Deep Learning based methods for remote sensing data
Doing more with buildings

Sylvain Lobry[1], Diego Marcos Gonzalez[1] , John E. Vargas-Muñoz[2] , Benjamin Kellenberger[1], Shivangi Srivastava[1], Devis Tuia[1]

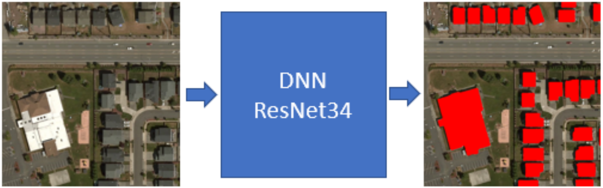
[1]Laboratory of Geo-information Science and Remote Sensing, Wageningen University & Research, The Netherlands
[2]Institute of computing, University of Campinas, Brazil

ESA Φ-week
13/11/2018

Building segmentation with deep learning

June 28th 2018: *Bing releases 125 million Building Footprints in the US as Open Data*

June 28th 2018: *Bing releases 125 million Building Footprints in the US as Open Data*
How?



Apply ResNet [He et al., 2015] + smart postprocessing

Building segmentation with deep learning

June 28th 2018: *Bing releases 125 million Building Footprints in the US as Open Data*

IGARSS 2018: *Large-scale semantic classification: outcome of the first year of Inria aerial image labeling benchmark* [Huang et al., 2018]

June 28th 2018: *Bing releases 125 million Building Footprints in the US as Open Data*

IGARSS 2018: *Large-scale semantic classification: outcome of the first year of Inria aerial image labeling benchmark* [Huang et al., 2018]
Winner:

Apply U-Net [Ronneberger et al., 2015] with a modified inference method

Question

Is it always sufficient to apply off the shelf methods?

Semantic segmentation vs Instance segmentation

Semantic segmentation vs Instance segmentation

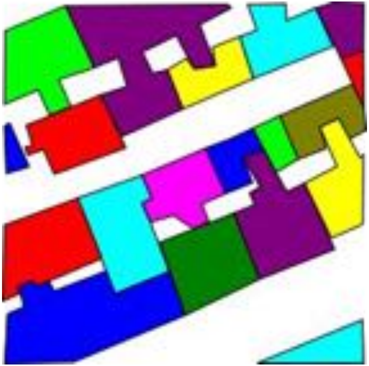
Semantic segmentation

- Many off the shelf algorithms
- No info about structure

Semantic segmentation vs Instance segmentation

Semantic instances segmentation



- Can encode geometry priors
- Can export GIS footprints
- No off the shelf algorithm

Segmenting buildings

Based on:

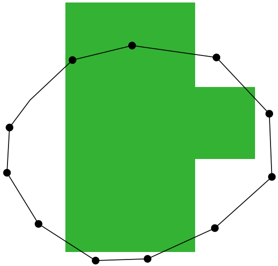
Learning deep structured active contours end-to-end
Diego Marcos, Devis Tuia, Benjamin Kellenberger, Lisa Zhang, Min Bai, Renjie Liao, Raquel Urtasun
in CVPR 2018

Semantic instances segmentation

In the 80's: people used Active Contour Models

Example of snakes [Kass et al., 1988]:

- A contour = set of points



Semantic instances segmentation

In the 80's: people used Active Contour Models

Example of snakes [Kass et al., 1988]:

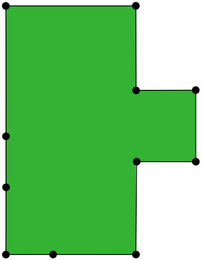
- A contour = set of points

Semantic instances segmentation

In the 80's: people used Active Contour Models

Example of snakes [Kass et al., 1988]:

- A contour = set of points
- Model enforce:
  - A data term (e.g. gradients)
  - Penalization of length
  - Penalization of curvature
  - Balloon term

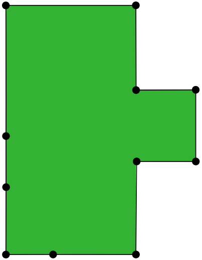


Semantic instances segmentation

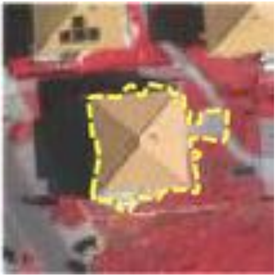
In the 80's: people used Active Contour Models

Example of snakes [Kass et al., 1988]:

- A contour = set of points
- Model enforce:
  - A data term (e.g. gradients)
  - Penalization of length
  - Penalization of curvature
  - Balloon term
- Each term is balanced



Tuning ACM parameters



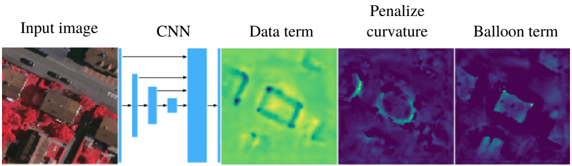
Should we penalize more length and curve?

