Rule-Based, Noisy Labels for Overhead Imagery Segmentation

Conrad M Albrecht[1,2*], Chenying Liu[1,2], Yi Wang[1,2], Levente Klein[3], and Xiaoxiang Zhu[1,2]

*e–mail: Conrad.Albrecht@dlr.de

1: *Remote Sensing Technology Institute, German Aerospace Center, Germany*

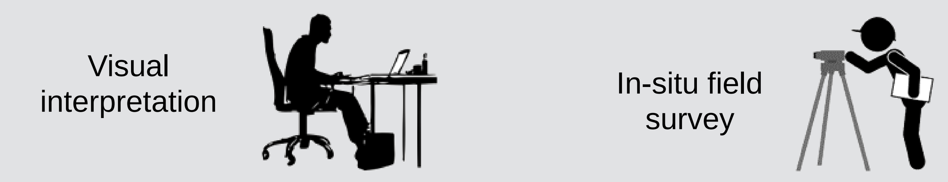
2: *Data Science in Earth Observation, Technical University of Munich, Germany*

3: *TJ Watson Research Center, IBM Research, Yorktown Heights, NY, USA*

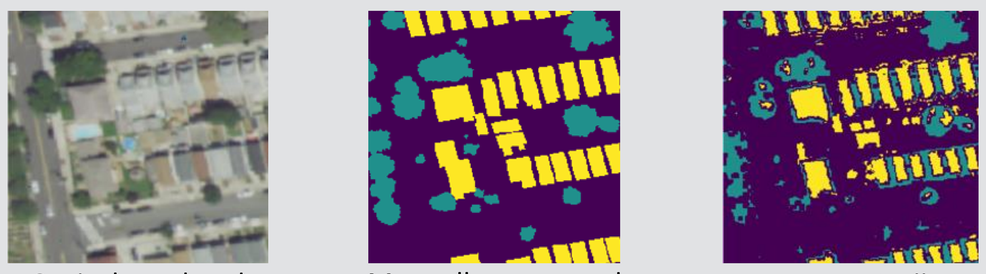
Introduction & Motivation

- With an increasing amount of satellites launched into space, **massive amounts** of overhead remote sensing (RS) imagery is **globally available**.

- In contrast, corresponding **large-scale, manually annotated RS data** is generally **hard to come by** and require significant financial resources.



- Development of **automatic labeling tools** is promising, yet challenging. A key aspect is **noise in automated labels**, i.e., inaccurate labels

