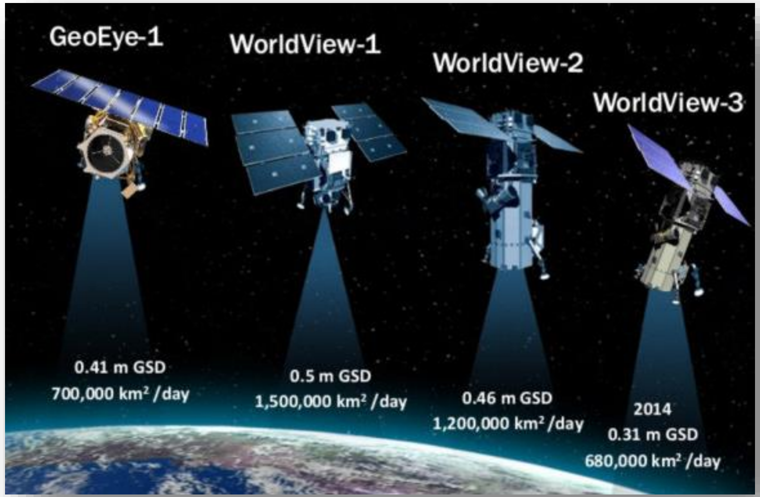
Towards an End-to-end LoD1 Building Reconstruction Pipeline from VHR Satellite Imagery

Chloé THENOZ

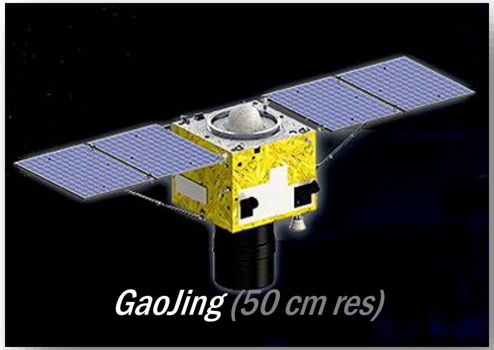
Context

More and more VHR EO satellites...



DigitalGlobe

Airbus Defense and Space

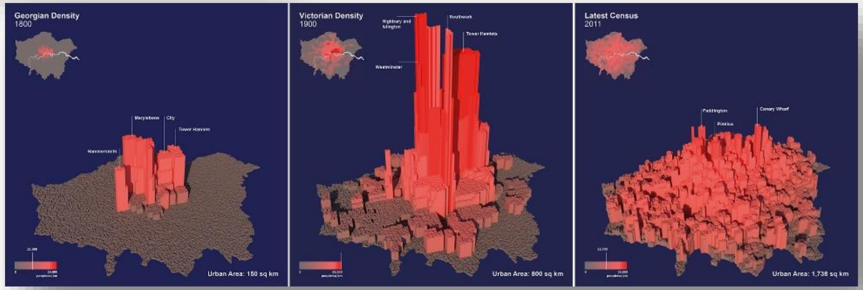


Beijing Space View Tech Co Ltd

... For numerous civil applications

Climate Change

Urban Growth monitoring



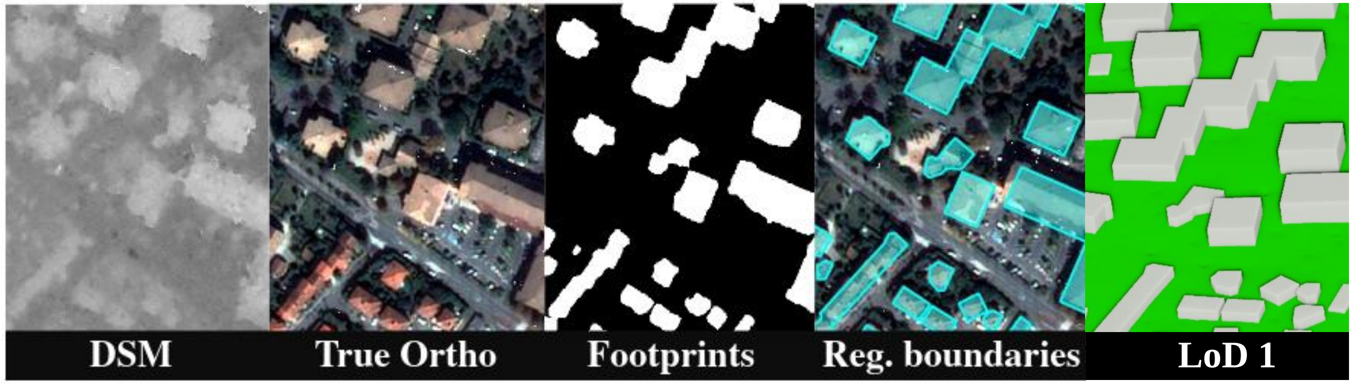
Smart Cities

...and so on!

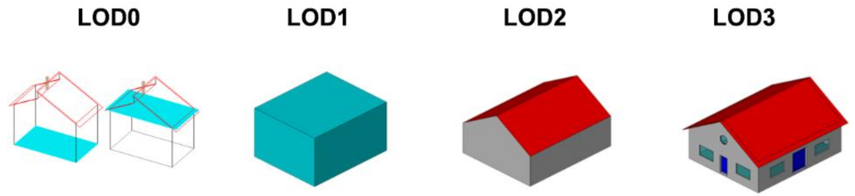
Towards large scale automatic reconstruction

Main contributors: Chloé Thenoz[1], Pierre Lassalle[2], Gwennaël Matot[1], Clément Dechesne[1], David Youssefi[2]
[1]Magellium, [2]CNES

Work inspired by Leotta, et al. "Urban Semantic 3D Reconstruction From Multiview Satellite Imagery." 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) (2019): 1451-1460.



