Semantic interpretation of optical remote sensing data by computer vision and machine learning

Michele Volpi

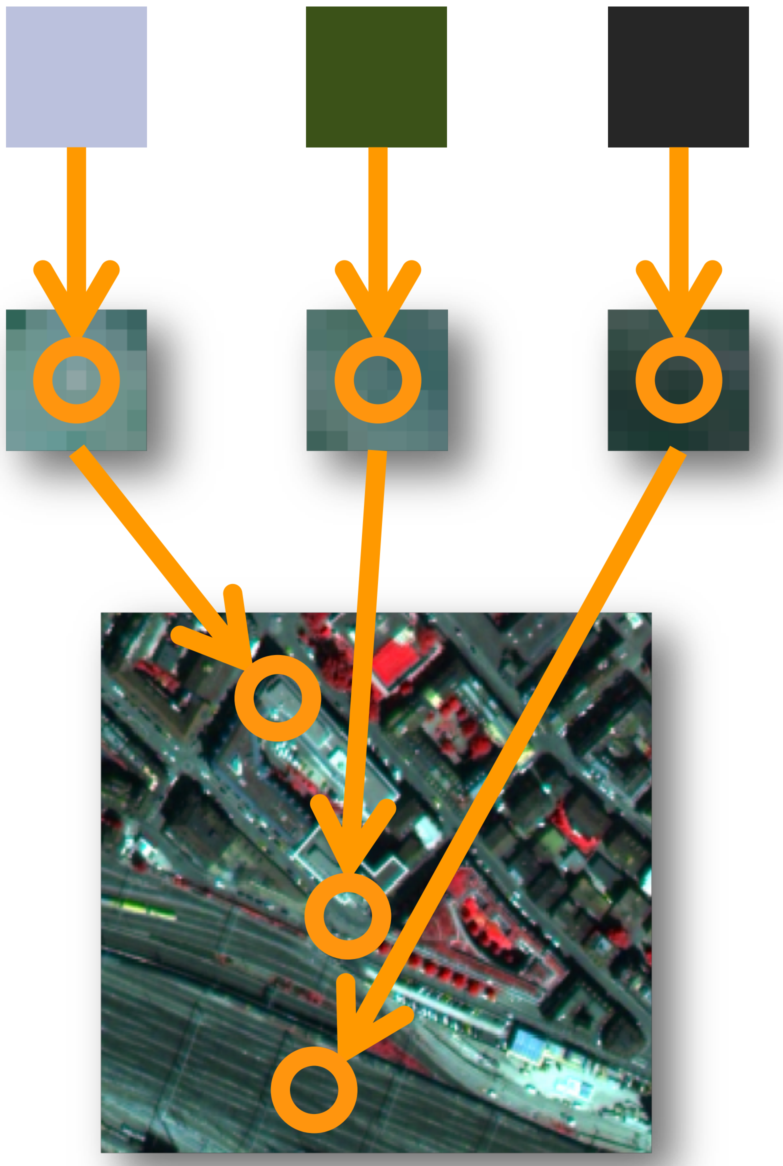
Swiss Data Science Center – ETHZ + EPFL

AMLD19, Lausanne

Semantic interpretation

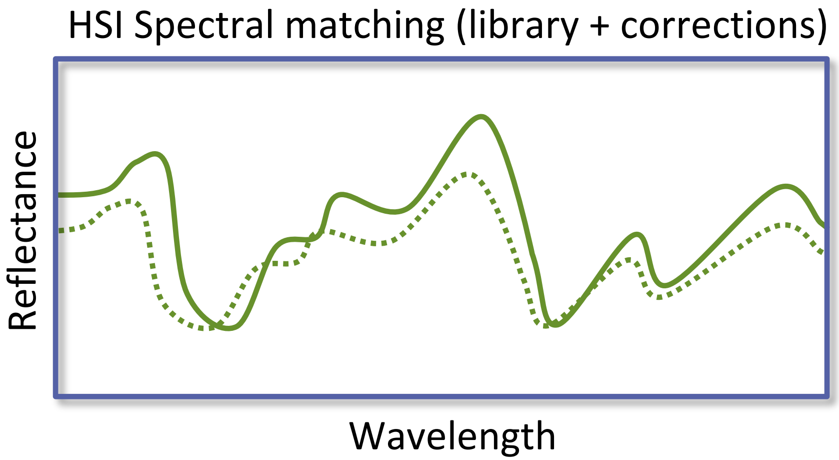
- Assign semantic classes to “objects” in the image

