	12311	Indoor objects	Synthetic
ShapeNet	Chang et al. (2015)	55	51190	Indoor objects	Synthetic
ScanNet	Dai et al. (2017)	17	12283	Indoor scenes	Real World
ScanObjectNN	Uy et al. (2019)	15	2902	Indoor objects	Real World
Objaverse	Deitke et al. (2022)	21 K +	10 million +	Indoor objects	Synthetic
ModelNet40-C	Sun et al. (2022)	15	185000	Indoor objects	Synthetic
ESTATE (Our)	Bayrak et al. (2024)	13	**6528**	Outdoor scenes	Real World

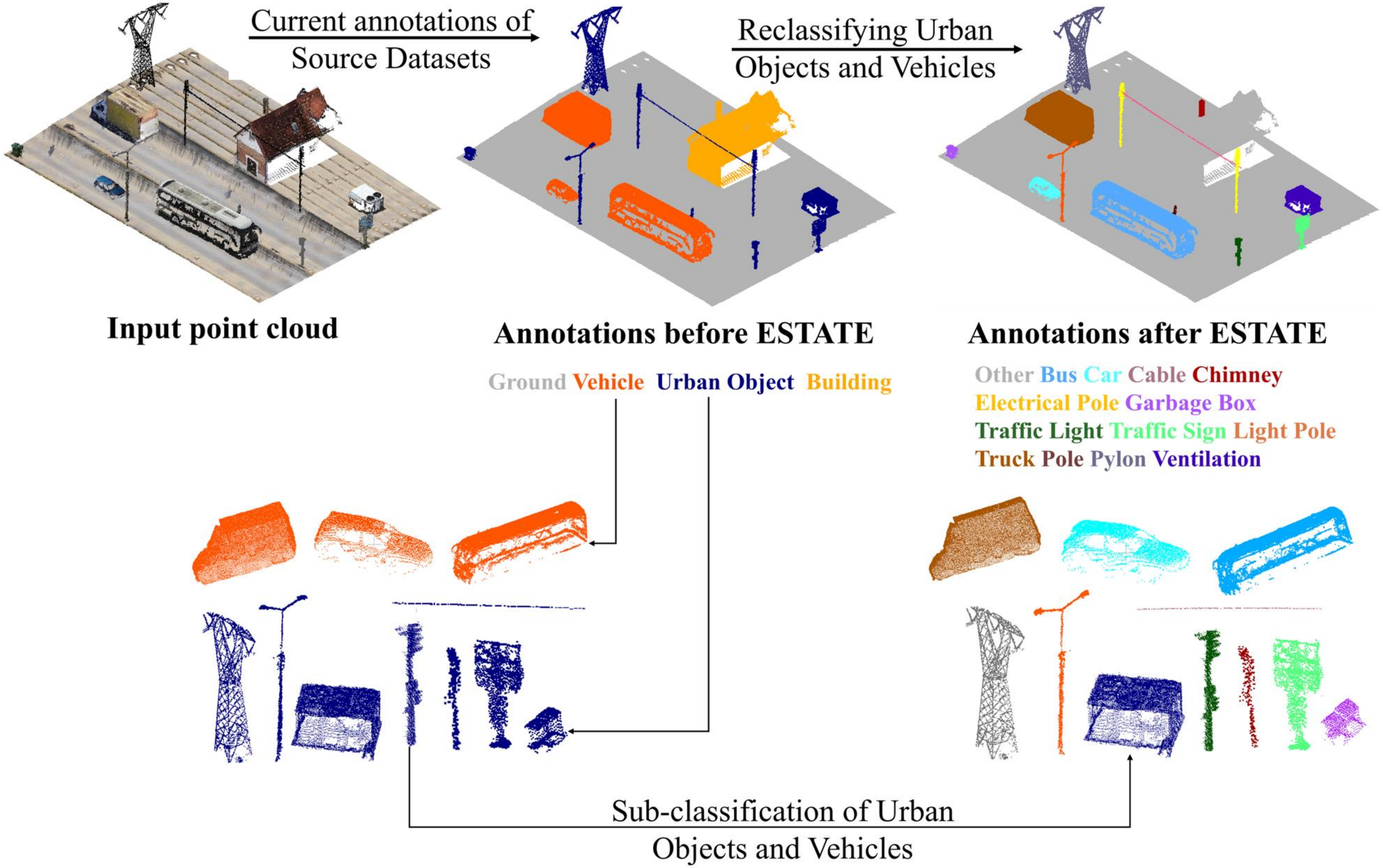
[Bayrak, O. Ma, Z., Farella, E.M., Remondino, F., Uzar, M., 2024: ESTATE: A Large Dataset of Under-Represented Urban Objects for 3D Point Cloud Classification. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2-2024, 25–32]

Point Cloud Classification - ESTATE

Dataset properties and classes	No Feature	With Intensity				With RGB & Intensity						Number of Objects in ESTATE
		ALS		MLS		UAV-Photogrammetry				ALS	MLS	
	WHU3D	DublinCity	FBK	Paris-Lille3D	TR-MLS	SensatUrban	Swiss3DCity	STPLS3D	YTU3D	Hessigheim	Toronto3D	
Approx. point density (pts/m²)	600	348	140	2000	700	400	1000	100	1000	800	1000	
Light Pole	337	258	70	52	48	8	5	116	346	32	64	1336
Traffic Light	79	3	2	15	17	-	-	16	6	-	26	164
Pole	135	71	27	24	28	-	13	67	39	18	32	454
Electr. Pole	7	-	83	2	-	-	-	43	-	-	41	176
Traffic Sign	231	5	74	124	114	9	-	36	20	14	43	670
Pylon	-	8	125	-	-	-	-	-	-	-	-	133
Cable	-	81	43	-	-	-	-	-	-	-	183	307
Garbage Box	87	-	-	162	13	369	17	-	120	66	-	834
Car	85	80	-	274	7	130	-	-	801	28	78	1483
Truck	10	-	5	-	2	20	14	6	64	-	-	121
Bus	7	30	-	-	-	3	2	2	38	-	-	82
Chimney	-	54	65	-	-	-	232	-	234	40	-	625
Ventilation	-	-	-	-	-	38	-	-	105	-	-	143
Total	978	590	494	653	229	577	283	286	1773	198	467	6528

Point Cloud Classification - ESTATE

- Support classification of classes with under-represented objects

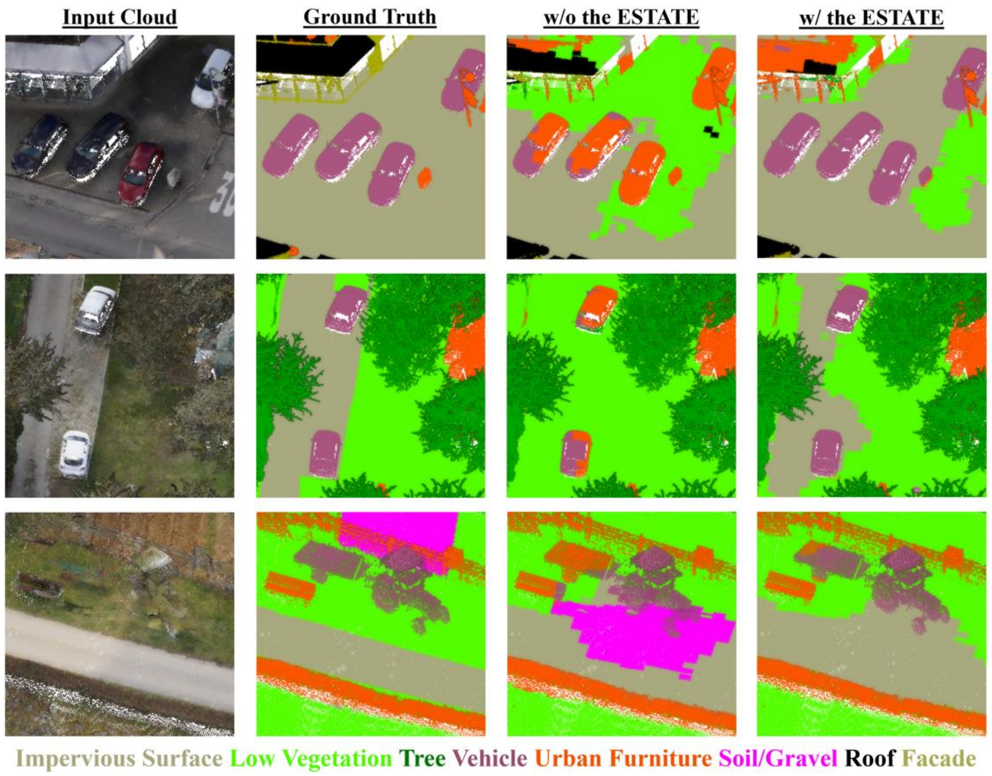


[Bayrak, O. Ma, Z., Farella, E.M., Remondino, F., Uzar, M., 2024: ESTATE: A Large Dataset of Under-Represented Urban Objects for 3D Point Cloud Classification. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2-2024, 25–32]

Point Cloud Classification - ESTATE

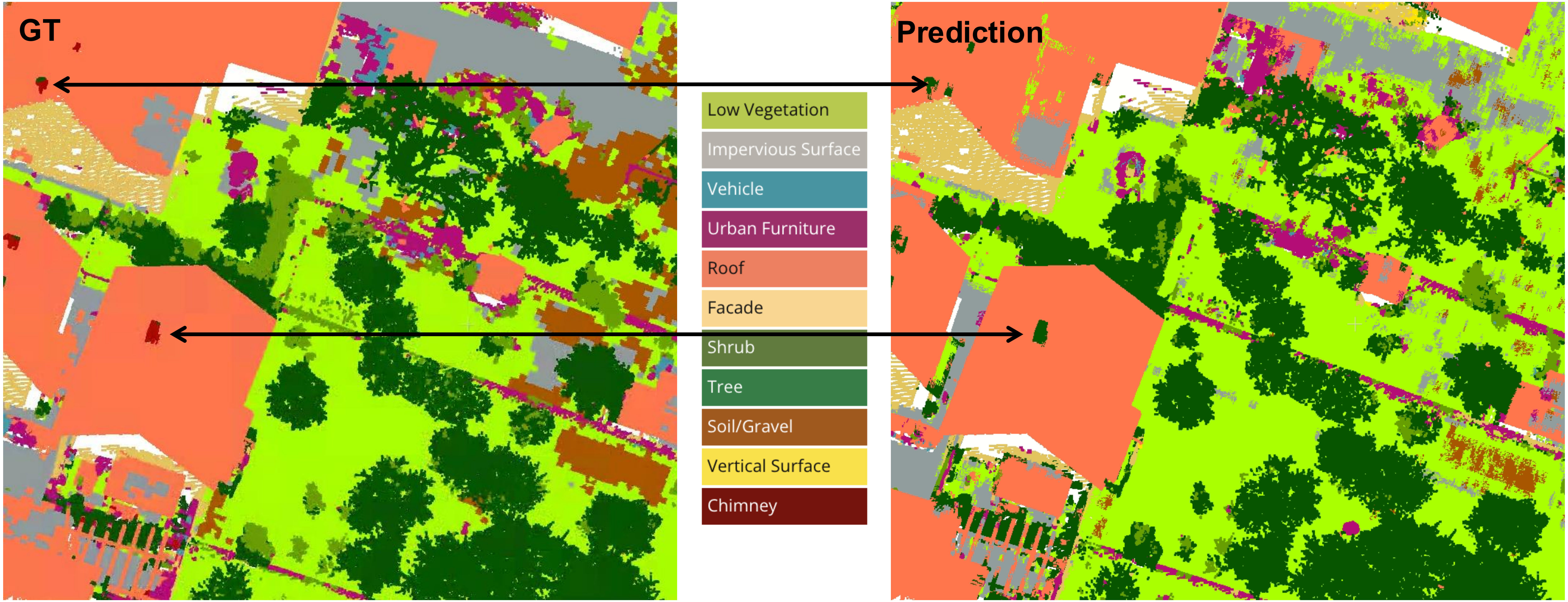
- **Hessigheim3D** benchmark: the augmentation of the Vehicle and Urban Furniture classes resulted in an improvement in classification performance by 26% and 7%, respectively