Our LoD1 pipeline illustration

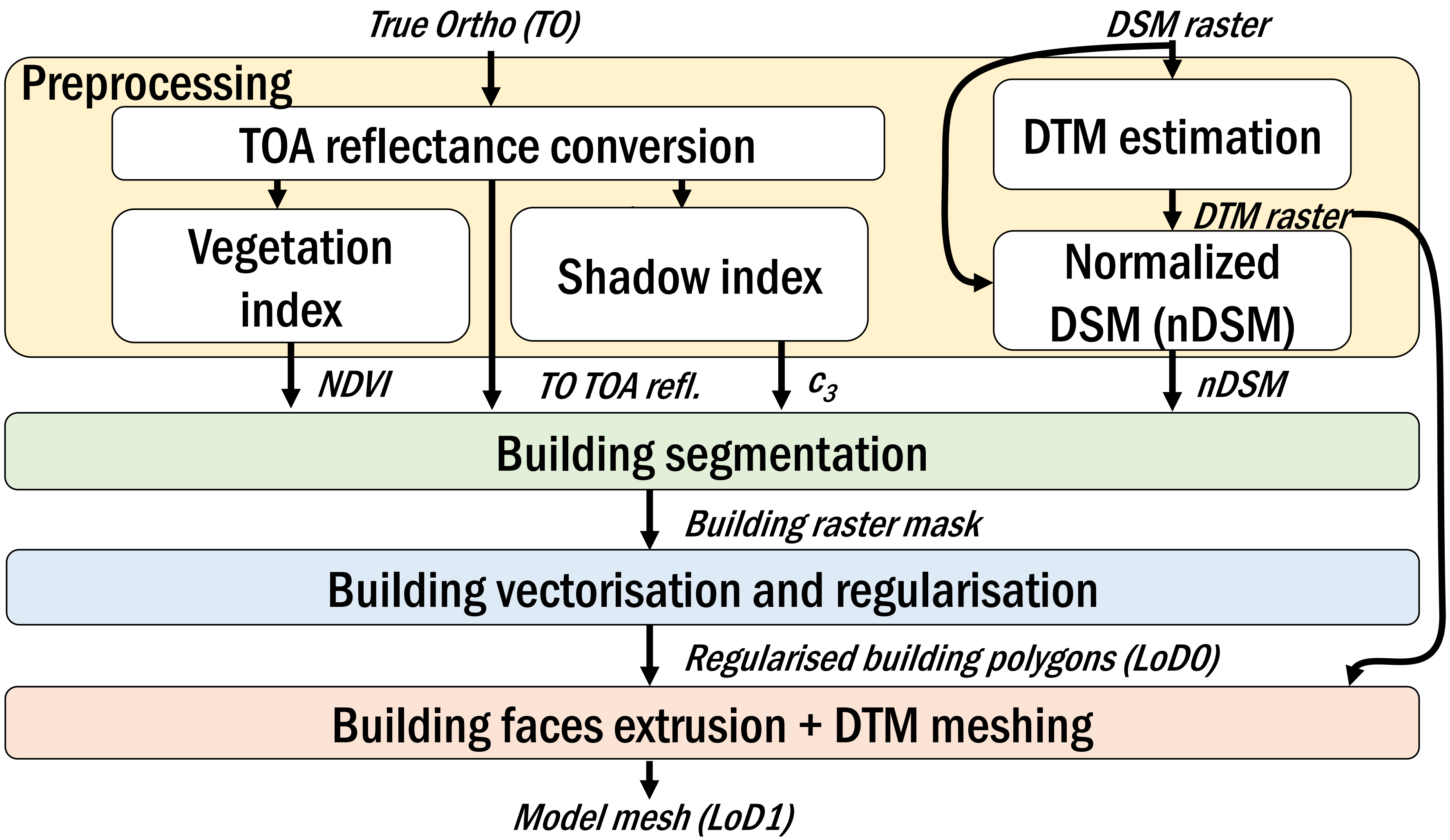


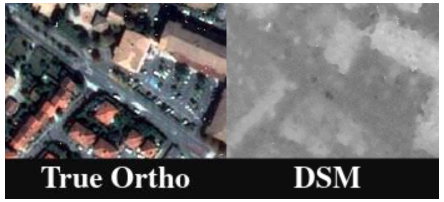
CityGML Level of Detail (LoD) standard (Open Geopsatial Consortium)

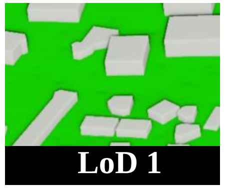
Challenges:

- Imagery resolution (Pleiades, 50cm, 70cm native GSD)
- Automatic
- Large scale

LoD1 Building Reconstruction Pipeline Overview







Preprocessing

Preprocessing on True Ortho

TOA Reflectance

Vegetation Index (NDVI)

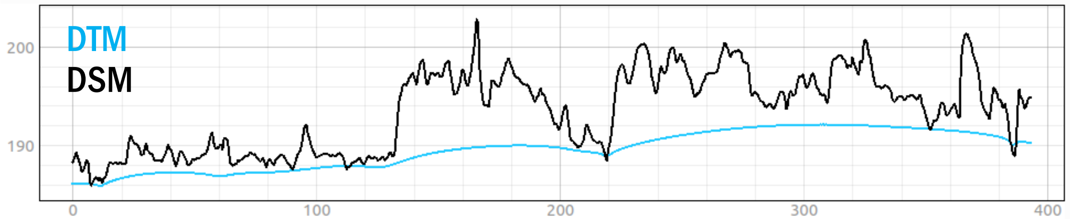
$$\mathrm{NDVI} = \frac{\mathrm{Nir} - \mathrm{Red}}{\mathrm{Nir} + \mathrm{Red}}$$

Shadow index (c3 band)

$$c_3 \text{ band} = \arctan\left(\frac{\mathrm{Red}}{\max(\mathrm{Green}, \mathrm{Blue})}\right)$$

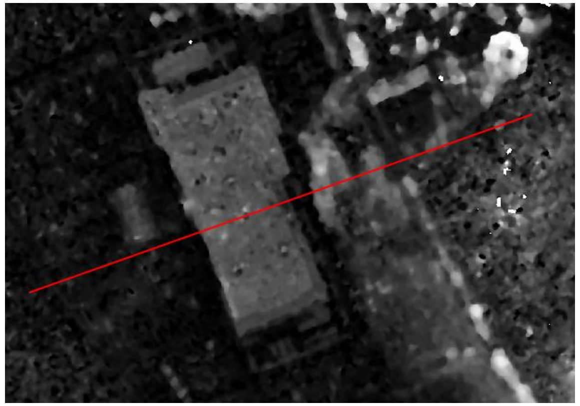
Preprocessing on DSM

DTM estimation with Leotta et al. DrapCloth image-based implementation



Altimetric profile along the red line (see below)

True Ortho



nDSM

nDSM (normalized DSM) = DSM - DTM

Building footprint segmentation: Search for better contours

Deep learning approach

Four axes of study

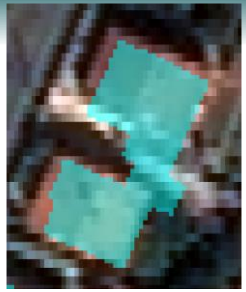
How to get better contours?

- Input bands
- Network architecture
- Database quality
- Loss function

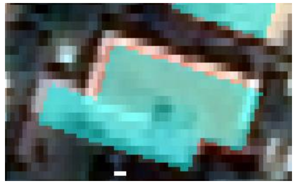
Training database used:

- **Inputs**: True ortho + DSM extracted from Pleiades Mono Stereo on Toulouse, France (~80km²)
- **Groundtruth**: OpenStreetMap building mask roughly spatially calibrated by a uniform translation (imperfect GT)

Inconsistency (destroyed building)