- Semantic concepts more abstract than radiometric classes (human vs physical concept)

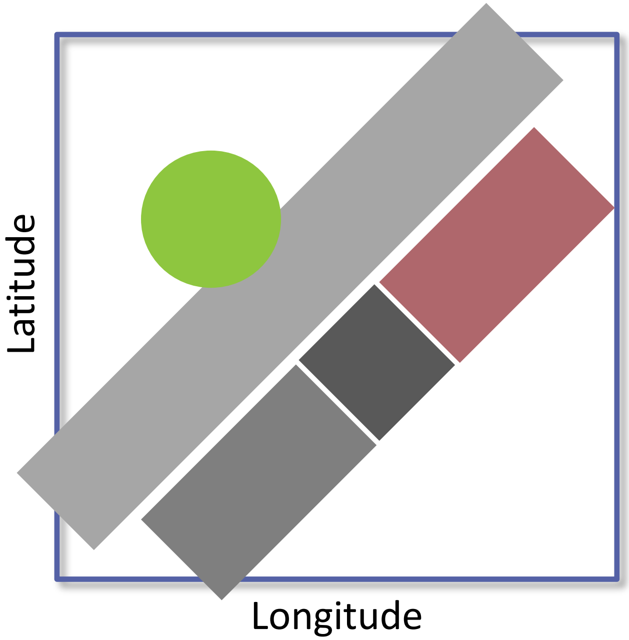


Semantic interpretation

- Why not just radiometric-based interpretation?
  - As resolution increases, smaller objects can be resolved
  - Semantics *might* not be directly derived from spectral signatures



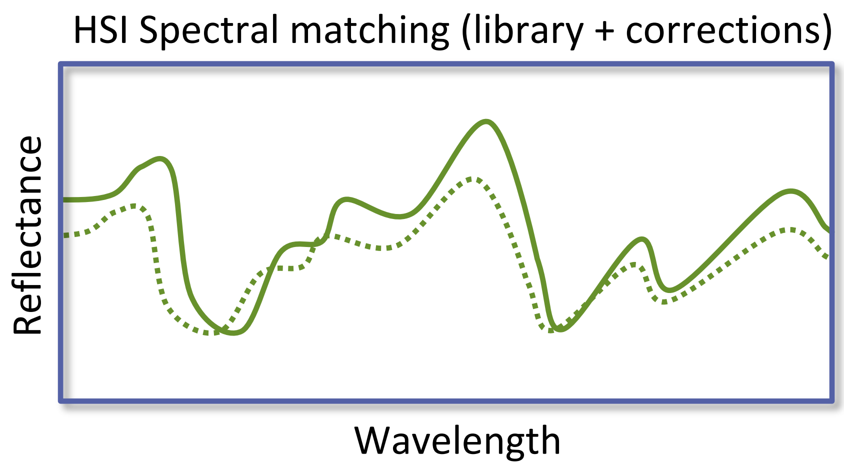
- **Semantic segmentation**: Assign a semantic label to every single mapping unit (pixel, window, segment, etc.)