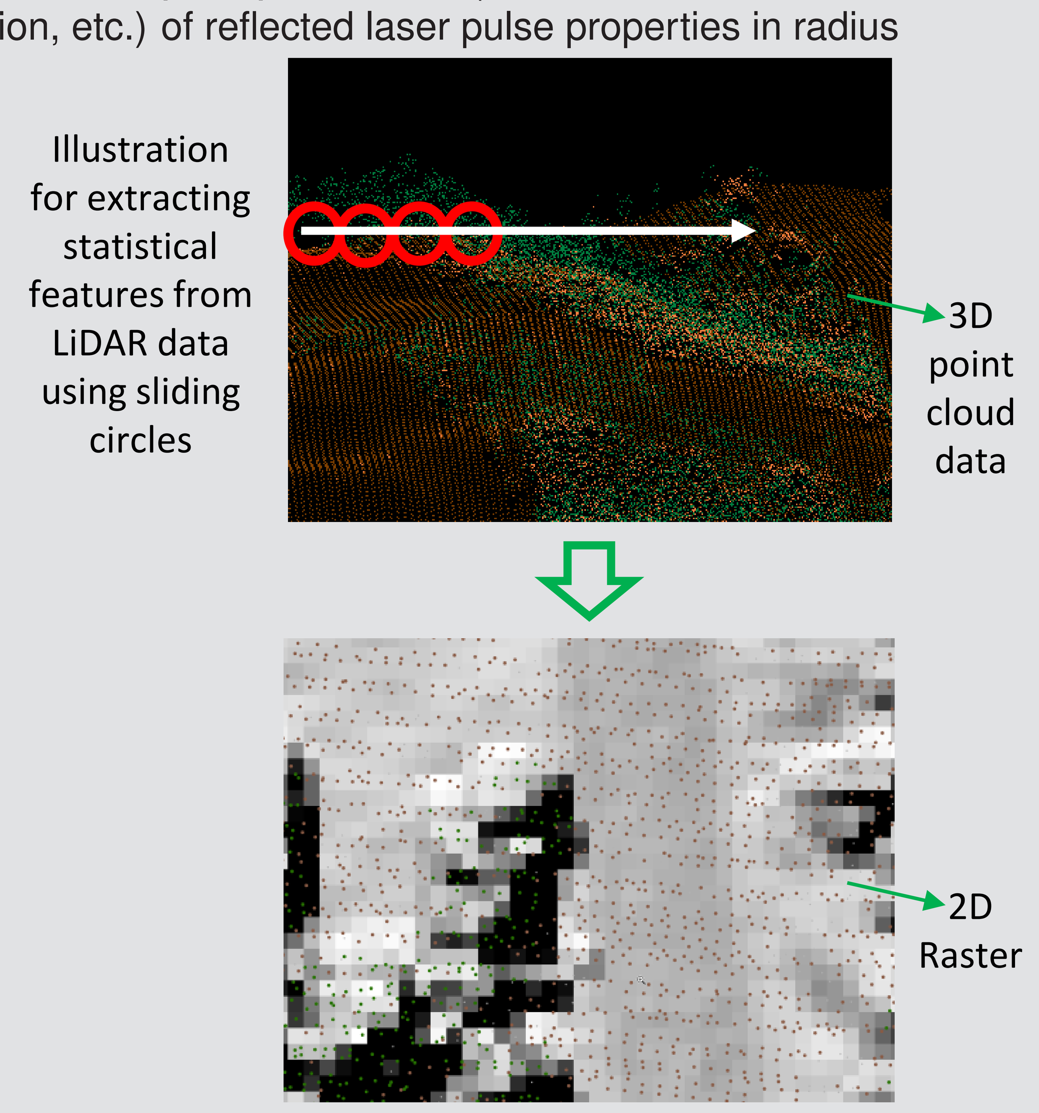


Optical overhead image — Manually annotated ground-truth labels — Automatically generated noisy labels

- In this work, we
  1. introduce a novel rule-based automatic label generation framework based on LiDAR data
  2. illustrate the effectiveness of the automatically generated noisy labels for semantic segmentation tasks

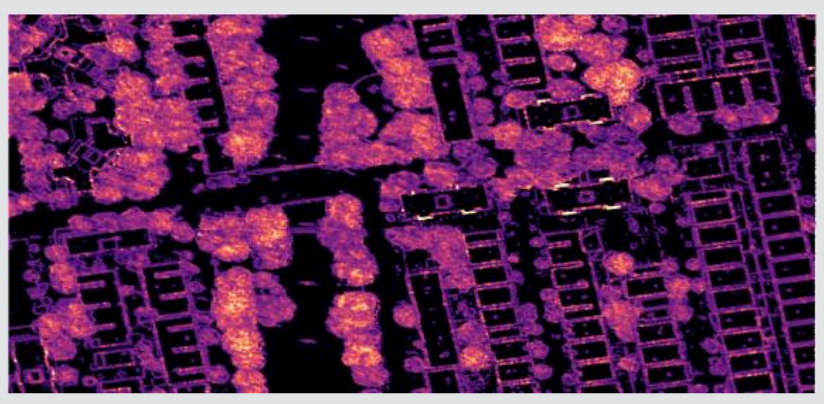
Automatic Label Generation

- LiDAR: **Li**ght **D**etection **A**nd **R**anging
  - send laser pulse to an object to measure the time-of-flight, the number of pulses, and light intensity of the light reflected
- statistical feature extraction: 3D → 2D
  - **point cloud rasterization**: slide a circle of radius 1.5 m across the geo-location of interest
  - calculate **(local) statistics** (min, max, mean, standard deviation, etc.) of reflected laser pulse properties in radius



- visualization of statistical features:

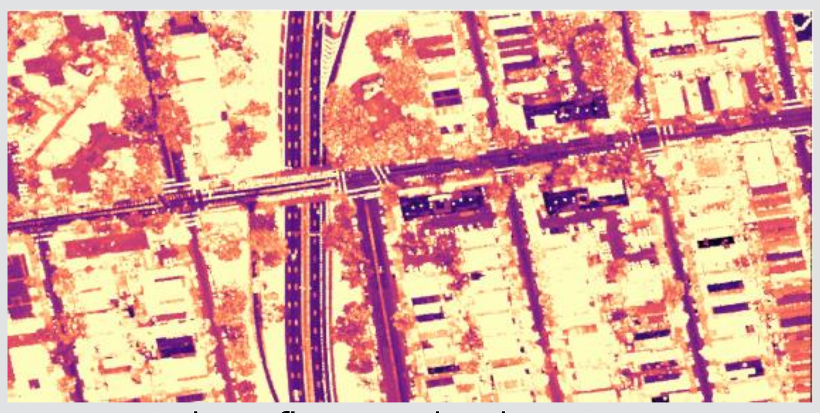
Optical image



Pulse count standard deviation (**trees** highlighted)