Wrong input

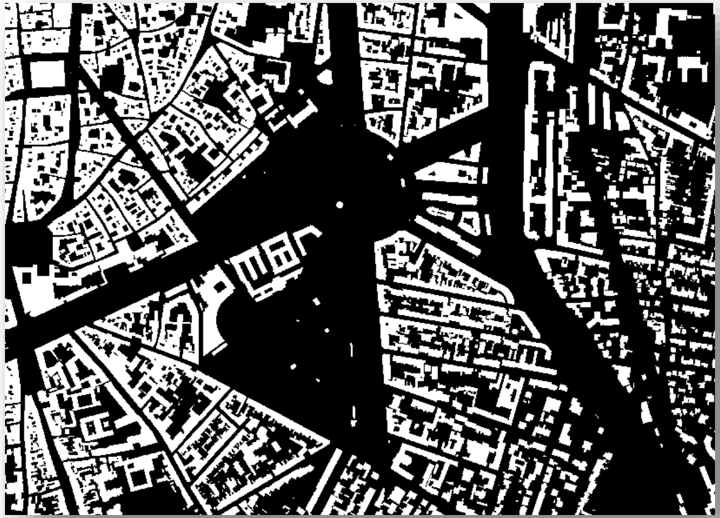


Shift between footprint and building

Common mismatches between inputs and GT

True Ortho

DSM



Building segmentation groundtruth

Building footprint segmentation: Four axes of study

Input bands

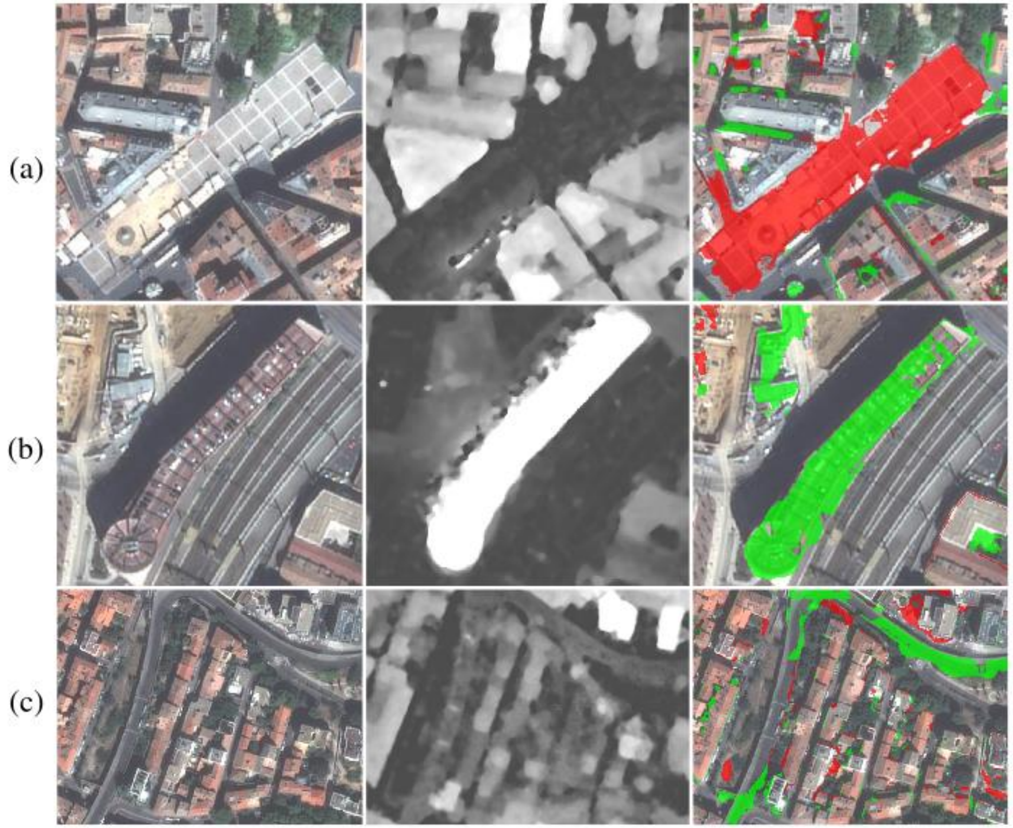
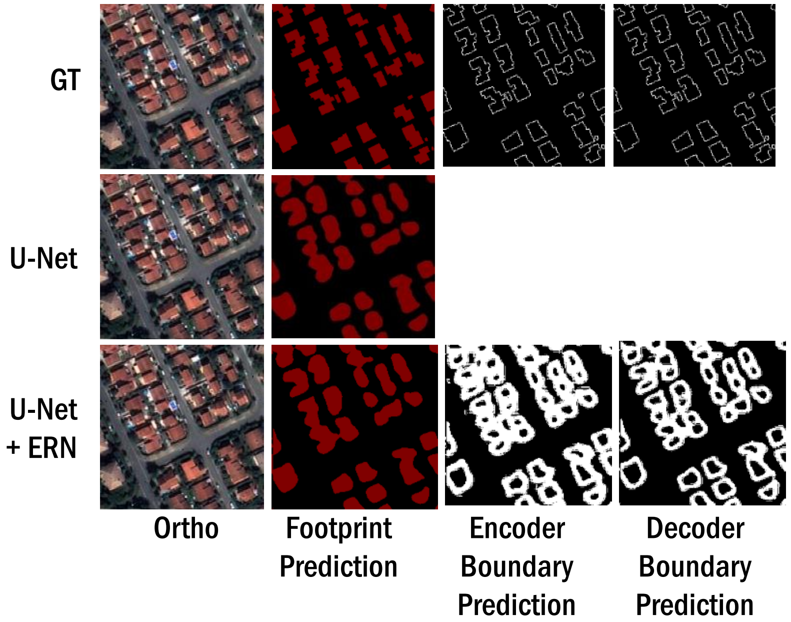


Figure 7: Difference of prediction masks between the U-Net RGBNir and the U-Net RGBNir+nDSM on Montpellier (red: removed by the nDSM, green: added by the nDSM). From left to right: TO image, nDSM, TO image overlaid with the prediction difference. (a) A square with roof-like radiometry (b) A building with a dark and complex roof (c) An elevated road.

Network architecture

Guided contour refinement



More suitable encoder-decoder association

Encoder \ Decoder	U-Net [7]	RefineNet [42]	DeepLabV3+ [43]
U-Net [7]	X		
ResNet-101 [40]		X	
ResNet-152 [40]	X		X
Xception [41]			X
EfficienNet [8]		X	

Building footprint segmentation: Four axes of study

Loss function

Cross entropy loss

$$L_{CE} = \sum_{i=0}^{n} -w_i \log \left(\frac{\exp(y_{pred_i})}{\sum_{j=0}^{n} \exp(y_{pred_j})} \right)$$

Dice loss

$$L_{dice} = 1 - \frac{2 \sum_{(x,y)} y_{true}(x,y) y_{pred}(x,y)}{\sum_{(x,y)} y_{true}^2(x,y) + \sum_{(x,y)} y_{pred}^2(x,y)}$$

Weighted sum

$$L_{hybrid} = 0.5 \cdot L_{CE} + 0.5 \cdot L_{dice}$$

Model	Acc	mIoU	P	R	F1	IoU
U-Net (cross-entropy)						
RGBNir	0.930	0.845	0.924	0.836	0.878	0.782
RGBNir+nDSM+NDVI+c3	0.926	0.838	0.924	0.826	0.872	0.773
U-Net (cross-entropy + dice)						
RGBNir	0.932	0.847	0.913	0.848	0.879	0.785
RGBNir+nDSM+NDVI+c3	**0.937**	**0.858**	0.915	**0.863**	**0.888**	**0.799**