Class	w/o the ESTATE		w/ the ESTATE	
	IoU (%)	F1-Score (%)	IoU (%)	F1-Score (%)
Low Vegetation	65.07	78.84	67.98	80.94
Impervious Surface	61.84	78.28	61.21	77.94
Vehicle	32.87	49.47	**58.99**	**74.21**
Urban Furniture	31.52	48.08	**38.67**	**55.58**
Roof	85.48	92.17	82.96	90.69
Façade	63.97	78.03	65.43	79.11
Shrub	47.31	64.23	48.39	65.22
Tree	91.48	95.55	91.45	95.54
Soil/Gravel	0.04	0.07	23.18	37.64
Vertical Surface	56.88	72.51	56.75	72.41
Chimney	0.090	16.55	0.00	0.00
Mean	48.87	61.25	54.09	66.29



[Bayrak, O. Ma, Z., Farella, E.M., Remondino, F., Uzar, M., 2024: ESTATE: A Large Dataset of Under-Represented Urban Objects for 3D Point Cloud Classification. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2-2024, 25–32]

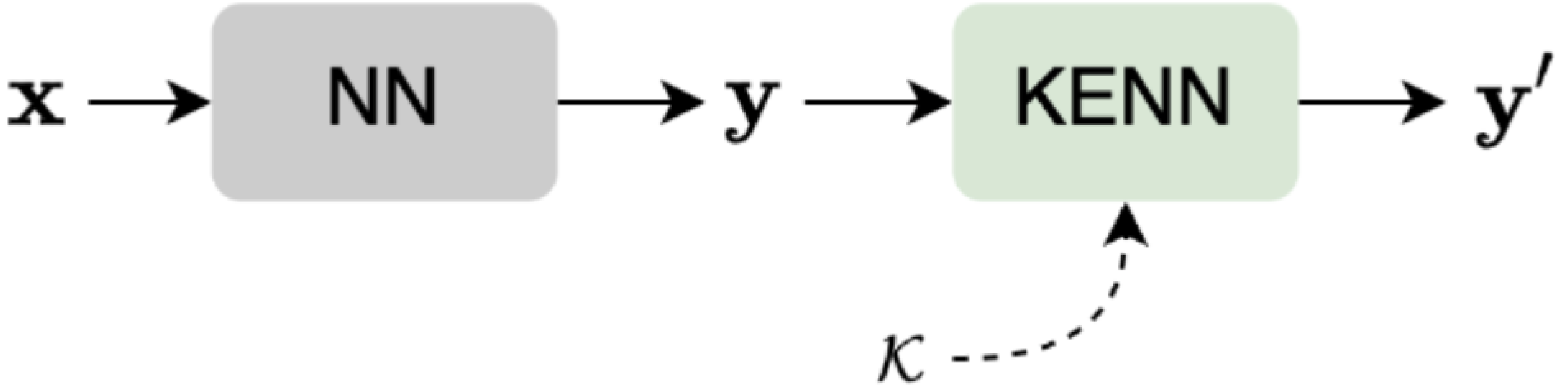
Point Cloud Classification - KENN



- KENN - Knowledge Extended Neural Network to support better prediction of under-represented classes
- inject Prior Knowledge into a neural network via First Order Logic
  - a chimney can only be on a roof
  - a tree cannot be on a roof
  - a car can only be on the ground

Point Cloud Classification - KENN

- A neural network (NN) takes a set of features ***x*** as inputs and produces an initial output ***y*** containing predictions for ***n*** classes
- The KENN layer refines the initial predictions in order to increase knowledge satisfaction and release ***y'***
- ***K*** is described as a set of logical rules (or constraints) that represent restrictions on the ***n*** classes to be predicted



$\forall u. \neg Linear(u) \vee \neg Vertical(u) \vee Pole(u)$ “every point *u* that is linear and vertical must be a pole”

$\forall u, v\ \neg Building(u) \vee \neg Over(v, u) \vee \neg Pole(v)$ “point building cannot have above a point pole”

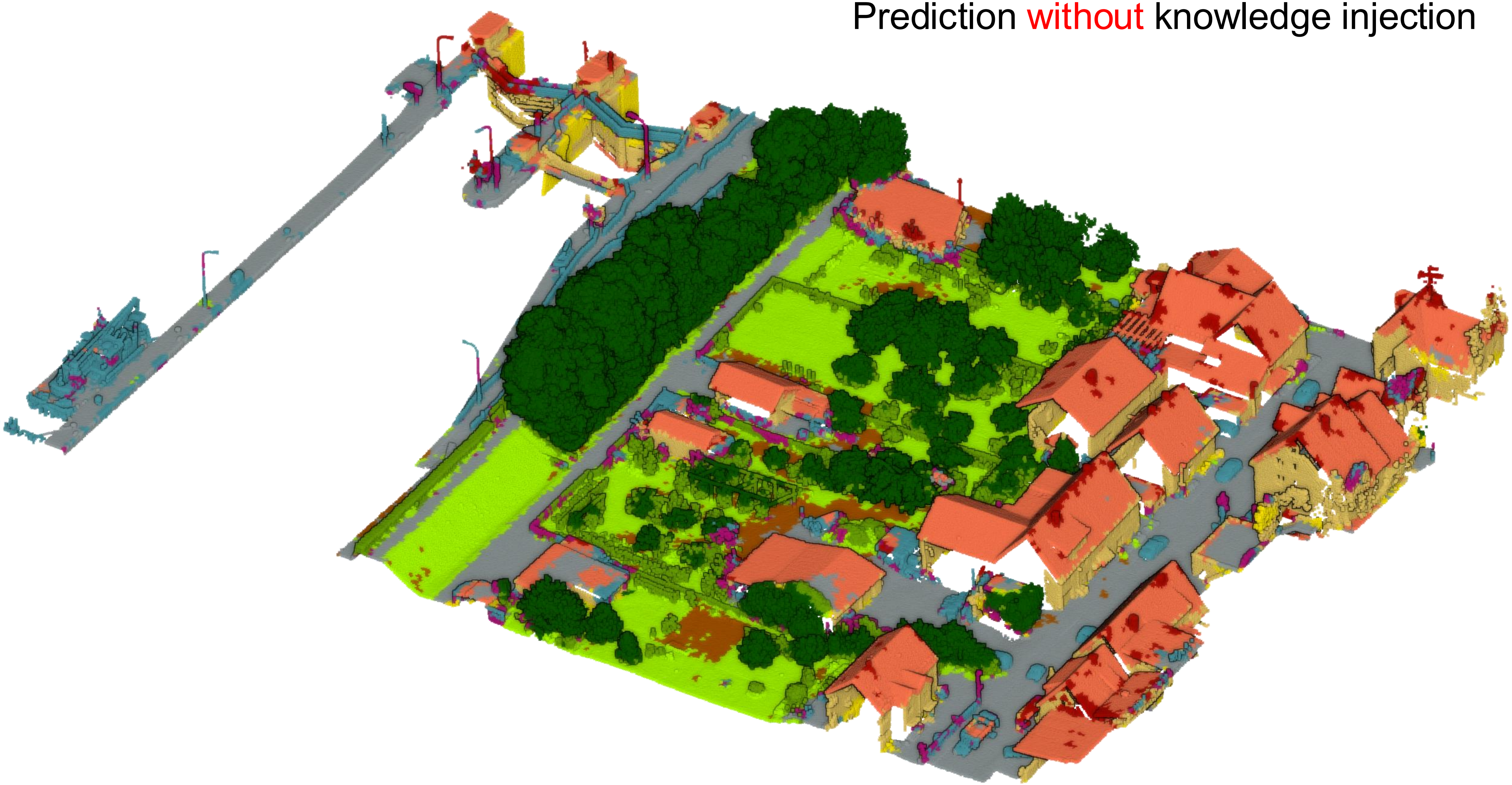
$\_ : nLinear(x), nVertical(x), Pole(x)$ “everything that is both linear and vertical (in terms of Covariance Features) within the dataset is likely to belong to the class pole”

[Grilli, E., Daniele, A., Bassier, M., Remondino, F., Serafini, L., 2023: Knowledge Enhanced Neural Networks for Point Cloud Semantic Segmentation. Remote Sensing, 15(10):2590]

Point Cloud Classification - KENN

Low Vegetation
Impervious Surface
Vehicle
Urban Furniture
Roof
Facade
Shrub
Tree
Soil/Gravel
Vertical Surface
Chimney