Tuning ACM parameters

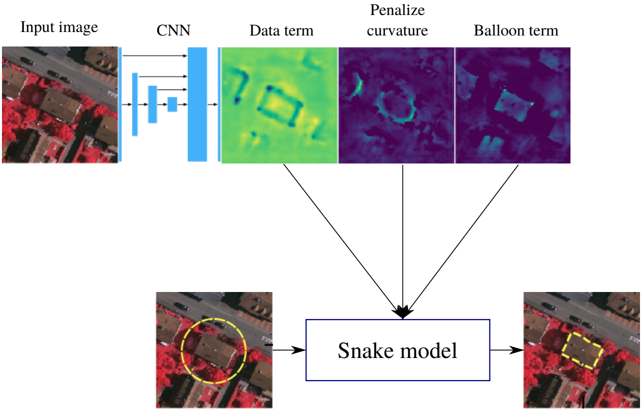
Tuning ACM parameters

What about the other buildings?

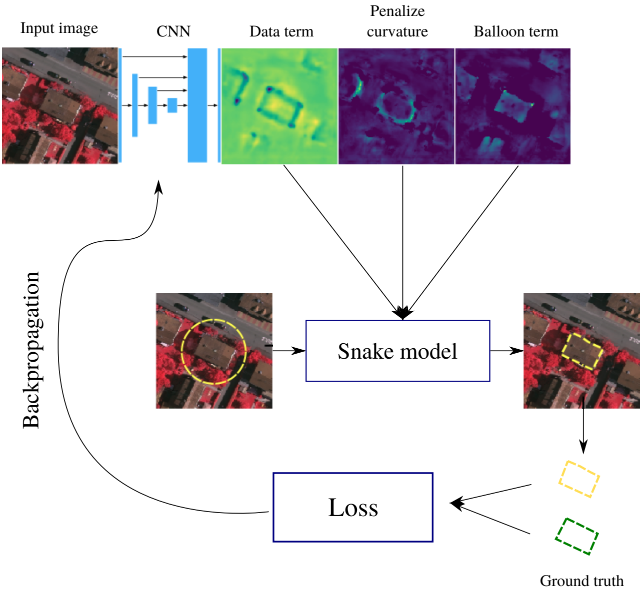
Tuning ACM parameters



Tuning ACM parameters



Tuning ACM parameters



Results and conclusion

- In a nutshell: learning the ACM parameters leading to desired convergence
- Comparison on the TorontoCity dataset [Wang et al., 2016] (over 12000 building instances):

Method	WeighCov	PolySim
ResNet [He et al., 2015]	0.40	0.29
Deep Watershed [Bai and Urtasun, 2016]	0.52	0.24
Proposed model	0.58	0.27

WeighCov: IoU-based weighted coverage
PolySim: shape similarity
(see [Wang et al., 2016])

Correcting building annotations

Based on:

Correcting rural building annotations in OpenStreetMap using convolutional neural networks

John Edgar Vargas Muñoz, Sylvain Lobry, Alexandre Xavier Falcão, Devis Tuia

in ISPRS Journal of Photogrammetry and Remote Sensing (in press)

Problem

Building annotations can be:

1. Misaligned (because imagery has changed)
2. Missing
3. There, but building has disappeared

Problem

Building annotations can be:

1. Misaligned (because imagery has changed)
2. Missing
3. There, but building has disappeared

Question

Can we correct these annotations instead of starting from scratch?