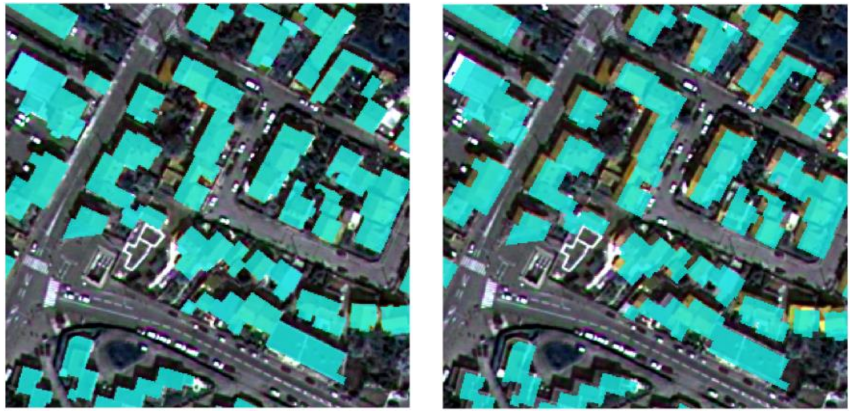
Table 2: Building segmentation networks evaluation. Accuracy (Acc) and mean IoU (mIoU) are computed for all classes. Precision (P), Recall (R), F1-score (F1) and IoU are computed for the building class.

Building footprint segmentation: Four axes of study

Database quality

Preliminary study on the impact of shifted groundtruth footprints on the prediction boundary quality

- Database: SemCity (WorldView-2, 50cm, on Toulouse)
- 2 training procedures:
  → Images WV2 + GT SemCity
  → Images WV2 + GT SC shifted per patch (2px ± 2px)



(a) Original GT (a) Shifted GT

SemCity Patch example

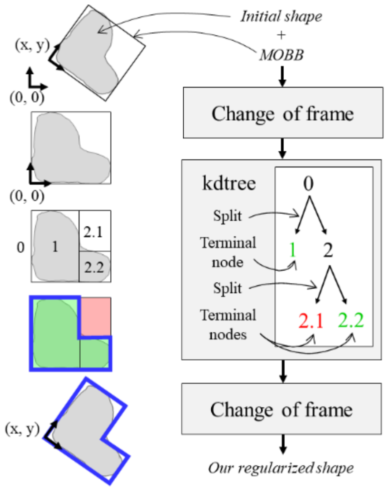
(a) Image SC (b) Original GT (c) Shifted GT

Predictions with Images WV2

Building vectorization and regularisation: Recover sharp faces

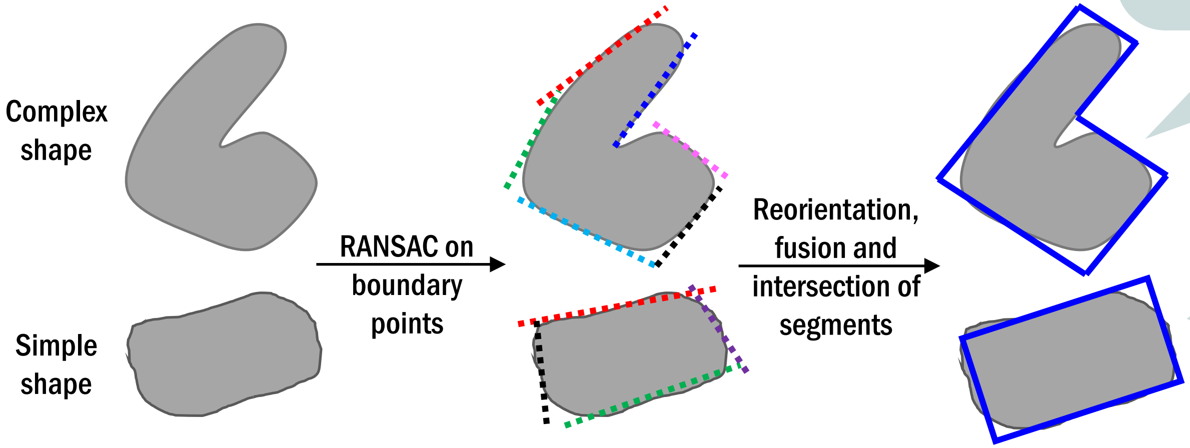
Two developed methods

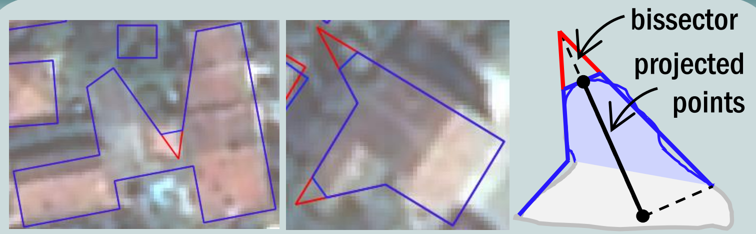
→ KDTree



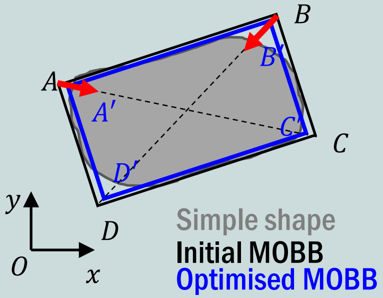
→ Photogrammetric Building Boundary (PBB)