Prediction without knowledge injection

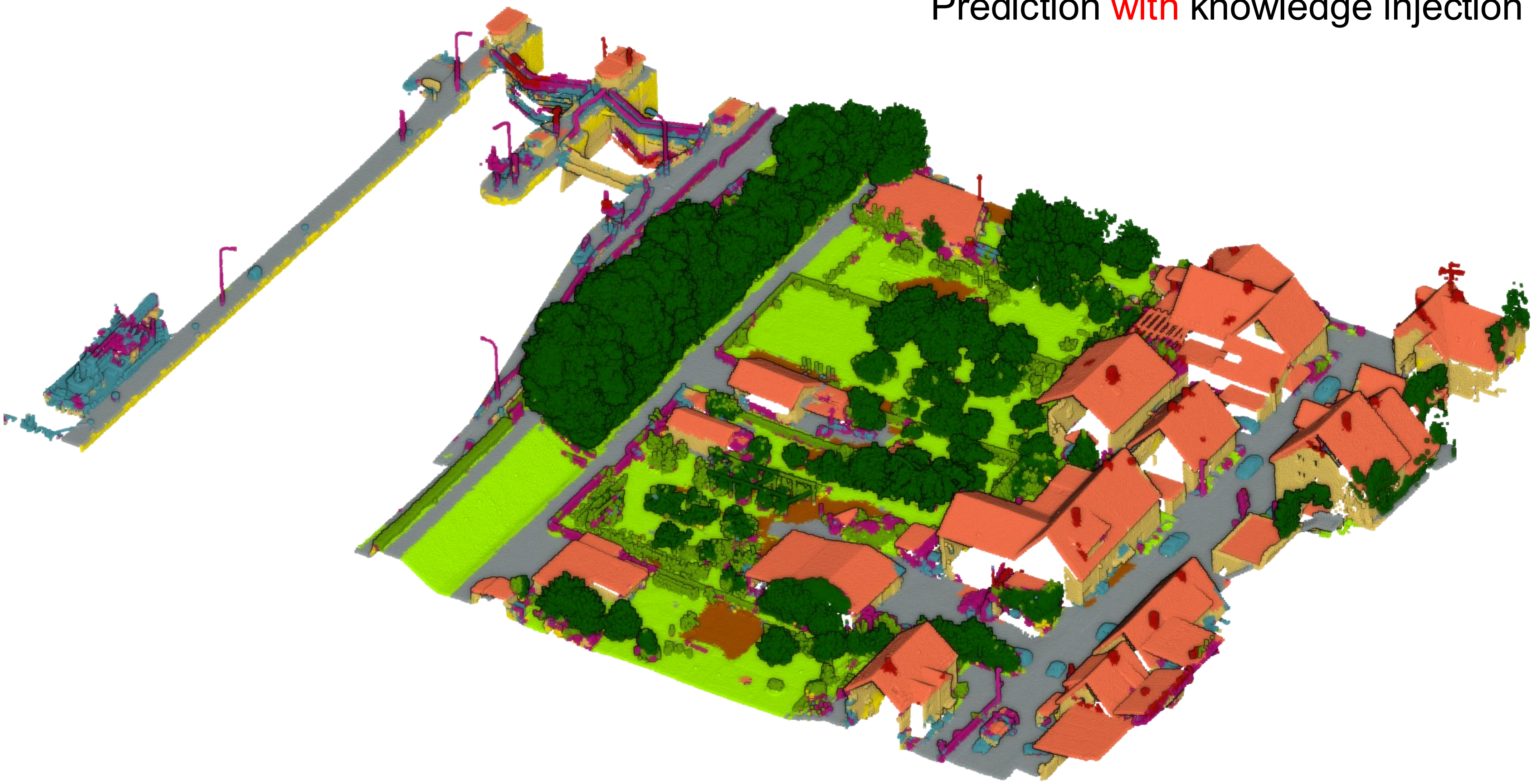


[Grilli, E., Daniele, A., Bassier, M., Remondino, F., Serafini, L., 2023: Knowledge Enhanced Neural Networks for Point Cloud Semantic Segmentation. Remote Sensing, 15(10):2590]

Point Cloud Classification - KENN

Low Vegetation
Impervious Surface
Vehicle
Urban Furniture
Roof
Facade
Shrub
Tree
Soil/Gravel
Vertical Surface
Chimney

Prediction with knowledge injection



[Grilli, E., Daniele, A., Bassier, M., Remondino, F., Serafini, L., 2023: Knowledge Enhanced Neural Networks for Point Cloud Semantic Segmentation. Remote Sensing, 15(10):2590]

Point Cloud Classification - KENN

OUTPUT ' (Before KENN)

	low veg	street	vehicle	urban furniture	roof	façade	shrub	tree	soil	vertical surface	chimney	Mean	Weighted
PRECISION	0.8519	0.786	0.352	0.3106	0.877	0.692	0.328	0.971	0.4072	0.3978	0.0895	0.5511	0.8073
RECALL	0.7966	0.827	0.783	0.133	0.812	0.744	0.557	0.875	0.4653	0.5977	0.9012	0.681	0.777
F1	0.8233	0.806	0.485	0.1862	0.843	0.717	0.413	0.92	0.4343	0.4777	0.1629	0.5699	0.7856
												OA	**0.5511**

OUTPUT ' (After KENN)

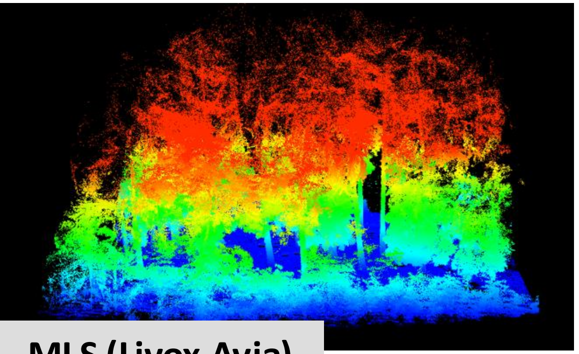
	low veg	street	vehicle	urban furniture	roof	façade	shrub	tree	soil	vertical surface	chimney	Mean	Weighted
PRECISION	0.8821	0.863	0.402	0.4302	0.894	0.776	0.437	0.976	0.5349	0.5001	0.9886	0.6995	0.8485
RECALL	0.8242	0.852	0.703	0.4474	0.921	0.735	0.651	0.918	0.4377	0.6365	0.9233	0.7319	0.8305
F1	0.8522	0.857	0.511	0.4386	0.907	0.755	0.523	0.946	0.4815	0.5601	0.9601	0.7085	0.8372
												OA	**0.8305**

- ❑ Overall improvement everywhere, in particular for under-represented classes
- ❑ Role of rules and generalization aspects

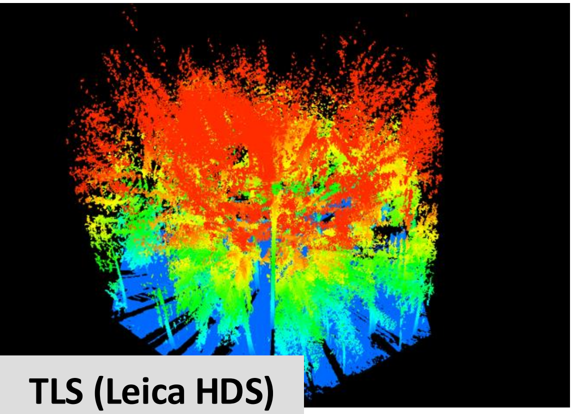
[Grilli, E., Daniele, A., Bassier, M., Remondino, F., Serafini, L., 2023: Knowledge Enhanced Neural Networks for Point Cloud Semantic Segmentation. Remote Sensing, 15(10):2590]

3D Forestry semantic segmentation

drone (DJI L1)

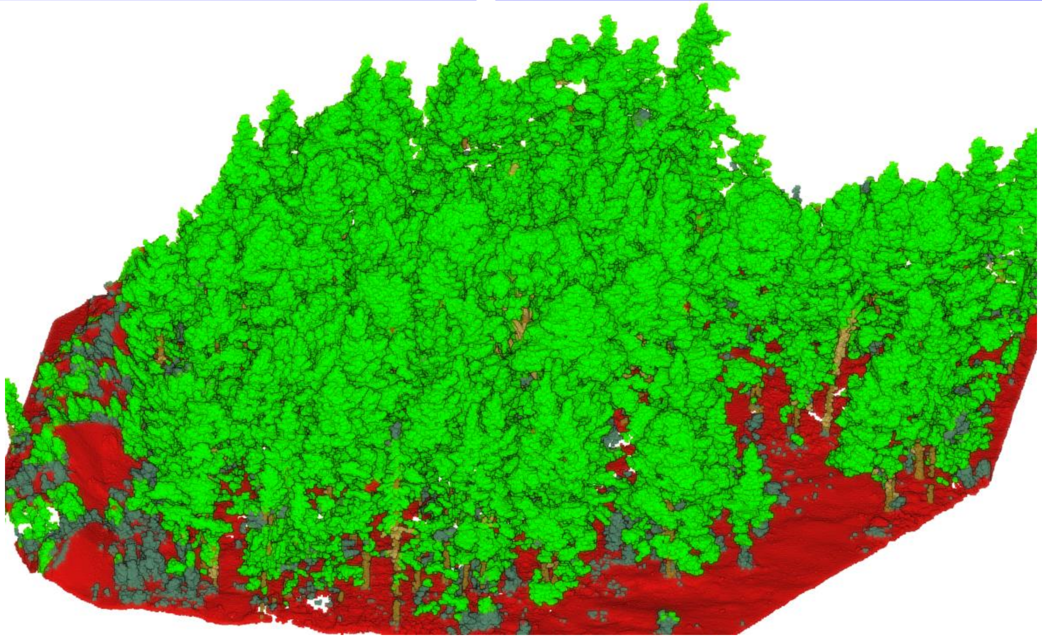


MLS (Livox Avia)

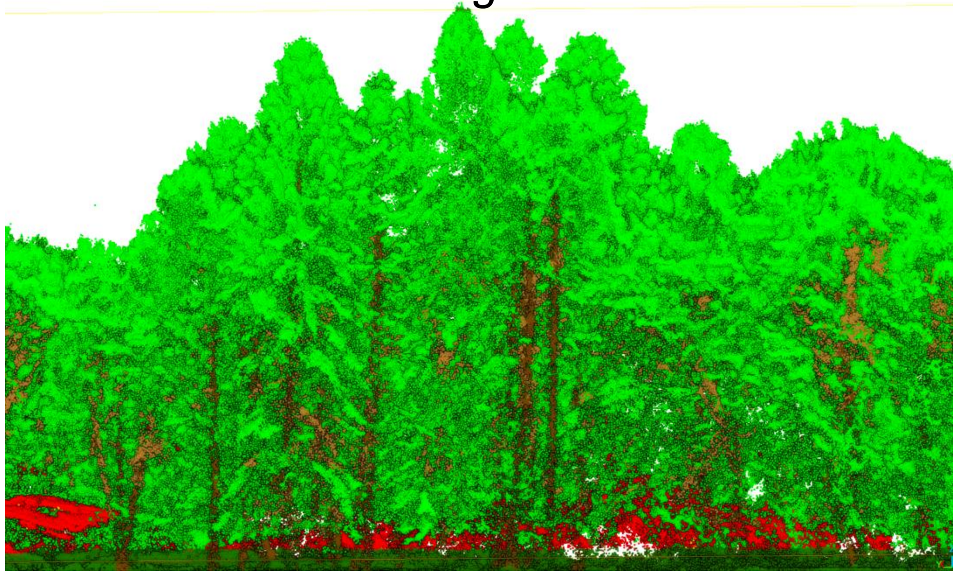


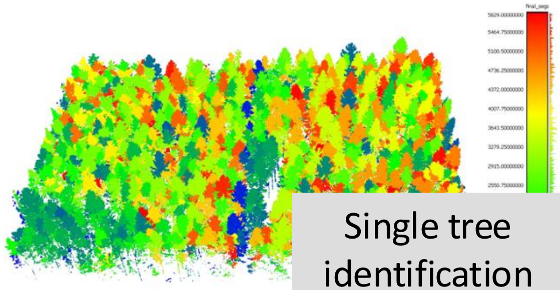
TLS (Leica HDS)

combined or stand-alone data

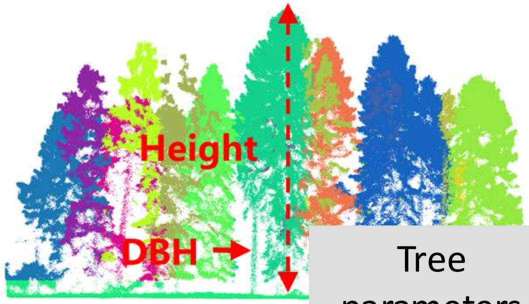


6 classes semantic segmentation with KPConv

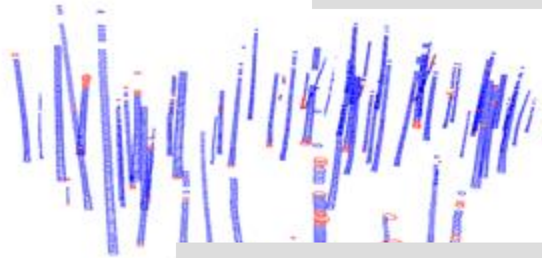




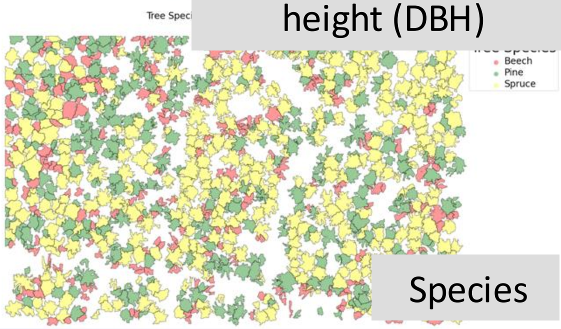
Single tree identification



Tree parameters



Diameter of breast height (DBH)