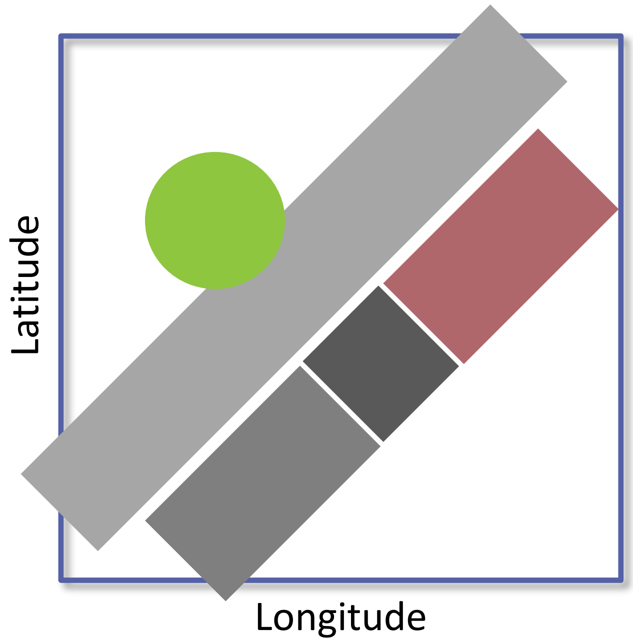


SDSC

Semantic segmentation

- Main differences:
  - Discriminative spatial context vs discriminative “colors”
  - Spatial ordering vs unordered (iid)
  - Different scales vs resolution
  - …
- How to achieve this?
  - (Learning) spatial features
  - Graphical models (MRF, CRF)
  - Inject domain specific information

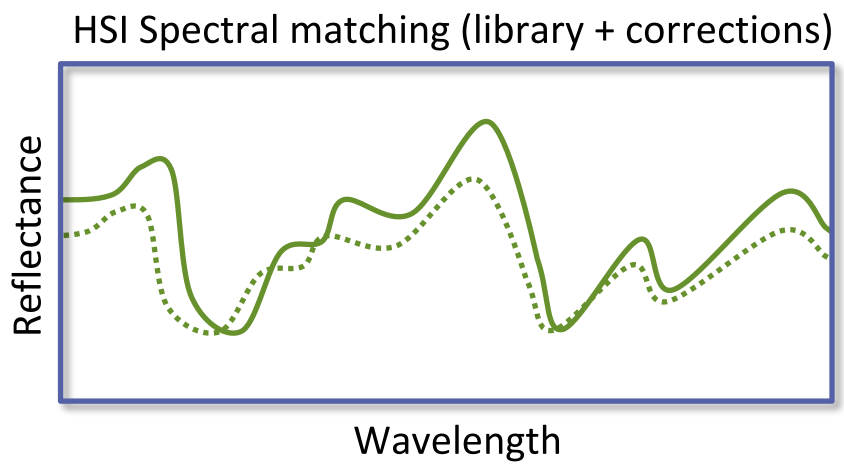


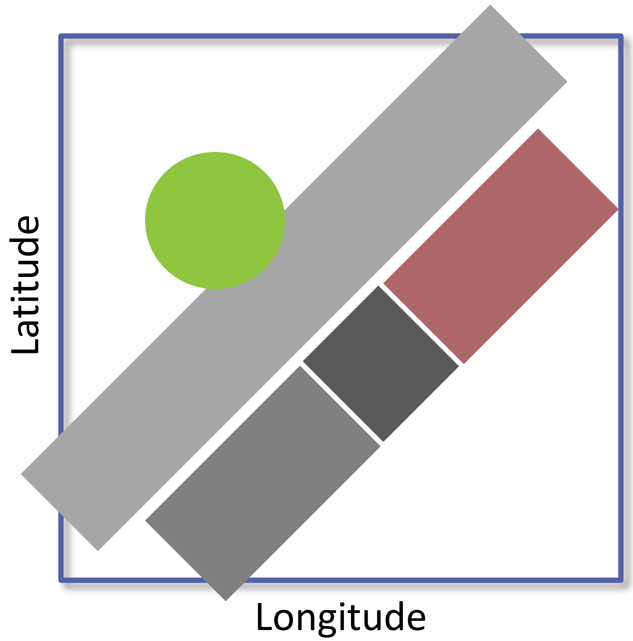


Semantic segmentation

- Main differences:
  - Discriminative **spatial context** vs discriminative “colors”
  - Spatial ordering vs unordered
  - Different scales vs resolution
  - …