based on Geodan's work: github.com/Geodan/building-boundary





« Spiked » angles regularisation

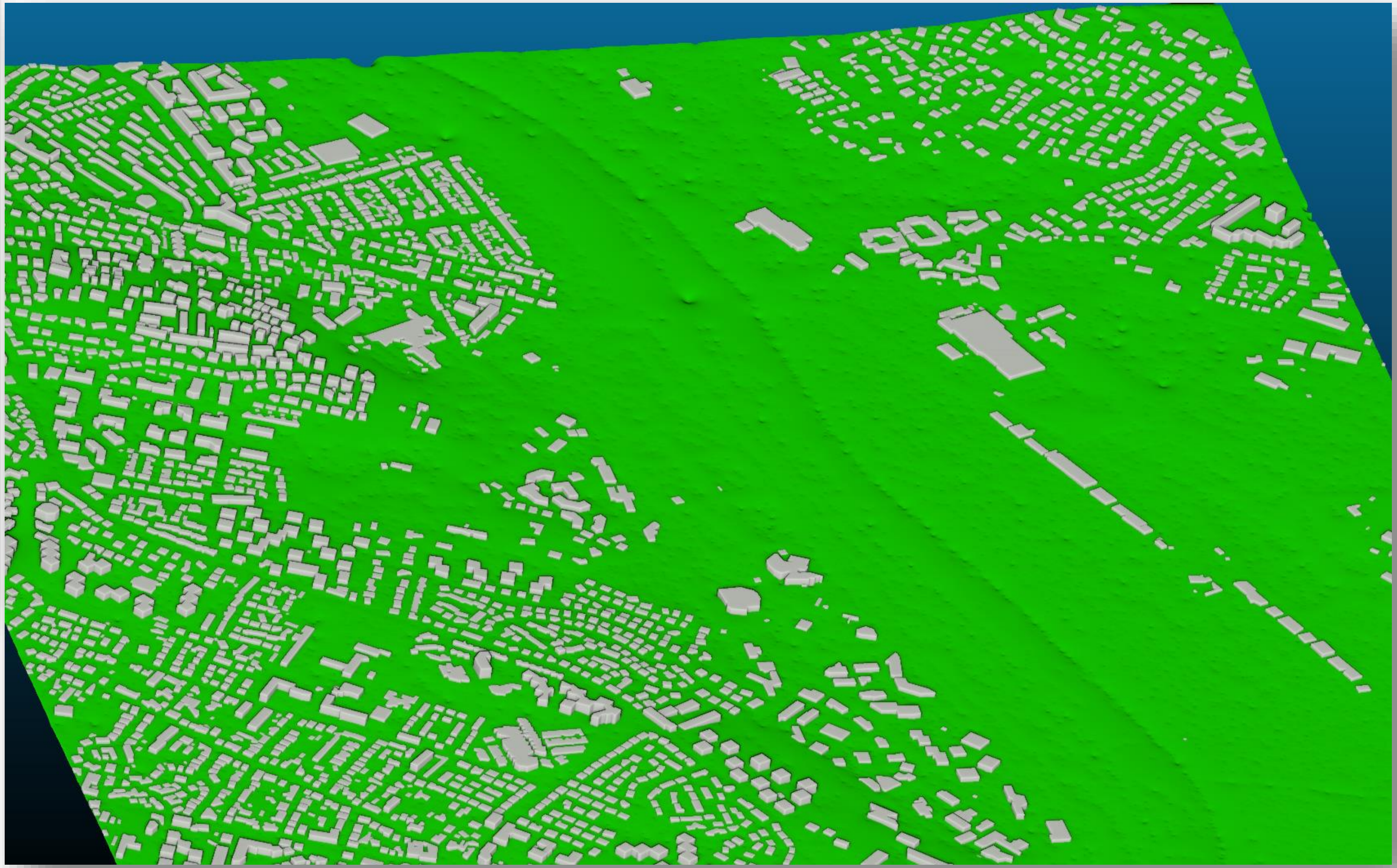


Optimisation of the Minimum Oriented Bounding Box (MOBB)

LPS - MAY - 2022

Building faces extrusion + DTM meshing

True Ortho

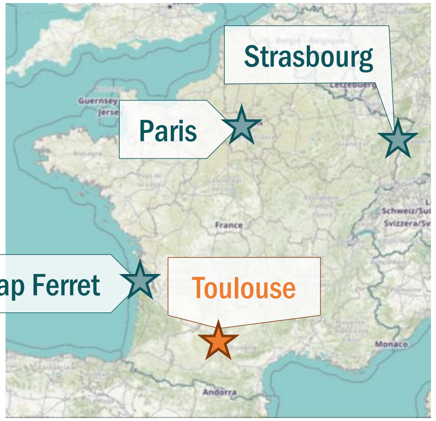


LoD1 building reconstruction

Toulouse, France

Evaluation: Generalisation ability of the pipeline

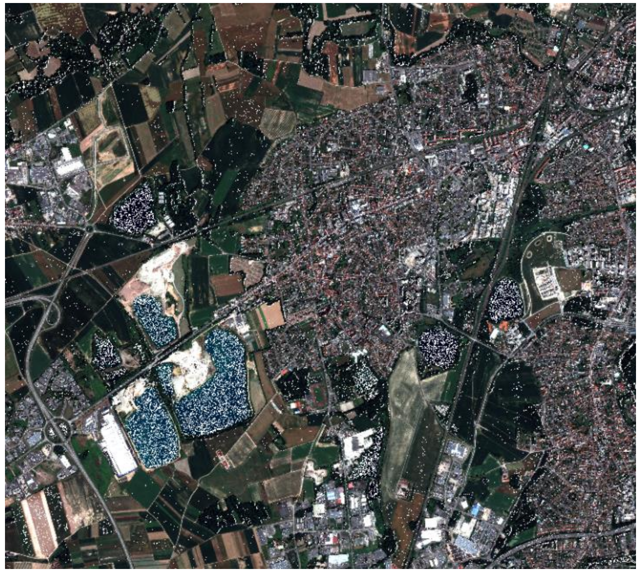
★ Test set
★ Training set

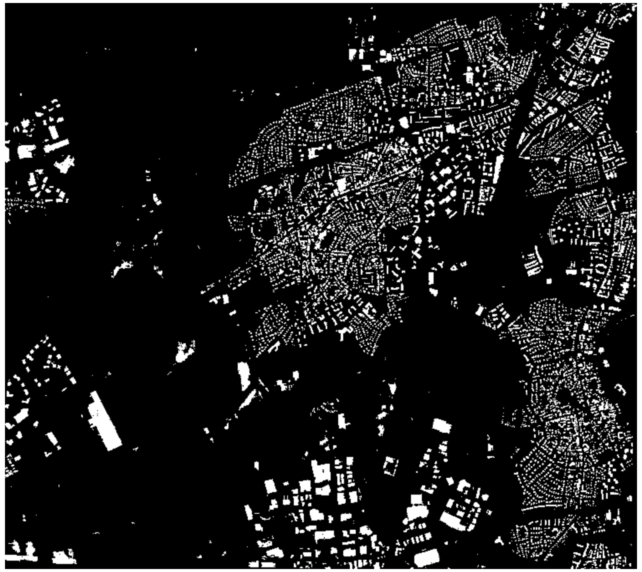


Metric	Value
Accuracy	0,88
mIoU	0,76
F1Score	0,84
Precision	0,91
Recall	0,77
IoU	0,72

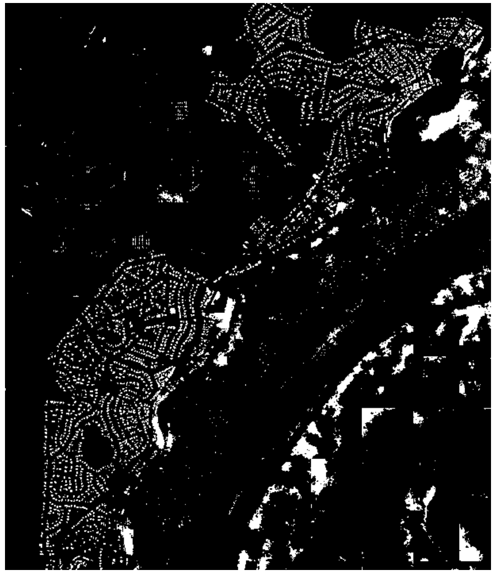
Paris

Strasbourg





Cap Ferret



Evaluation: Paris vectorisation and regularisation

Conclusion and perspectives

Conclusion

Implementation of a 3D building reconstruction pipeline from Pleiades imagery towards LoD1
Investigation of several aspects impacting the sharp restitution of contours