Species

3D Forestry semantic segmentation

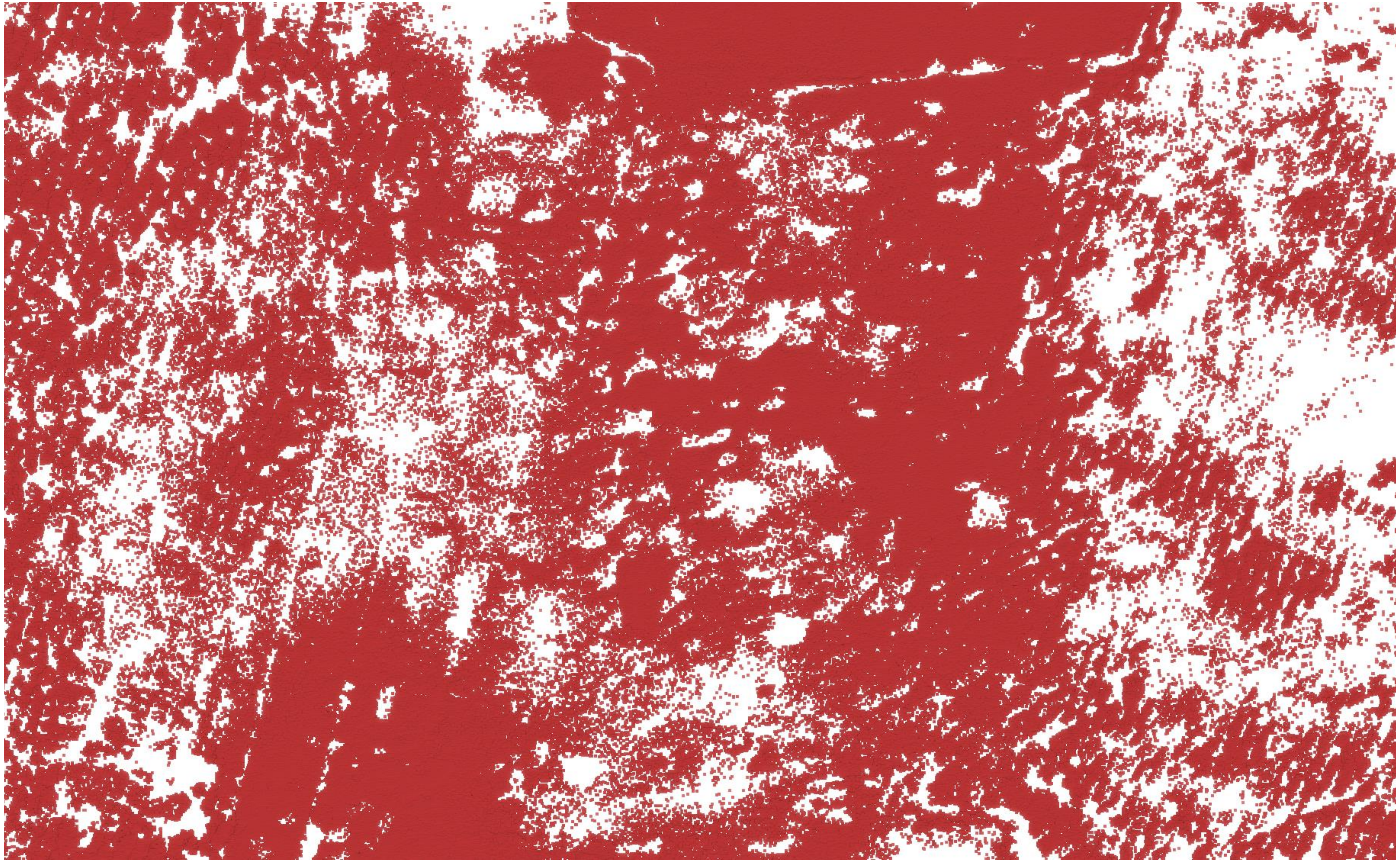
<u>GROUND</u>

TRUNKS

BRUNCHES

LOW-VEGETATION

CANOPY



3D Forestry semantic segmentation

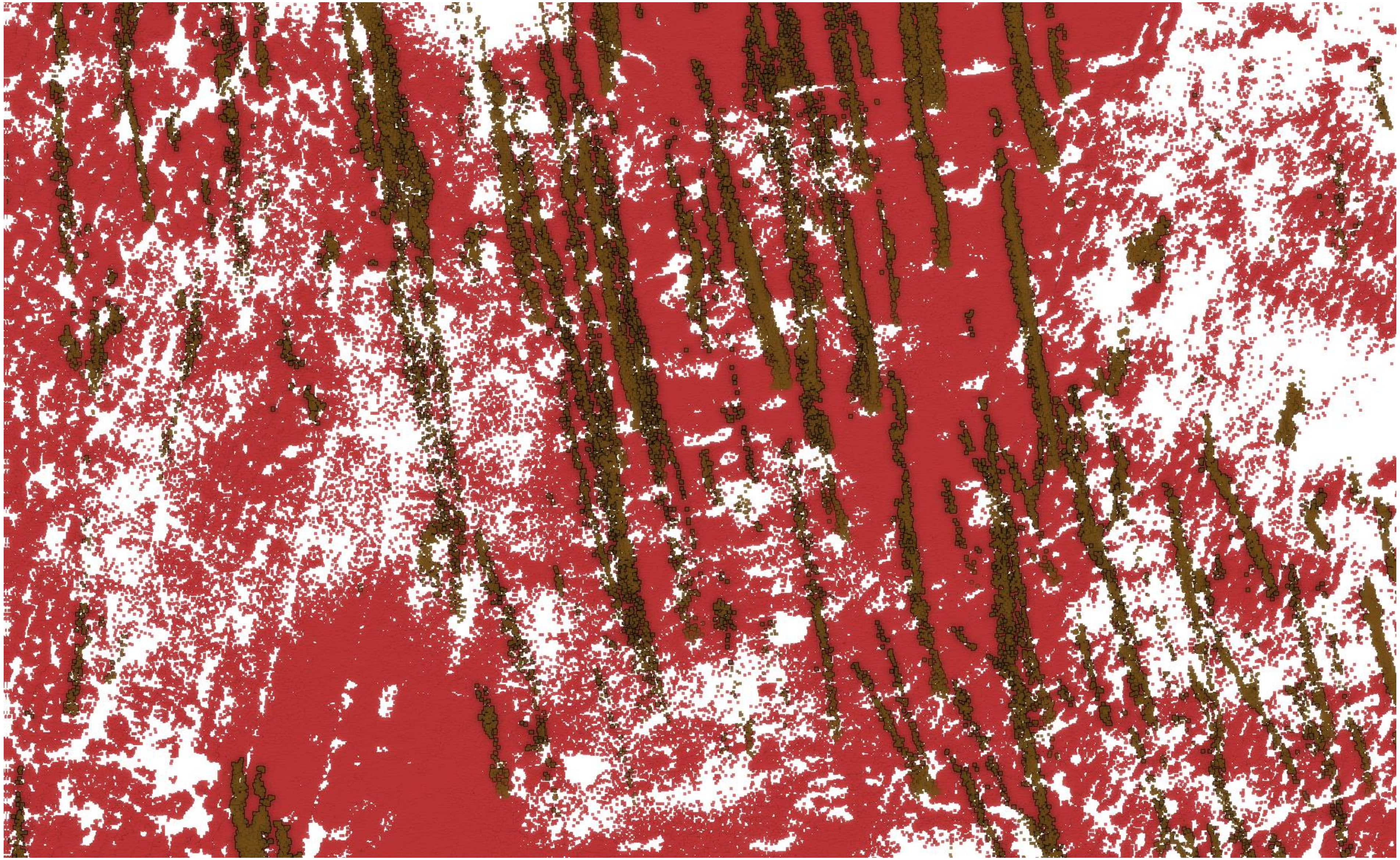
GROUND

TRUNKS

BRUNCHES

LOW-VEGETATION

CANOPY



3D Forestry semantic segmentation

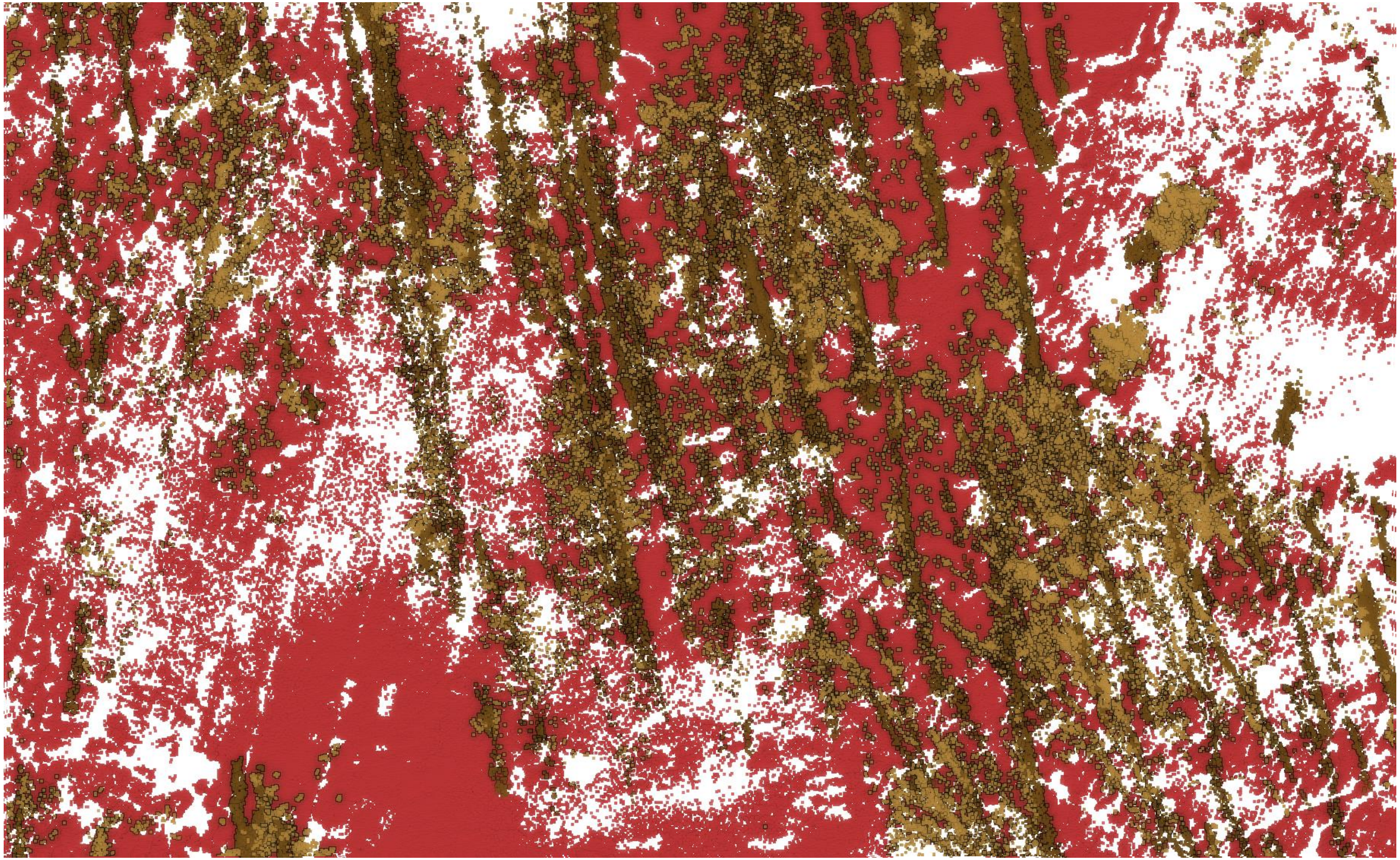
GROUND

TRUNKS

BRUNCHES

LOW-VEGETATION

CANOPY



3D Forestry semantic segmentation

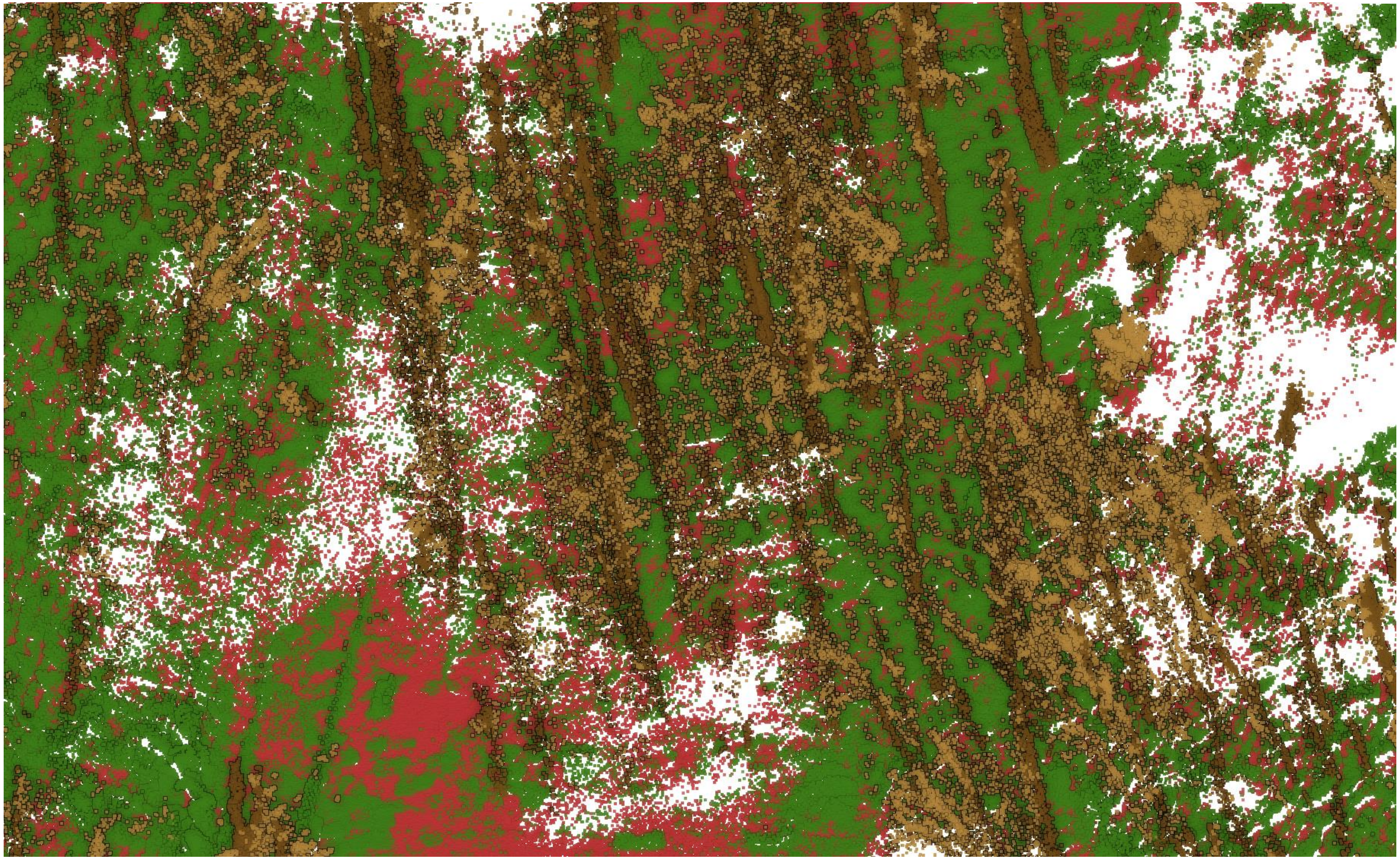
GROUND