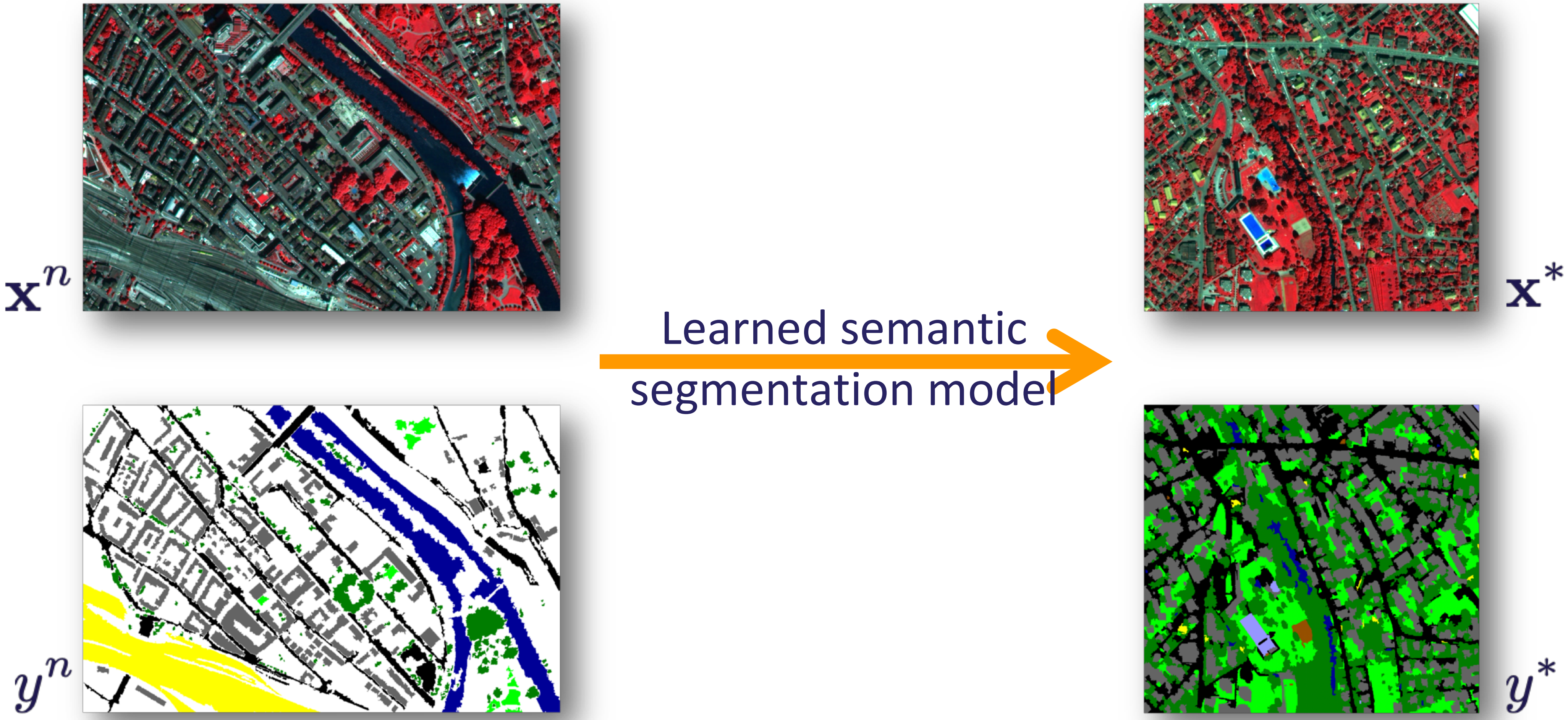
- How to achieve this?
  - (Learning) spatial features
  - Graphical models (MRF, CRF)
  - Inject domain specific information