Automatic Label Generation

Elevation local mean (**buildings** highlighted)



Pulse reflectance local maximum (**roads** highlighted)

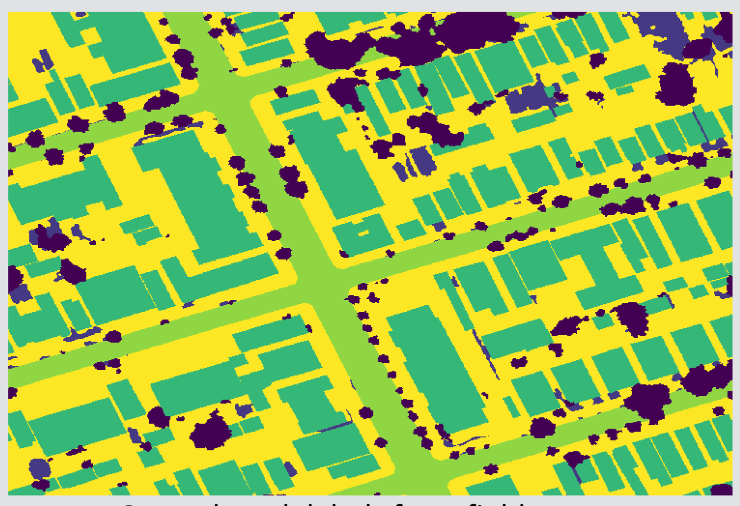
- Rule-based labeling process
  - Denotations:

	reflectance r	count c	elevation e
minimum	r_-	c_-	e_-
maximum	r_+	c_+	e_+
mean	$\bar{r}$	$\bar{c}$	$\bar{e}$
standard deviation	r_Δ	c_Δ	e_Δ

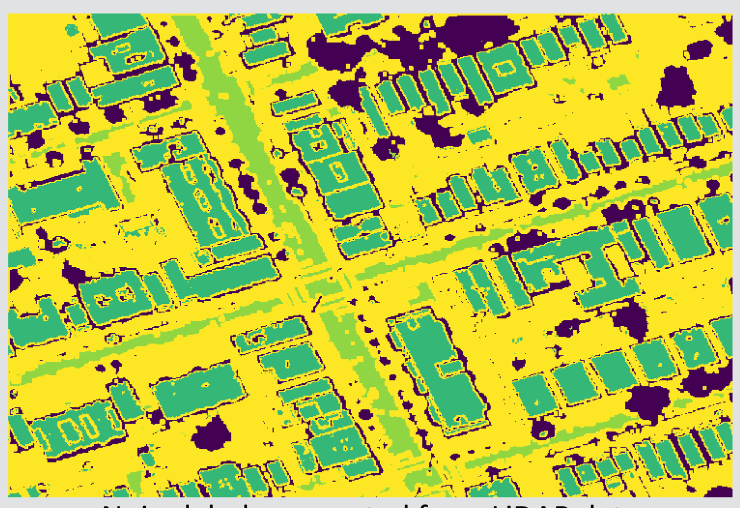
  - label generating rules:
    - **Trees**: the laser might get reflected multiple times ($c_\Delta \gg 0$) by branches and foliage at various elevation levels ($e_\Delta \gg 0$) as it penetrates a tree's canopy
    - **Buildings**: Rooftops reflect the laser pulse by a single return and show little variation in elevation ($e_+ \approx e_-$, $e_\Delta \approx 0$)
    - **Roads**: the minimum elevation is approx. 0 ($e_- \approx 0$) for roads. Besides, the black surface of asphalt absorbs a major portion of the laser pulse ($r_- \approx 0$).
  - Implementation:

class	pseudo (R,G,B)	binary classification formulas e
trees	$(c_+, e_\Delta, c_\Delta)$	$c_+ > \langle c_+ \rangle \wedge e_\Delta > \langle e_\Delta \rangle \wedge c_\Delta > \langle c_\Delta \rangle$
buildings	(e_-, e_Δ, e_+)	$e_- > \langle e_- \rangle \wedge e_\Delta < \langle e_\Delta \rangle \wedge e_+ > \langle e_+ \rangle$
roads	$(r_-, \bar{r}, e_-)$	$r_- > 0.1 \cdot r_+ \wedge \bar{r} < 0.6 \cdot r_+ \wedge e_- < 0.1 \cdot e_+$