TRUNKS

BRUNCHES

LOW-VEGETATION

CANOPY



3D Forestry semantic segmentation

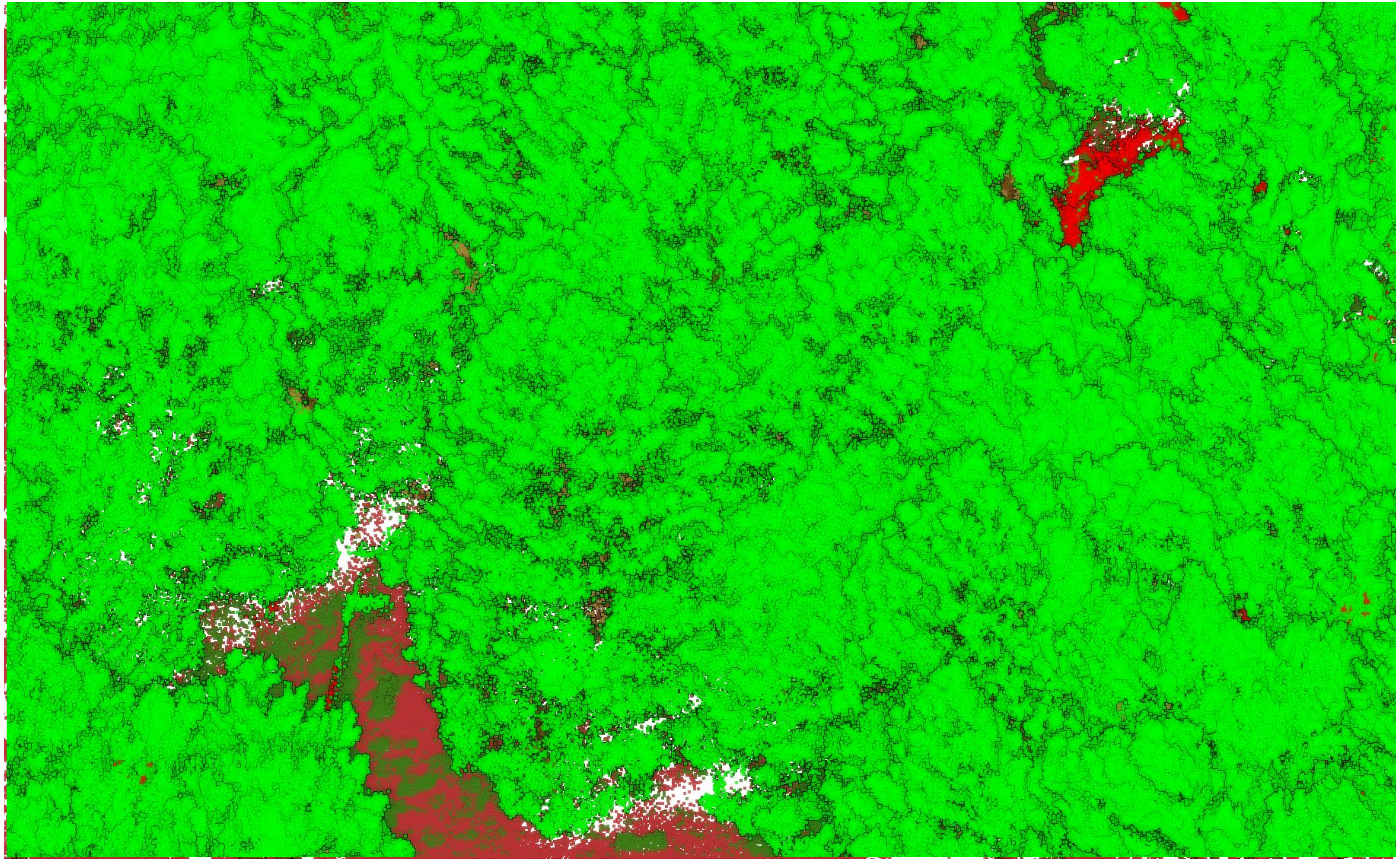
GROUND

TRUNKS

BRUNCHES

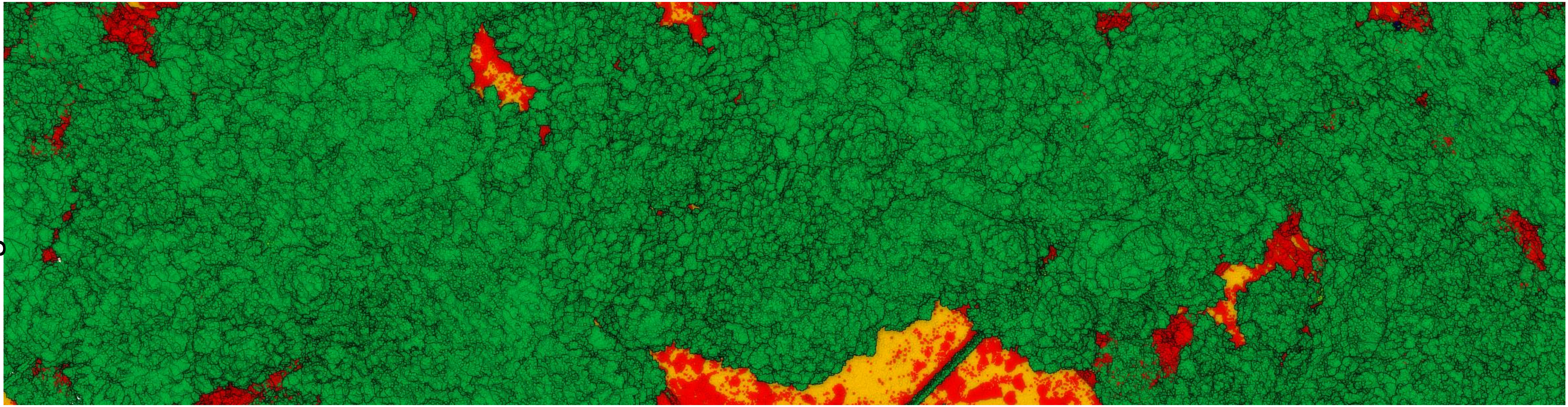
LOW-VEGETATION

CANOPY

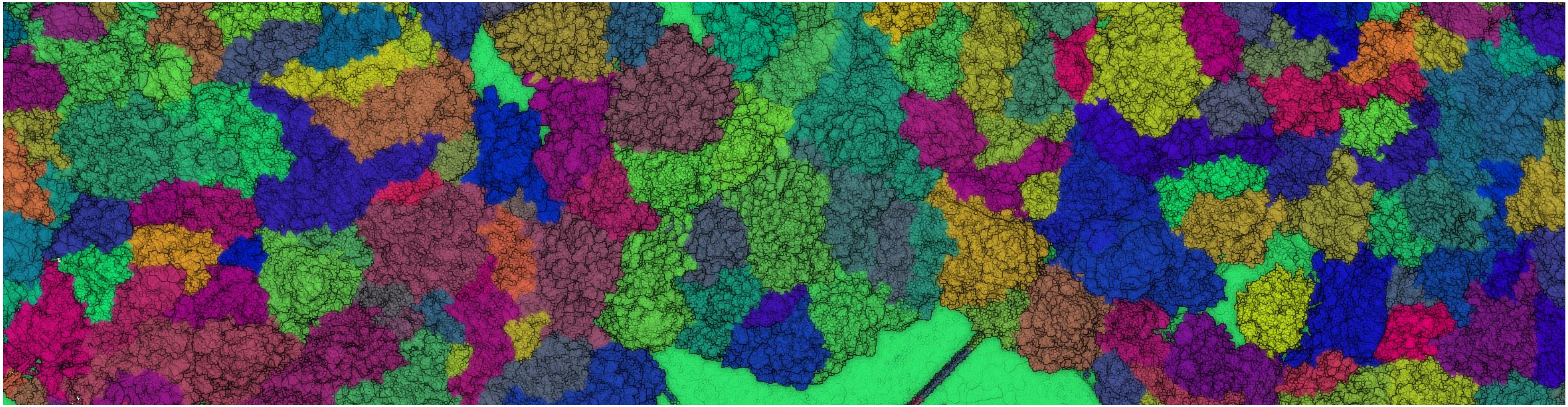


3D Forestry semantic segmentation

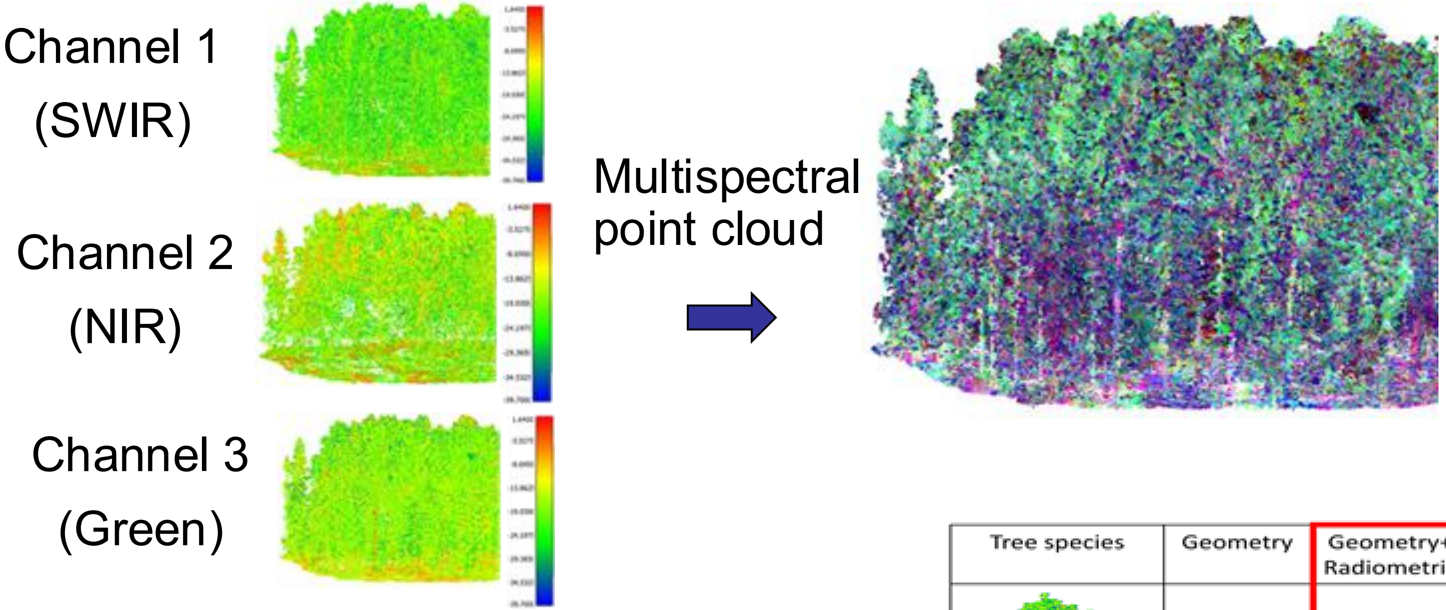
Semantic segmentation



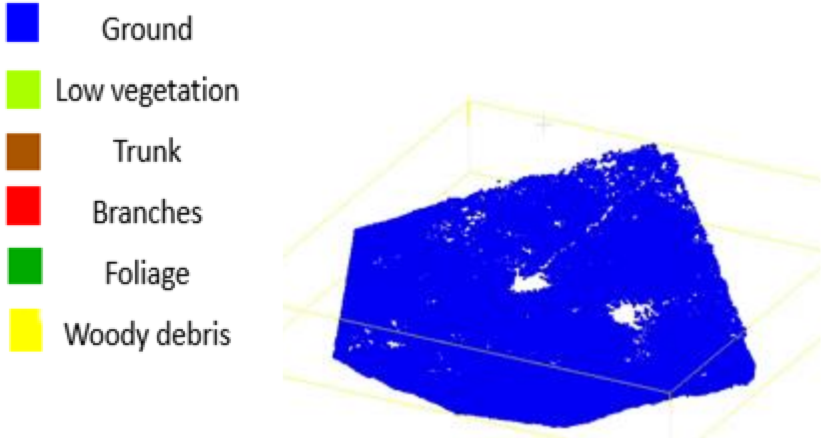
Single tree detection / tree instances



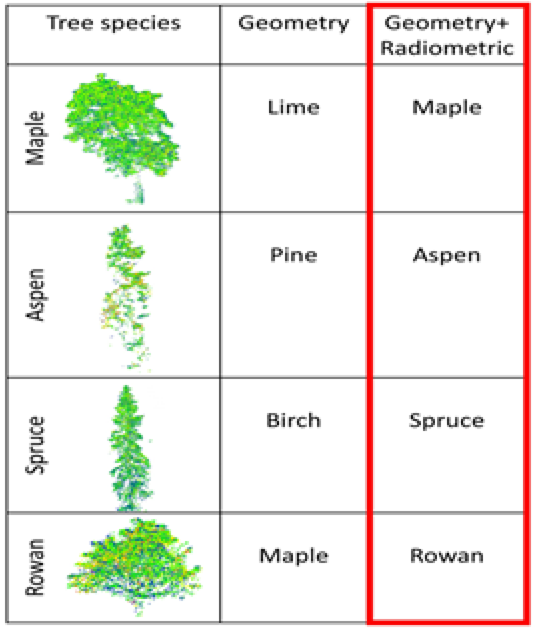