SDSC

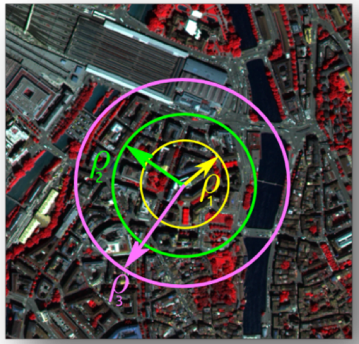
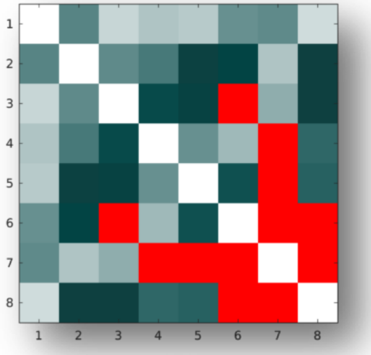
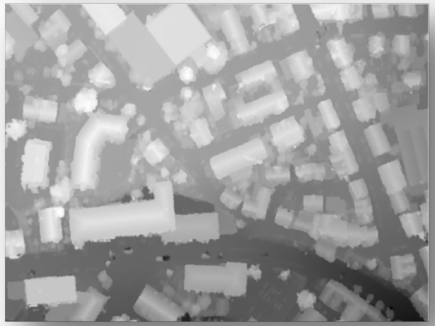
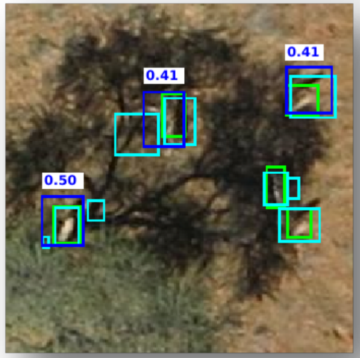
Learning the context

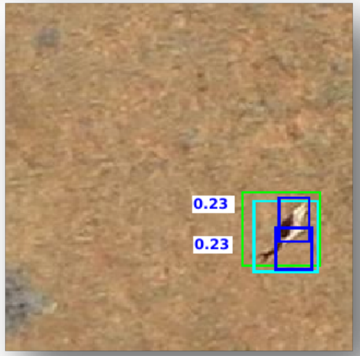


SDSC

Side information

- What can we use to inform models?
  - Additional priors related to the problem (e.g. spatial and geographic context)
  - Additional inputs (e.g. DSM, time series, image features, …)
  - Additional outputs (e.g. object detection, semantic boundaries, …)

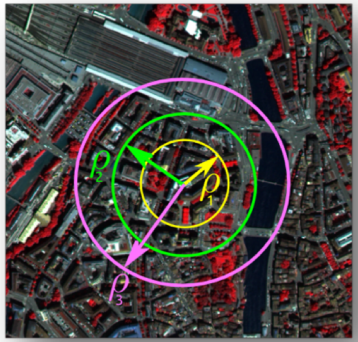





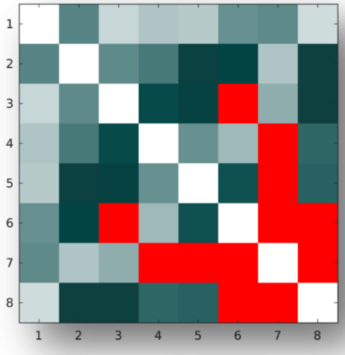


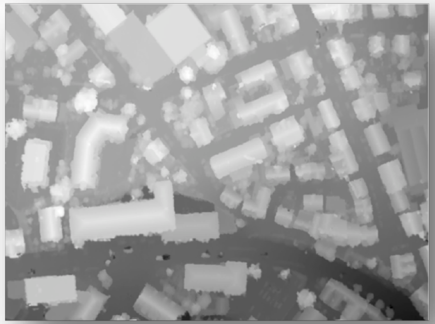
[Kellenberger et al., IEEE IGARSS 2017]

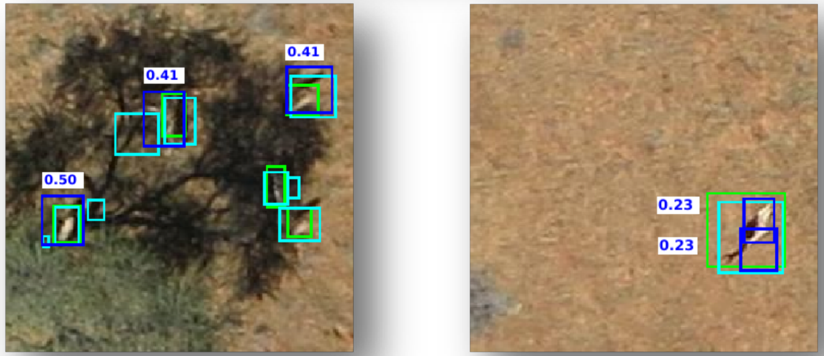
Side information

- What can we use to inform models?
  - **Additional priors** related to the problem (e.g. spatial and geographic context)
  - **Additional inputs** (e.g. DSM, time series, image features, …)
  - Additional outputs (e.g. object detection, semantic boundaries, …)