  - contrasting rule-based noisy labels vs. ground truth:



Ground-truth labels from field survey



Noisy labels generated from LiDAR data
Building contours are well defined

Trees Buildings Roads Background

Dataset

- *spatial & temporal coverage*: New York City (NYC) in 2017
- *input data*: multispectral orthophotos obtained from National Agriculture Imaging Program (NAIP)
  - 4 bands: near-infrared (NIR), red (R), green (G), and blue (B)
  - spatial resolution: 1 meter
- *LiDAR data for noisy label generation*: ~ 10 points per square meter
- *Ground-truth labels for evaluation*: generated based on geospatial surveys NYC cadastral information

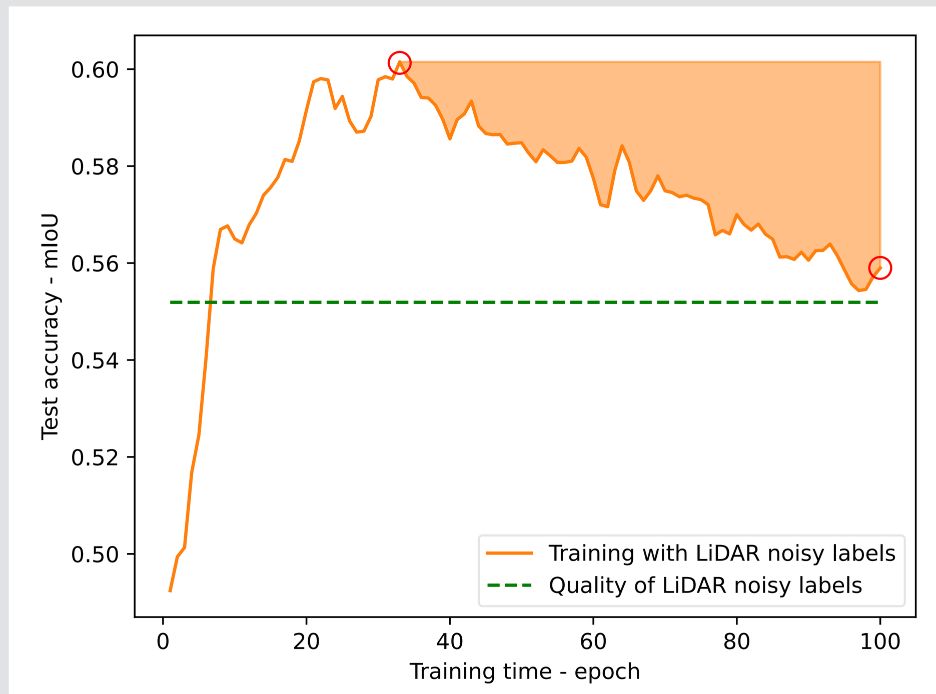
Segmentation with Noisy Labels

- Setting:
  - network architecture: **U-Net**
  - segmentation classes: **trees**, **buildings**, backgrounds
  - training & testing data: 6,650 patches of 256×256 pixels in total with 5,600 for training, and 1,050 patches for test
  - loss function: **Cross Entropy Loss + Dice loss**
  - optimization: **Adam optimizer** @ learning rate `1e-3`
- Quality of automatically generated noisy labels[a]:

class	trees	buildings	background	Overall
IoU	0.49(1)	0.46(2)	0.70(1)	0.55(1)[b]
precision	0.60(1)	0.66(1)	0.93(1)	0.77(1)[c]

[a]Uncertainty was estimated according to sets of 200 patches randomly picked
[b]mIoU: mean Intersection of Union
[c]OA: overall accuracy