3D Forestry with Multispectral LiDAR



Forest monitoring task	Overall accuracy (%)	
	Geometry	Geometry + Radiometric
Forest segmentation	62.8	**71.50**
Tree species classification	48.15	**82.72**



Forest component segmentation (KPConv)

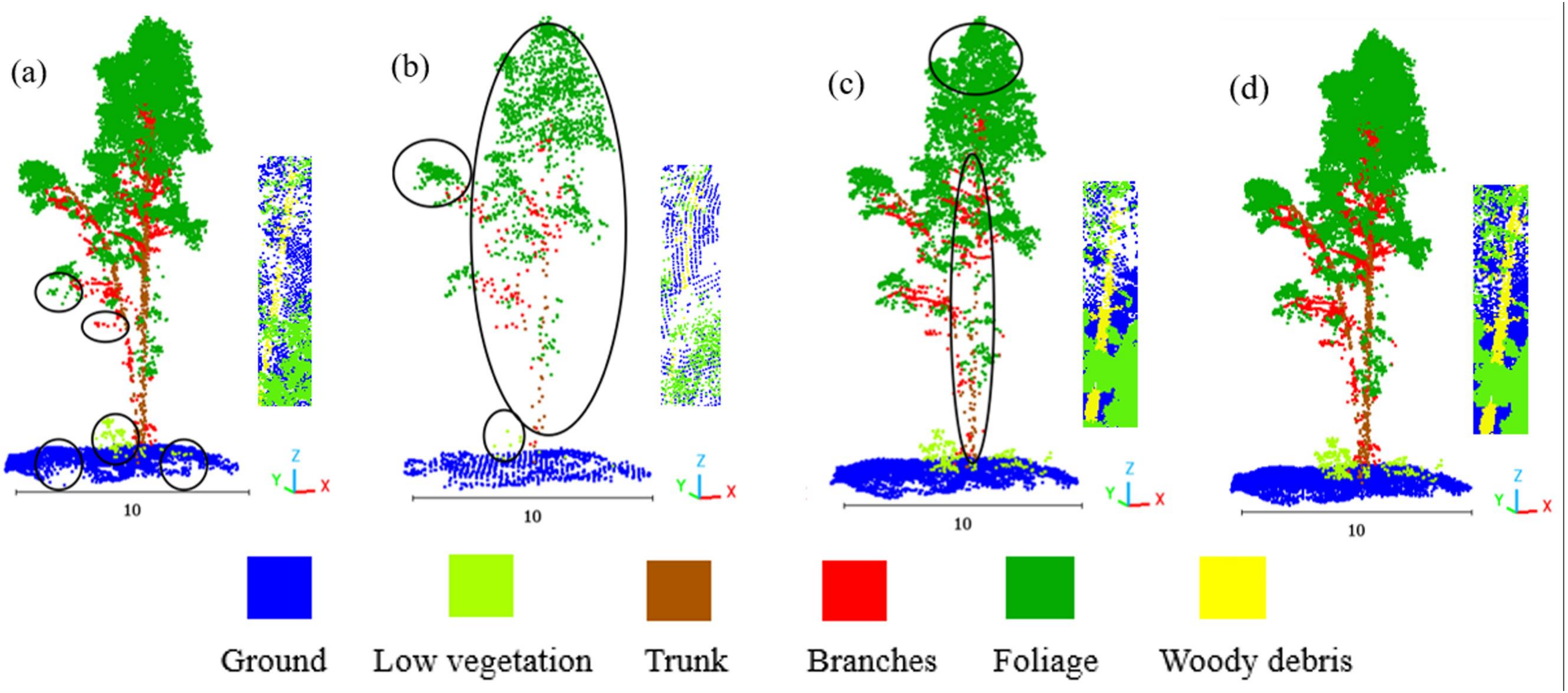


Tree species classification

[Takhtkeshha, N., Mandlburger, G., Remondino, F., Hyyppä, J., 2024: Multispectral Light Detection and Ranging Technology and Applications: A Review. Sensors; 24(5):1669]