Perspectives

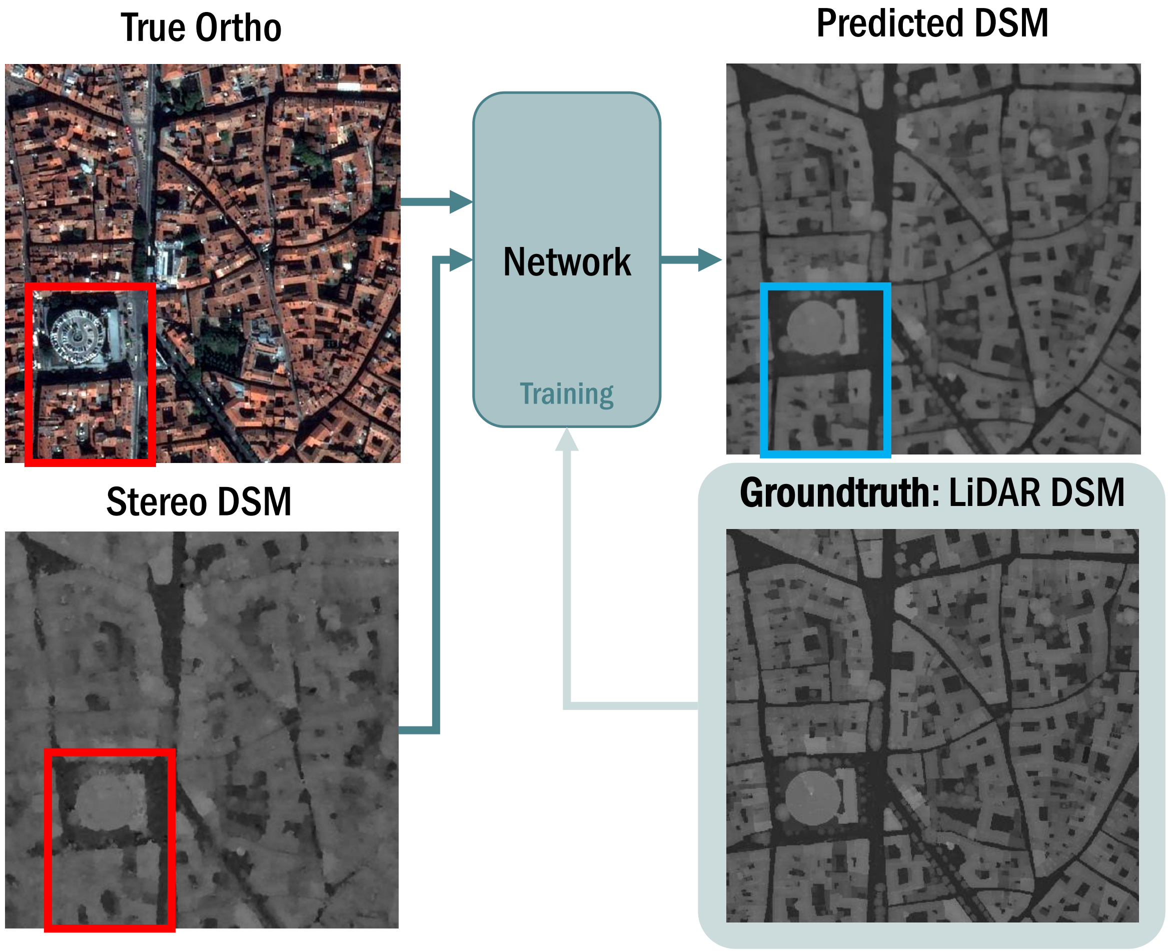
Work reused by CNES in the AI4GEO project

Allowed to identify the main limitations to work on

- Errors are additive along the pipeline
- DSM lack of quality

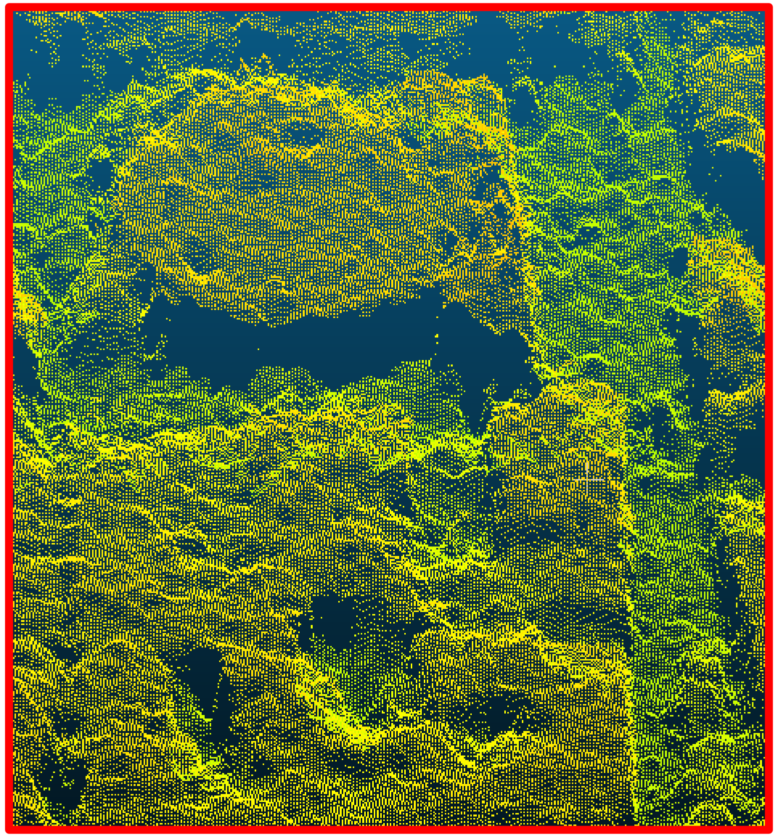
Current work on DSM improvement by deep learning, inspired by Stucker et al. "ResDepth"

Stucker, C., & Schindler, K. (2022). ResDepth: A deep residual prior for 3D reconstruction from high-resolution satellite images. ISPRS Journal of Photogrammetry and Remote Sensing.