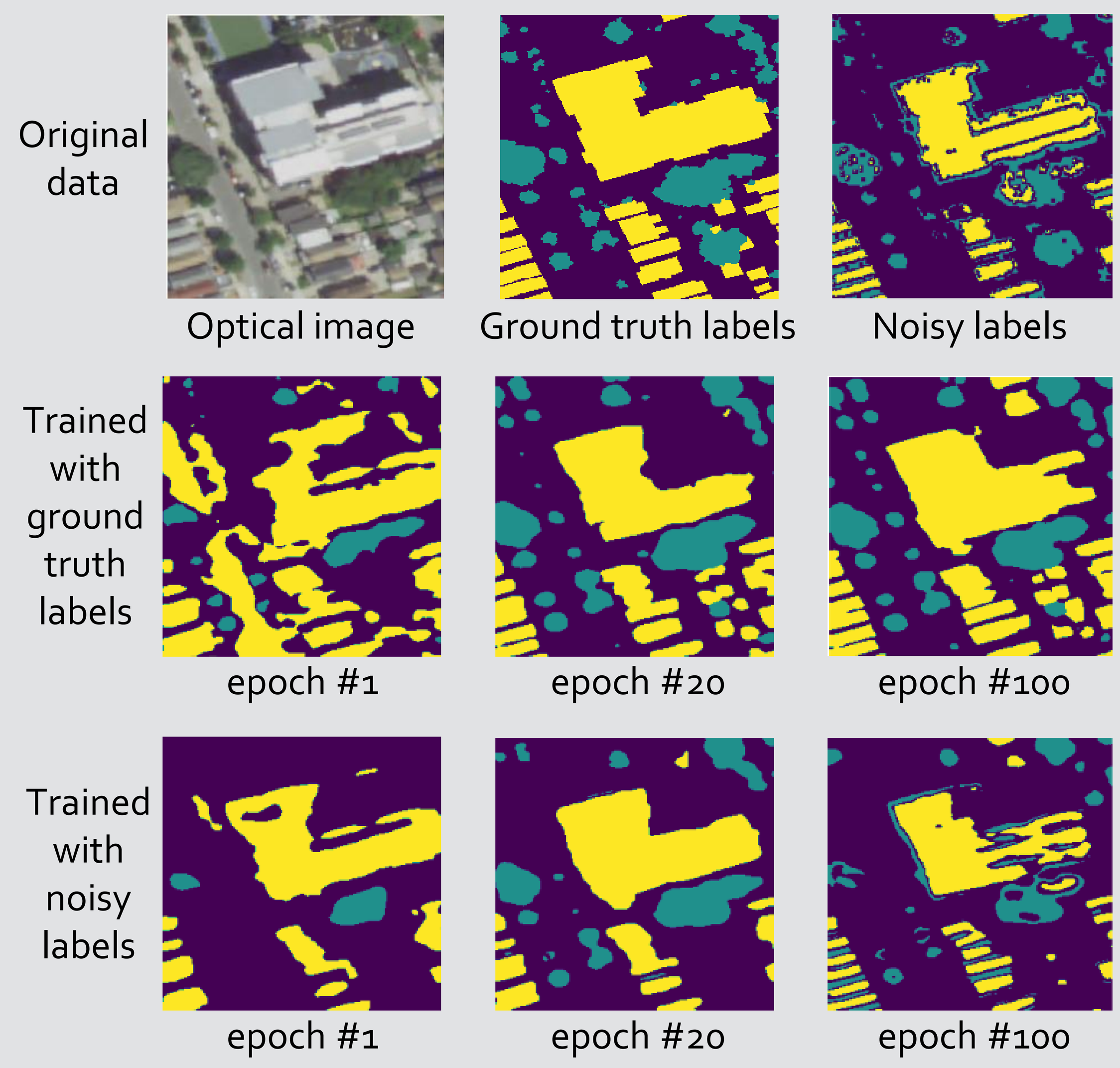
- Experimental results:



  - model accuracy in units of Intersection-over-Union (IoU) training on noisy LiDAR labels:

IoU	trees	buildings	background	mIoU
noisy labels	0.49(1)	0.46(2)	0.70(1)	0.55(1)
final	0.50(1)	0.47(1)	0.70(0)	0.56(1)
best	**0.55(1)**	**0.59(1)**	0.71(1)	**0.61(1)**
@epoch	@57	@29	@32	@32

  - Classification maps (samples):



Conclusion & Perspective

- The **rule-based**, automatic (noisy) **label generation** framework bears potential **for remote sensing** image processing **in** the context of **weak supvervision**.
- training **semantic segmentation with U-Net** allows to **reduce label noise** before overfitting
- *Perspective*: design deep learning strategies to handle label noise in Big Geo-Data

References

[1] "AutoGeoLabel: Automated Label Generation for Geospatial Machine Learning," 2021 IEEE International Conference on Big Data (Big Data), 2021, pp. 1779-1786.

[2] "Monitoring Urban Forests from Auto-generated Segmentation Maps," in press for IGARSS 2022.

conference: *2022 Living Planet Symposium*, May 23–27, 2022, World Conference Center Bonn, Bonn, Germany