Problem

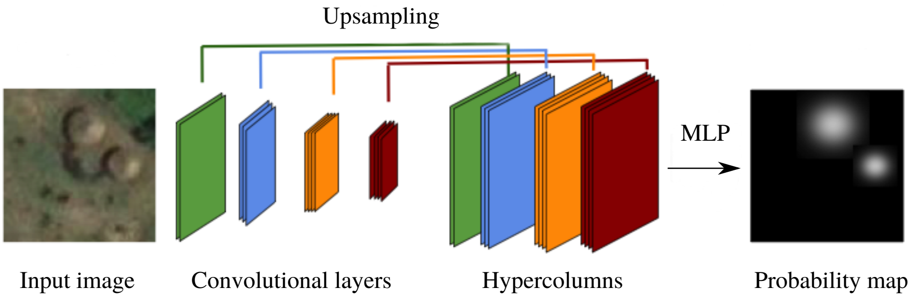
Building annotations can be:

1. Misaligned (because imagery has changed)
2. Missing
3. There, but building has disappeared

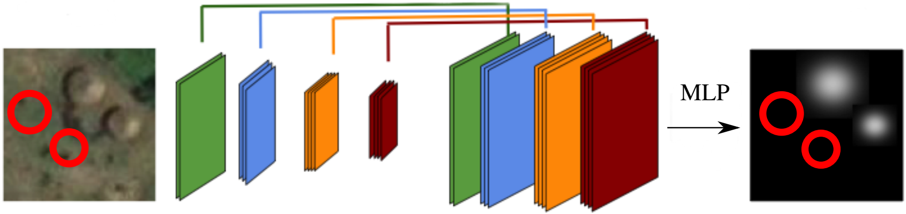
Question

Can we correct these annotations instead of starting from scratch?

Solution: Aligning

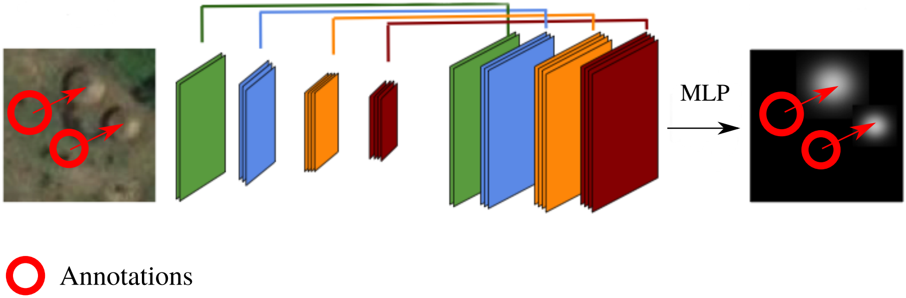


Solution: Aligning

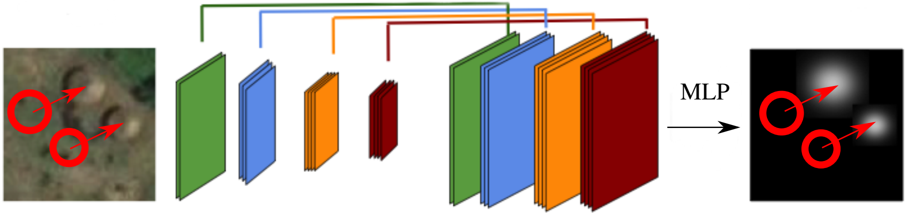


Annotations

Solution: Aligning



Solution: Aligning



Annotations

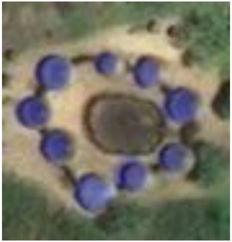
Use a Markov Random Field which:

- Maximize the correlation between annotations and probability map
- Enforce alignment vectors to be similar (in a group of buildings)

Results

Input

Semantic segmentation



Proposed method

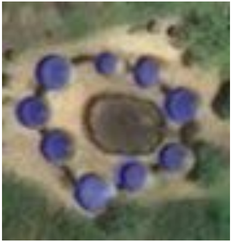
Method	F-score
Semantic segmentation [Maggiori et al., 2017]	0.657
Proposed method	0.725

F-score: harmonic mean of precision and recall
(higher is better)

Results

Input

Semantic segmentation



Proposed method

Conclusion

It is better to use (potentially inaccurate) OpenStreetMap data than starting from scratch

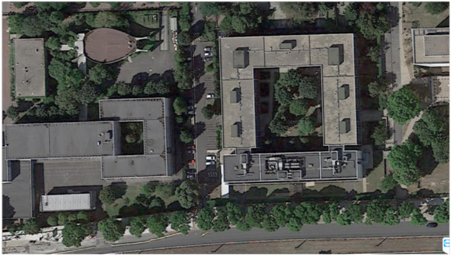
Land use classification

Based on:

Understanding urban landuse from above and ground perspectives: a deep learning, multimodal solution.