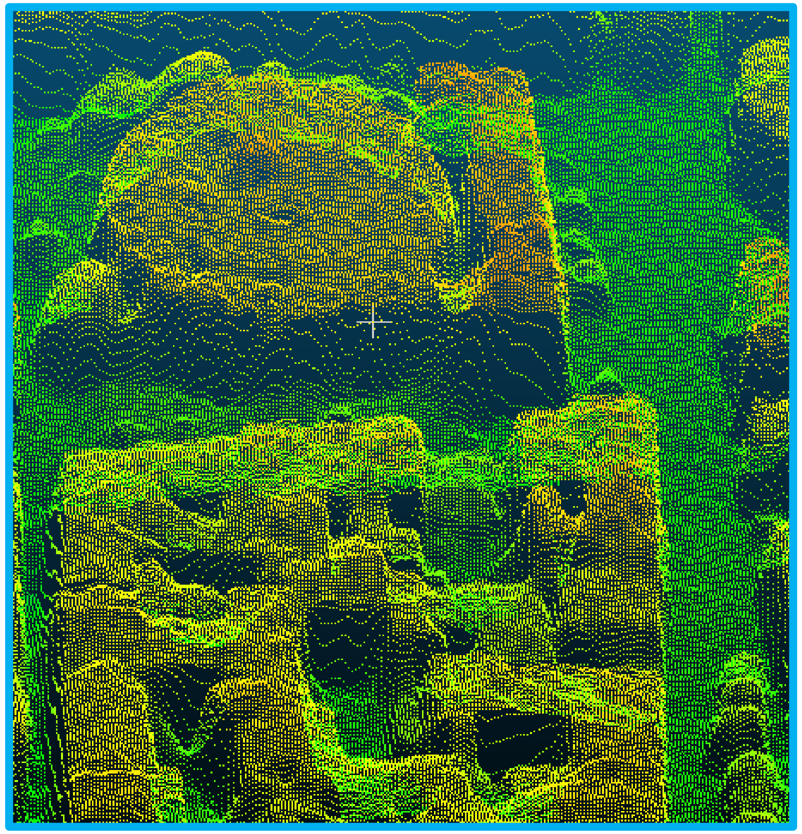
Preliminary results on DSM improvement

Conclusion and perspectives



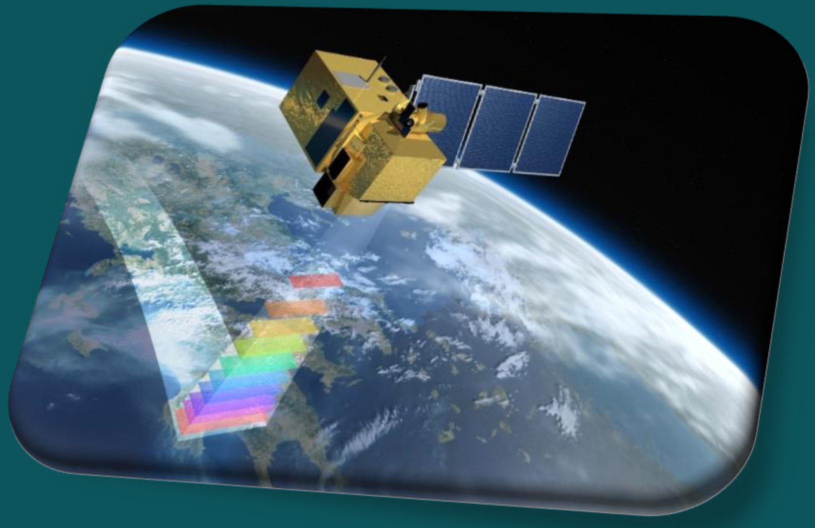
Stereo DSM 3D View

Google Maps 3D



Predicted DSM 3D View

Thank you for your attention.



earthobservation.magellium.com

eo@magellium.fr

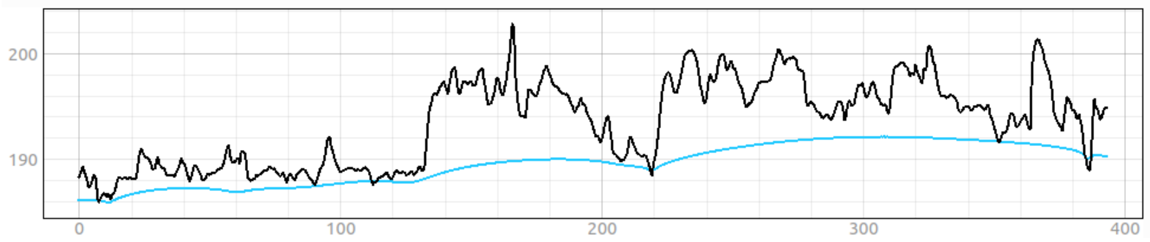
Appendices

Preprocessing on DSM

DTM estimation

Leotta et al. DrapCloth image-based implementation

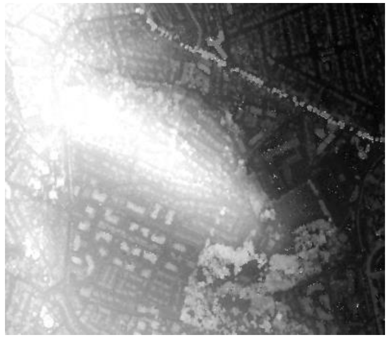
nDSM (normalized DSM) = DSM - DTM



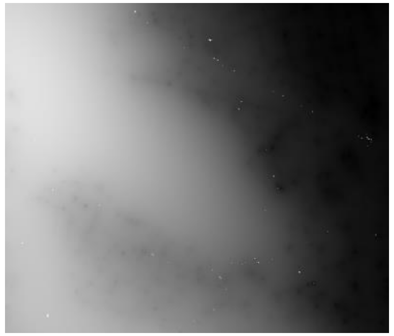
Altimetric profile along the red line (see below)

+ Good approximation, Fast

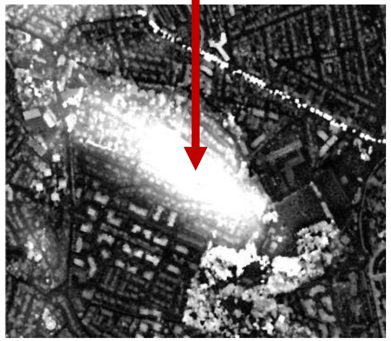
- Hill clipping effect



DSM

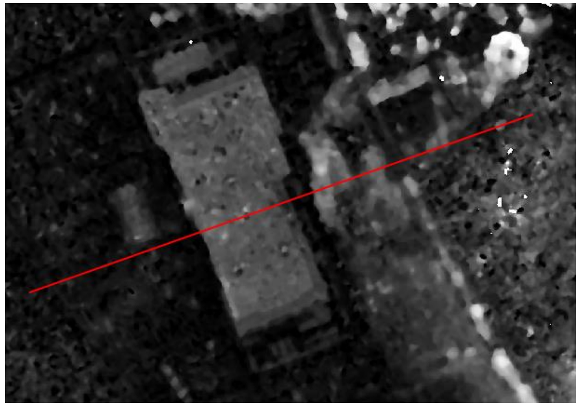


DTM (Median prefilter k=7)



nDSM

True Ortho

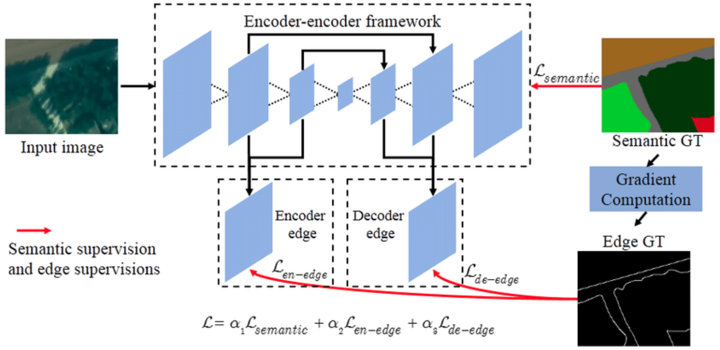


nDSM

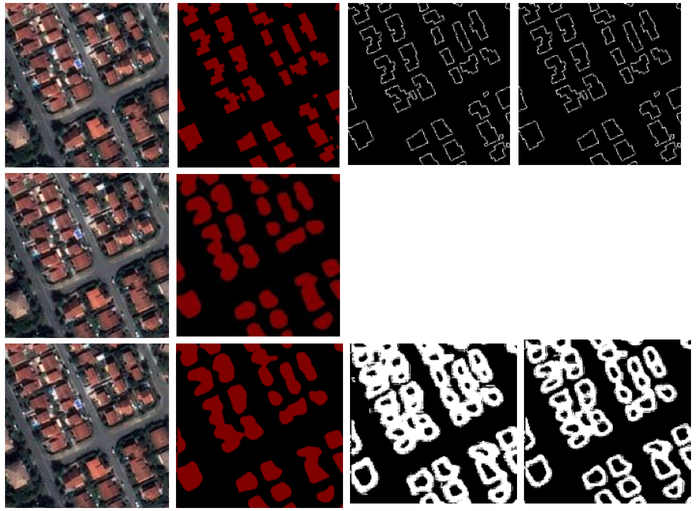
Building footprint segmentation: Four axes of study

Network architecture

Guided contour refinement



Liu, Shuo et al. "ERN: Edge Loss Reinforced Semantic Segmentation Network for Remote Sensing Images." Remote. Sens. 10 (2018): 1339.