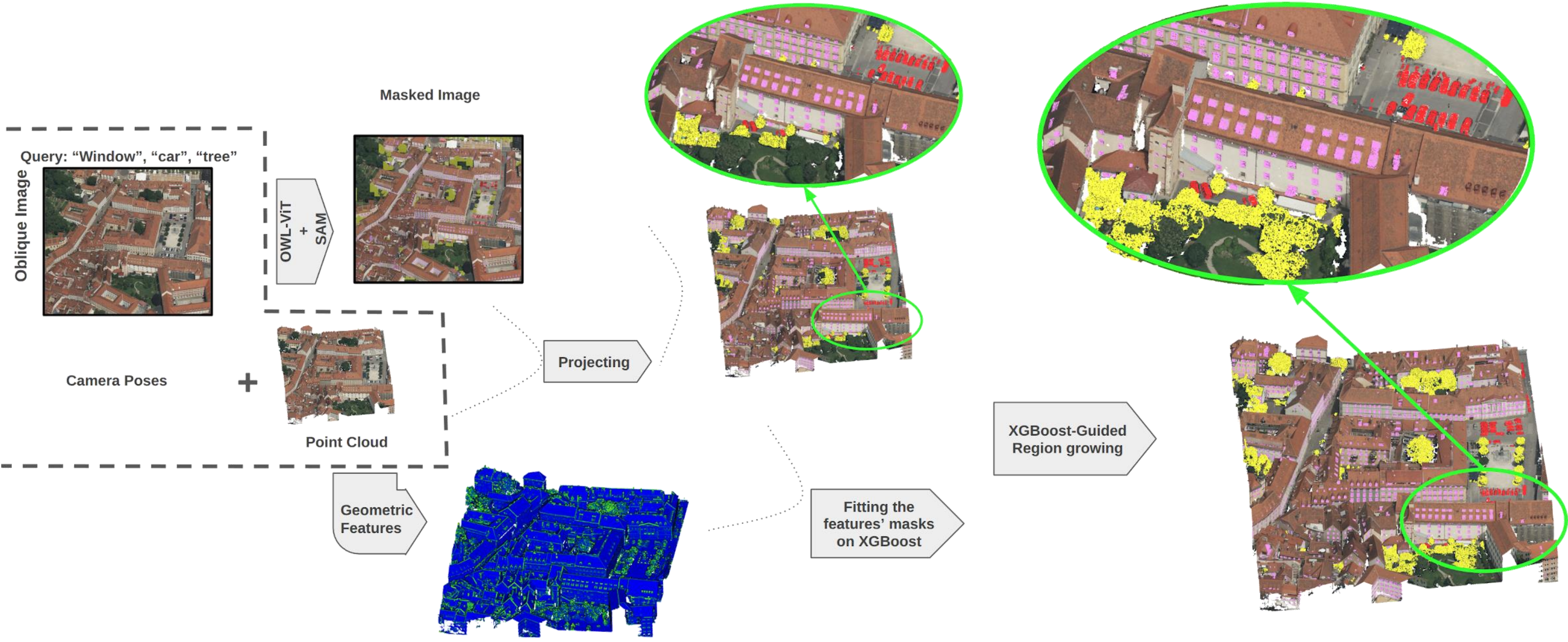
3D Forestry with Multispectral LiDAR

VUX-1HA (1550 nm) **miniVUX-1DL (905 nm)** **VQ840-G (532 nm)** **MS-LiDAR**



Semantic Segmentation with Open-Vocabulary

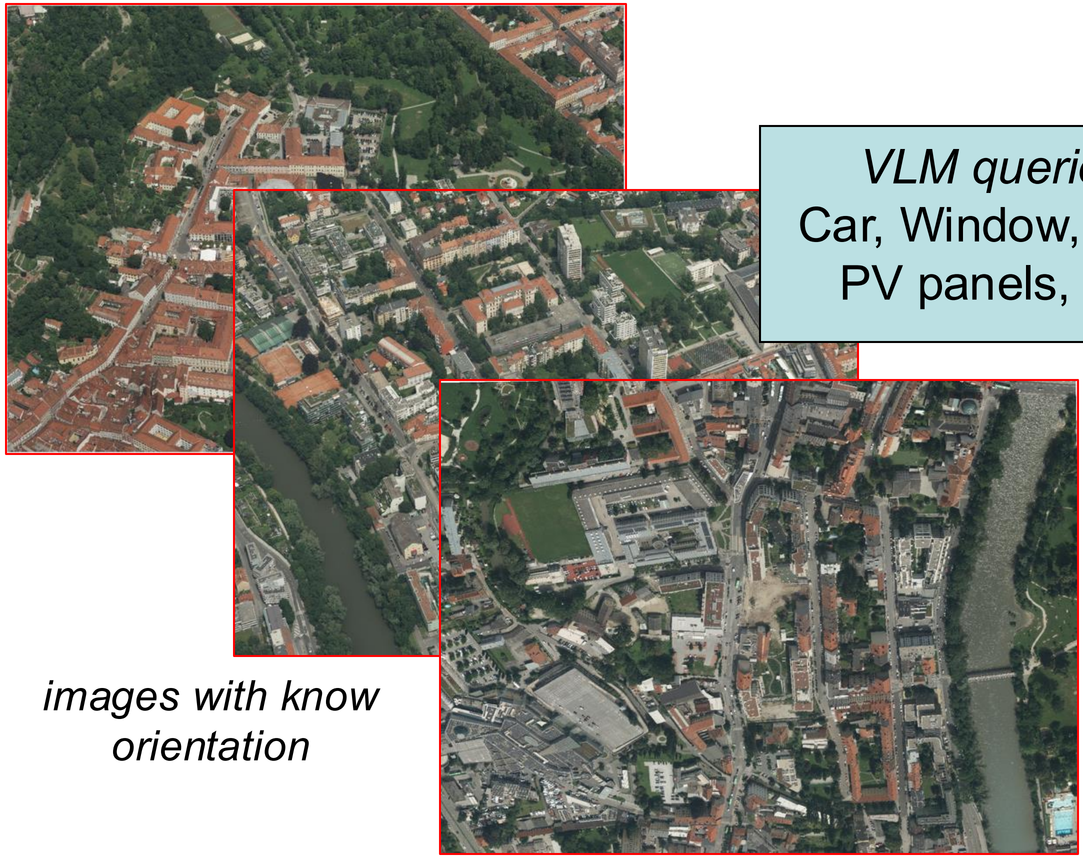
- **VLM-based** query of large scale scenarios through **oblique aerial images**



[Alami, A. and Remondino, F., 2024: **Querying 3D point clouds exploiting open-vocabulary semantic segmentation of images**. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2/W8-2024, 1–7]

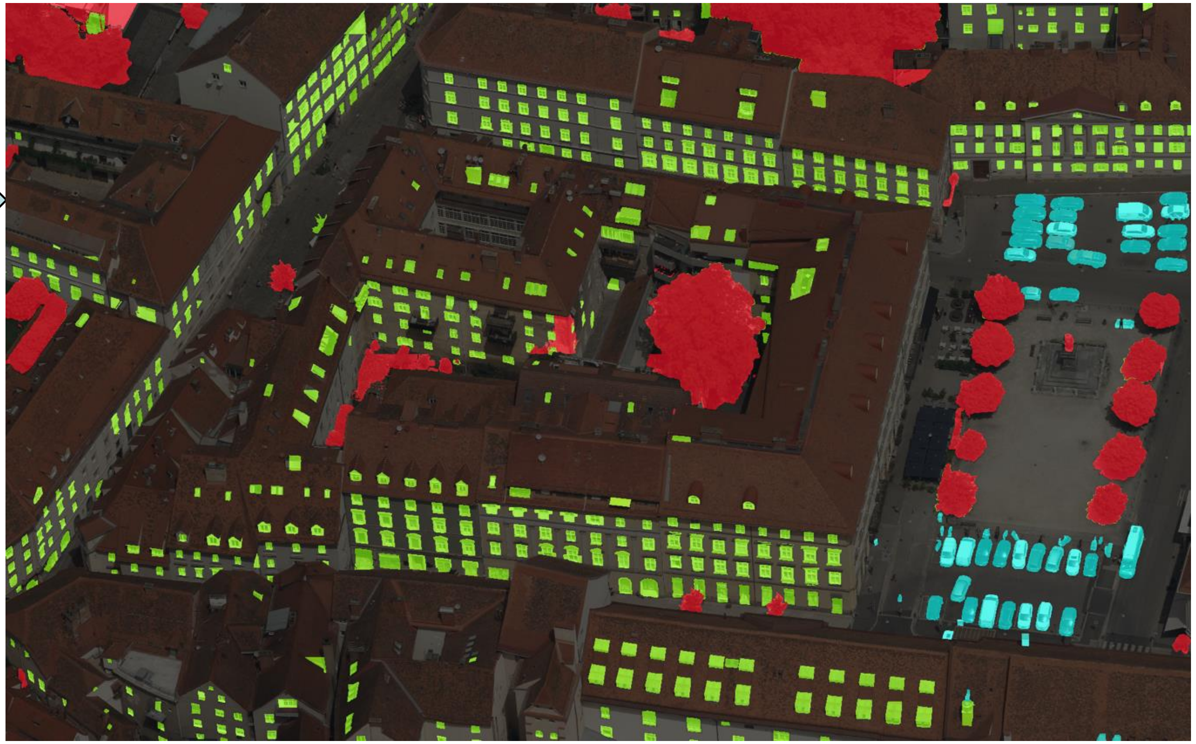
Semantic Segmentation with Open-Vocabulary

- **VLM-based** query of large scale scenarios through **oblique aerial images**
  - Yolo-World + Grounding DINO -> bounding box for the target object
  - SAM -> object mask around the target object
  - Projection onto the point clouds using camera parameters
  - Refinement using leveraging on geometric features



images with know orientation

VLM queries: Car, Window, Tree, PV panels, etc.