Shivangi Srivastava, John Edgar Vargas Muñoz, Devis Tuia

in Remote Sensing of Environment (under review)



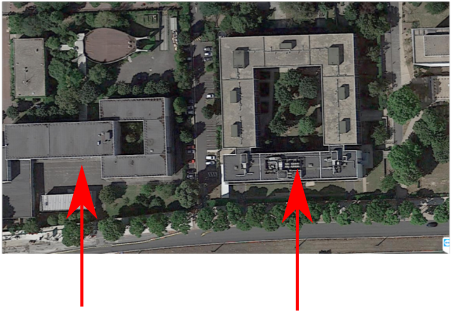
Land use classification

Based on:

Understanding urban landuse from above and ground perspectives: a deep learning, multimodal solution.

Shivangi Srivastava, John Edgar Vargas Muñoz, Devis Tuia

in Remote Sensing of Environment (under review)



Land use classification

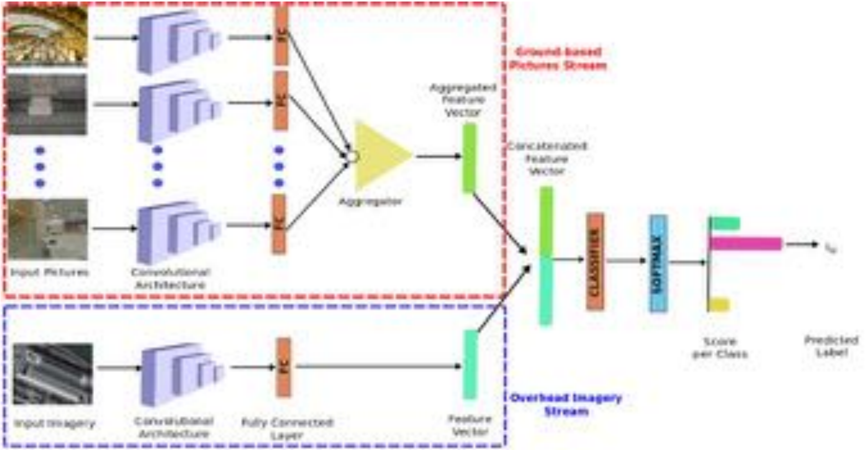
Problem

Using overhead imagery alone is not enough!

$\longrightarrow$ Use ground-based pictures (e.g. Google Street View)

a) Pictures located outside (GSV) b) Urban-Object Label (OSM) c) Pictures located inside (GSV)

Land use classification



Land use classification

Conclusion

Conclusion

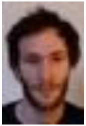
Applying off the shelf methods from computer vision to remote sensing data works but, as a community, we can do better

- We do not always have the same problems
- Using priors
- Using auxiliary data

Thank you!

The team:

Devis Tuia



Sylvain Lobry

Benjamin Kellenberger

Diego Marcos

John Vargas

Shivangi Srivastava

Supported by:

FNSNF
Fonds national suisse
Schweizerischer Nationalfonds
Fondo nazionale svizzero
Swiss National Science Foundation

http://www.sylvainlobry.com/phi-week-2018/
work@sylvainlobry.com

References

- Bai, M. and Urtasun, R. (2016).
  Deep watershed transform for instance segmentation.
  CoRR, abs/1611.08303.

- He, K., Zhang, X., Ren, S., and Sun, J. (2015).
  Deep residual learning for image recognition.
  CoRR, abs/1512.03385.

- Huang, B., Lu, K., Audebert, N., Khalel, A., Tarabalka, Y., Malof, J., Boulch, A., Le Saux, B., Collins, L., Bradbury, K., et al. (2018).
  Large-scale semantic classification: outcome of the first year of inria aerial image labeling benchmark.
  In *IEEE International Geoscience and Remote Sensing Symposium–IGARSS 2018*.

References (cont.)

- Kass, M., Witkin, A., and Terzopoulos, D. (1988).
  Snakes: Active contour models.
  International journal of computer vision, 1(4):321–331.

- Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P. (2017).
  Convolutional neural networks for large-scale remote-sensing image classification.
  IEEE Transactions on Geoscience and Remote Sensing, 55(2):645–657.

- Ronneberger, O., Fischer, P., and Brox, T. (2015).
  U-net: Convolutional networks for biomedical image segmentation.
  CoRR, abs/1505.04597.

- Wang, S., Bai, M., Máttyus, G., Chu, H., Luo, W., Yang, B., Liang, J., Cheverie, J., Fidler, S., and Urtasun, R. (2016).
  Torontocity: Seeing the world with a million eyes.
  CoRR, abs/1612.00423.