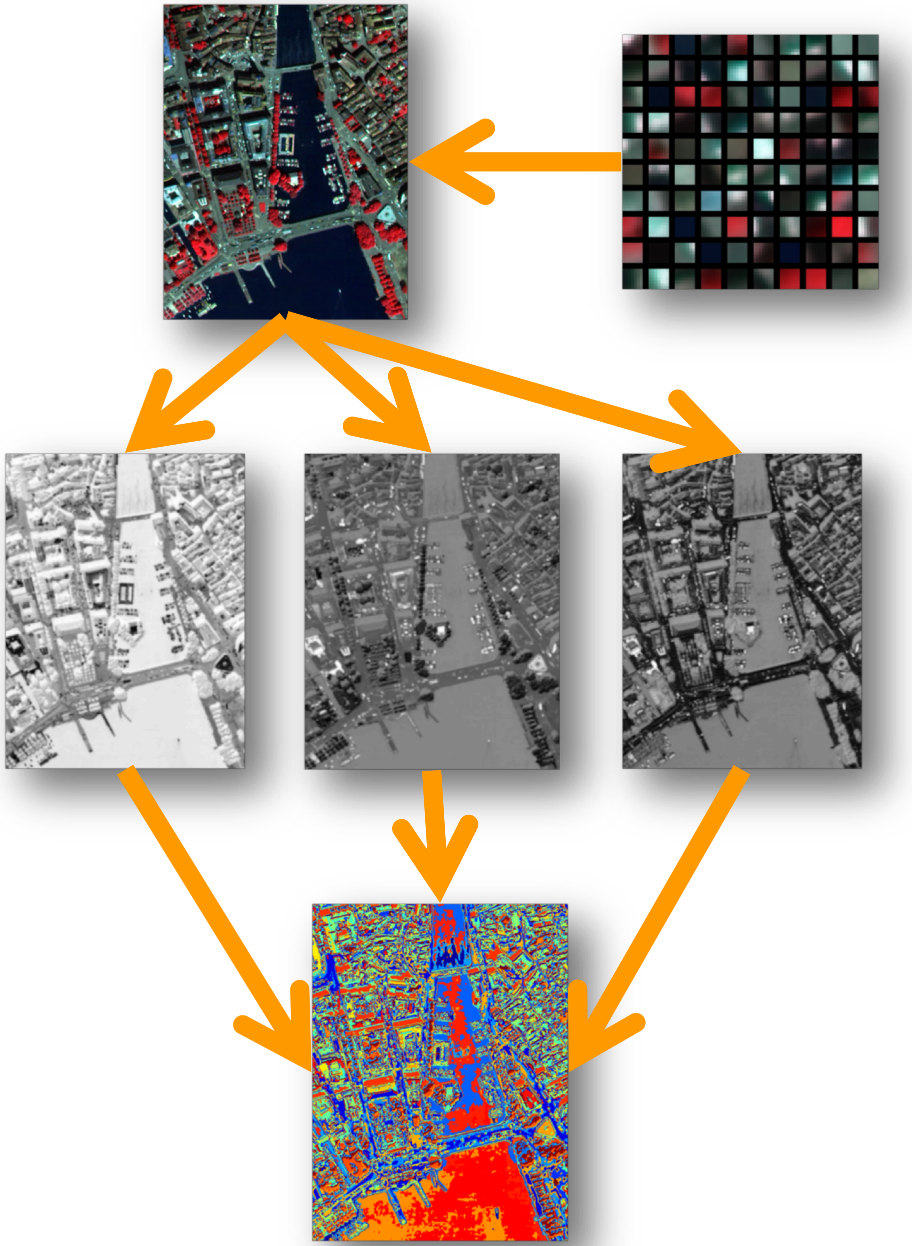
[Kellenberger et al., IEEE IGARSS 2017]