More suitable encoder-decoder association

Encoder \ Decoder	U-Net [7]	RefineNet [42]	DeepLabV3+ [43]
U-Net [7]	X		
ResNet-101 [40]		X	
ResNet-152 [40]	X		X
Xception [41]			X
EfficienNet [8]		X	

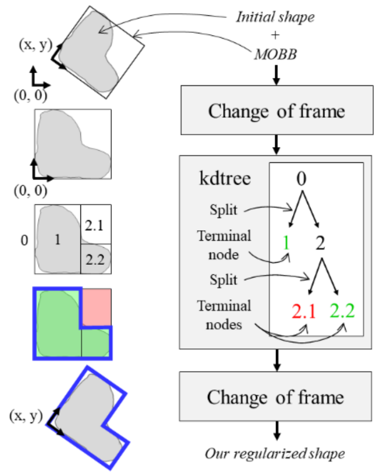
Model	Acc	mIoU	P	R	F1	IoU
U-Net (cross-entropy)						
RGBNir	0.930	0.845	0.924	0.836	0.878	0.782
RGBNir+nDSM	0.927	0.838	0.910	0.837	0.872	0.773
RGBNir+nDSM+NDVI	0.929	0.842	**0.931**	0.828	0.876	0.780
RGBNir+nDSM+NDVI+c3	0.926	0.838	0.924	0.826	0.872	0.773
ResNet-101 – RefineNet	0.926	0.84				
ResNet-152 – U-Net	0.888	0.76				
ResNet-152 – DeepLabV3+	0.912	0.81				
Xception – DeepLabV3+	0.882	0.75				
EfficientNet – RefineNet						
RGBNir+nDSM+NDVI+c3	0.931	0.846	0.921	0.841	0.879	0.784

Table 2: Building segmentation networks evaluation. Accuracy (Acc) and mean IoU (mIoU) are computed for all classes. Precision (P), Recall (R), F1-score (F1) and IoU are computed for the building class.

Building vectorization and regularisation: Recover sharp faces

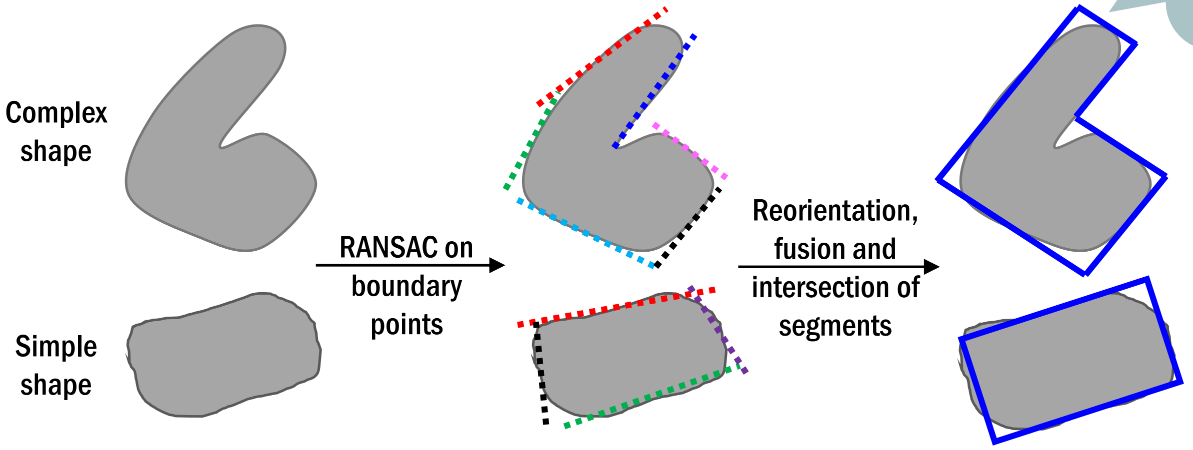
Two developed methods

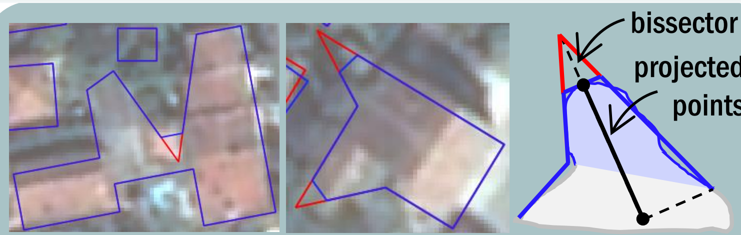
→ KDTree



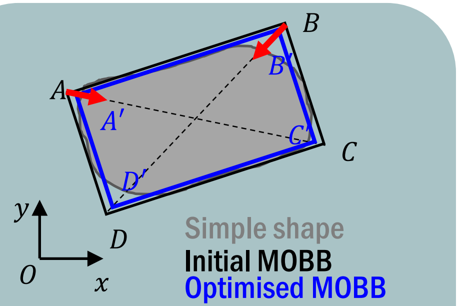
→ Photogrammetric Building Boundary (PBB)

based on Geodan's work: github.com/Geodan/building-boundary





« Spiked » angles regularisation



Optimisation of the Minimum Oriented Bounding Box (MOBB)

Douglas Peucker *KDTree* *PBB*

Loss	Method	IoU	Complexity	Orientation (in °)
c-e	DP (tol=2)	0.65 ± 0.10	0.87 ± 0.36	6.1 ± 6.9
	Kdtree	0.64 ± 0.11	2.02 ± 1.06	5.5 ± 6.3
	PBB	0.63 ± 0.12	0.72 ± 0.36	5.5 ± 6.4
c-e + dice	DP (tol=2)	0.64 ± 0.10	0.84 ± 0.39	5.5 ± 5.9
	Kdtree	0.63 ± 0.10	1.96 ± 1.08	4.5 ± 6.0
	PBB	0.63 ± 0.10	0.72 ± 0.35	5.0 ± 7.0

Table 3: Evaluation of building boundary regularization algorithms on a $1km^2$ Toulouse tile (~ 500 buildings): Douglas-Peucker (DP), Kdtree, Photogrammetric Building Boundary (PBB). Footprints are predicted by U-Net RGB-Nir+nDSM+NDVI+c3 models trained respectively with the cross-entropy (c-e) loss and the hybrid loss (c-e + dice).

Building vectorization and regularisation: PBB algorithm