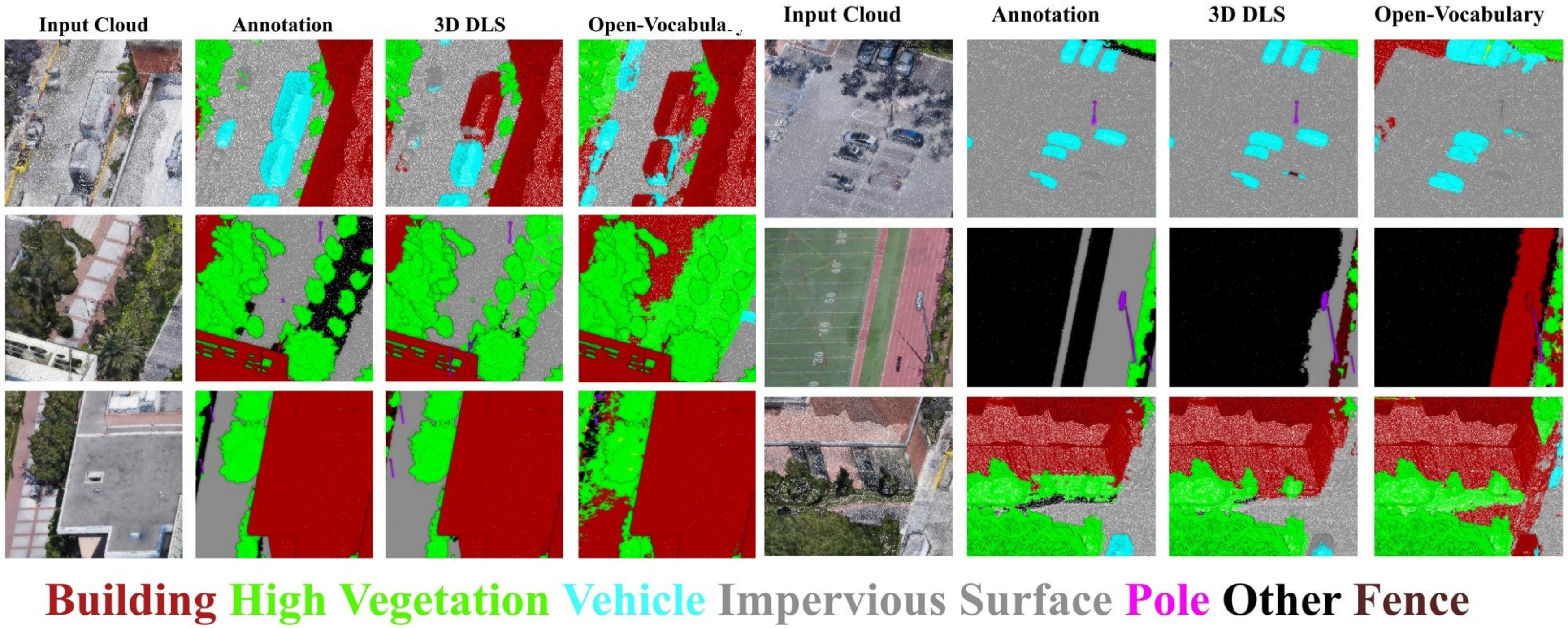
no need for 3D annotations…

[Alami, A. and Remondino, F., 2024: **Querying 3D point clouds exploiting open-vocabulary semantic segmentation of images**. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-2/W8-2024, 1–7]

Semantic Segmentation with Open-Vocabulary

- **3D deep learning** (KPConv) vs V**LM-based / Open-vocabulary** semantic segmentation

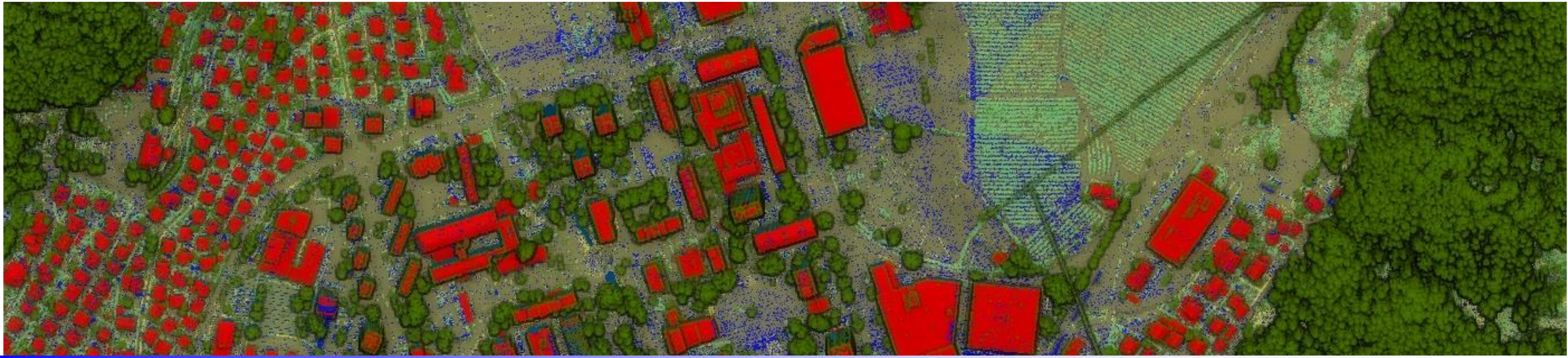
Class	3D DL		Open-vocabulary	
	IoU	F-1 Score	IoU	F-1 Score
Building	93.75	96.78	78.10	87.70
Vegetation	86.23	92.60	67.40	80.53
Vehicle	47.44	64.36	37.61	54.66
Poles	50.62	67.21	19.18	31.93
Fence	26.41	41.78	2.01	3.94
Imp. surface	65.92	79.46	12.87	22.80
Other (*)	31.32	47.71	23.18	37.63
Mean	57.38	69.94	34.34	45.60



- So far, 3D DL performs better than LLM/Open Vocabulary methods
- LLM could be complementary and / or support annotations

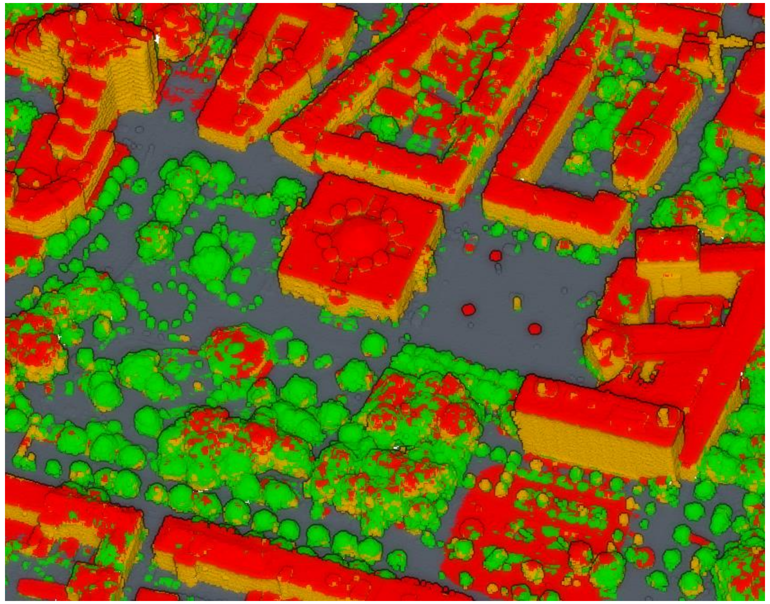
TAKE AWAY MESSAGES

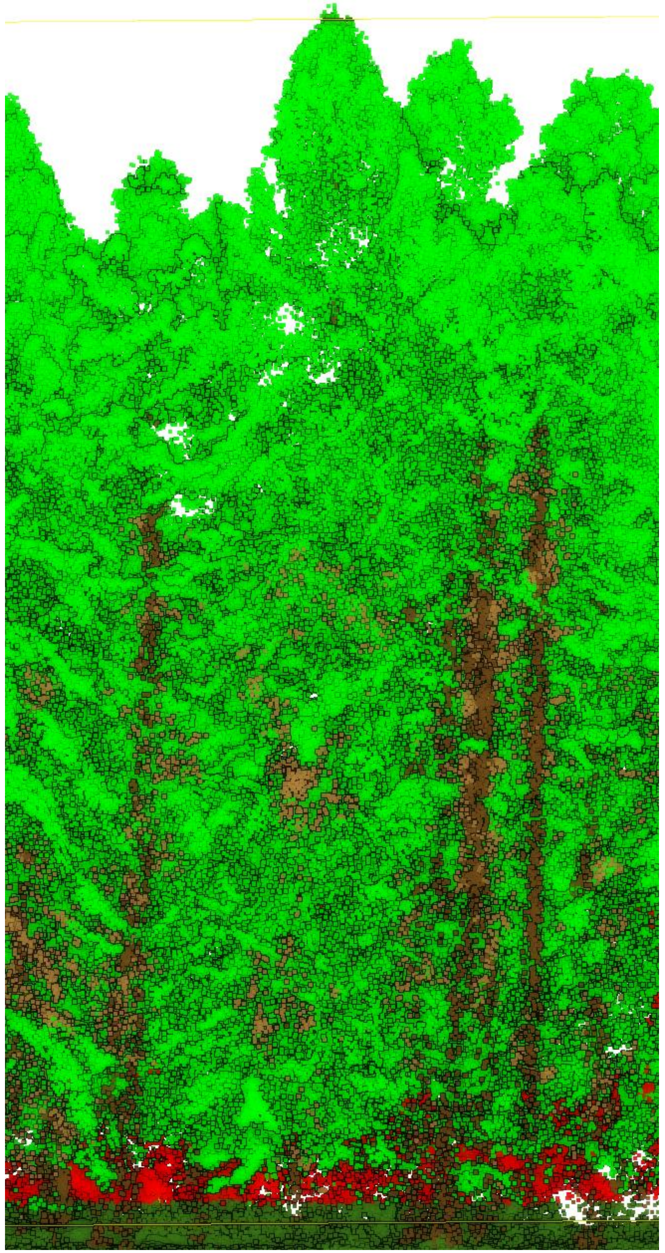
- Point cloud classification / semantic segmentation is still a very hot and open research task
- Methods (publications) fight for few % more (93.2% vs 93.4%), better consider:
  - generalization (ability to process any unseen scenario)
  - classes standardization (at least for urban mapping?)
  - fully unsupervised methods
  - under-represented objects
  - explainability



WHAT'S NEXT?

- Strong need to develop easy and standardized procedures for real-world daily-based large-scale 3D applications
- Role of LLM / VLM / MLLM, as complementary to deep learning networks
- Foundation models for 3D geospatial data
- More collaborations with colleagues in neighboring disciplines and with end-users (companies and NMCAs)





THANK YOU FOR YOUR ATTENTION

EUROSDR WORKSHOP ON MULTISPECTRAL LIDAR

23 June 2025 (Online only - 14:30 - 17:30 CET)

Organizers: Juha Hyyppä, Gottfried Mandlburger, Fabio Remondino, Narges Ttakhtkeshha

Light Detection and Ranging (LiDAR) is a well-established active technology for the direct acquisition of 3D data. In recent years, the emergence of Multispectral LiDAR (MSL) systems, which operate with two or more wavelengths, is revolutionizing the simultaneous acquisition of height and intensity information. MSL sensor properties (e.g., wavelength, instrument size and measurement range) are selected with respect to the intended application and domain which vary from forestry mapping, Land Use Land Cover (LULC) classification or change detection to bathymetry, topographic mapping, archaeology and geology.

MSL sensors provide information on the full 3D distribution of materials with improved penetration capacity hence they are becoming a valuable solution for geospatial data acquisition for NMCAs.

This online EuroSDR workshop on Multispectral LiDAR aims to review the emerging MSL and its possible application, with technical presentations and best practices from companies and mapping agencies.

NLS NATIONAL LAND SURVEY OF FINLAND

FONDAZIONE BRUNO KESSLER
FUTURE BUILT ON KNOWLEDGE

TU WIEN

GEO

EuroSDR