SDSC

Spatial context

- Injecting spatial context into a model:
  - **Inputs**: Spatial arrangement of patches of pixels (e.g. visual words)
  - **Outputs**: Interactions between classes (class co-occurrence: $p(y_i,y_j)$)
  - **Both**: Model-based invariances (or equivariance, covariance)

    e.g. [Marcos et al., ICCV 2017] for rotations



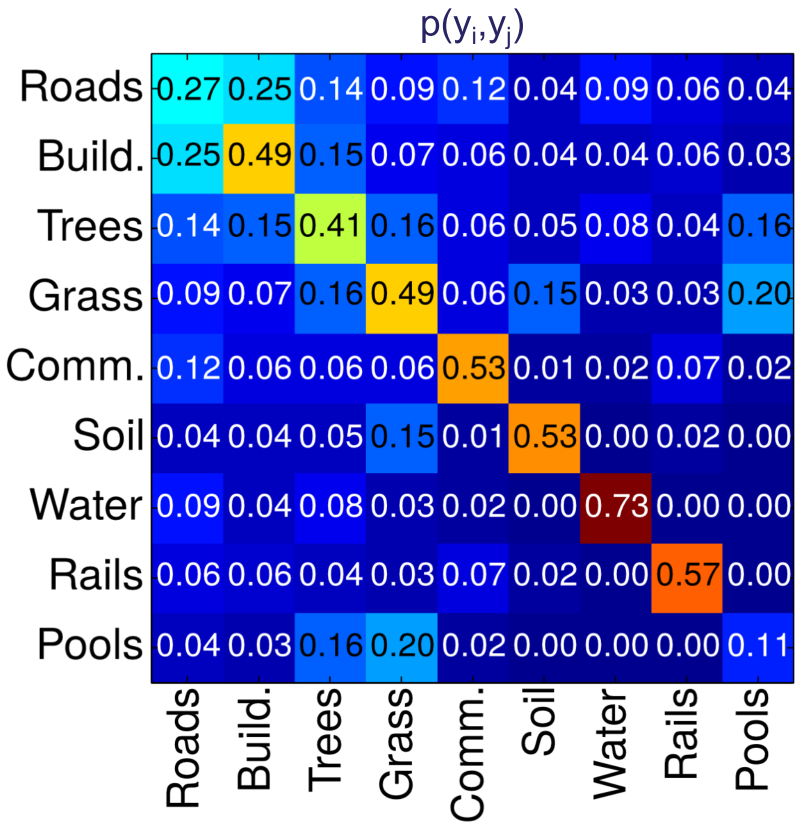
SDSC

Spatial context

[Volpi and Ferrari, JURSE 2015]

- Injecting spatial context into a model:
  - **Inputs**: Spatial arrangement of patches of pixels (e.g. visual words)
  - **Outputs**: Interactions between classes (class co-occurrence: $p(y_i,y_j)$)
  - **Both**: Model-based invariances (or equivariance, covariance)

    e.g. [Marcos et al., ICCV 2017] for rotations



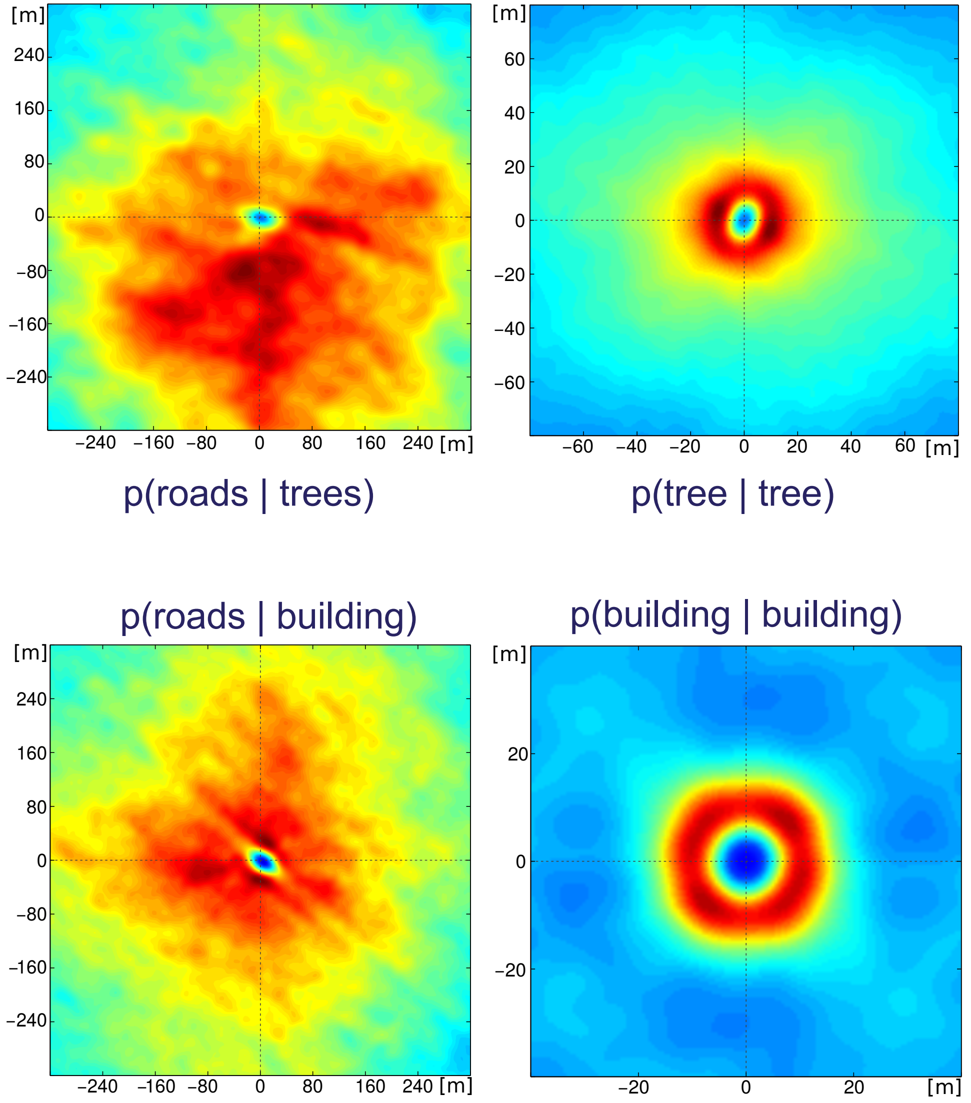
Spatial context

[Volpi and Ferrari, JURSE 2015]

- Injecting spatial context into a model:
  - **Inputs**: Spatial arrangement of patches of pixels (e.g. visual words)
  - **Outputs**: Interactions between classes (class co-occurrence: $p(y_i,y_j)$)
  - **Both**: Model-based invariances (or equivariance, covariance)

    e.g. [Marcos et al., ICCV 2017] for rotations

- Closer things are (generally) more related than those far apart
  - Geographic context



SDSC

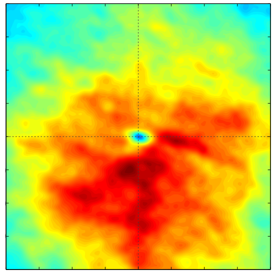
Spatial context

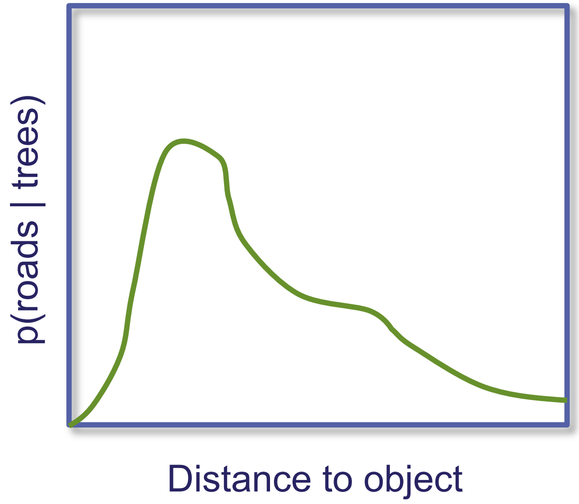
[Volpi and Ferrari, JURSE 2015]
[Volpi and Ferrari, CVPRW 2015]

- Injecting spatial context into a model:
  - **Inputs**: Spatial arrangement of patches of pixels (e.g. visual words)
  - **Outputs**: Interactions between classes (class co-occurrence: $p(y_i,y_j)$)
  - **Both**: Model-based invariances (or equivariance, covariance)

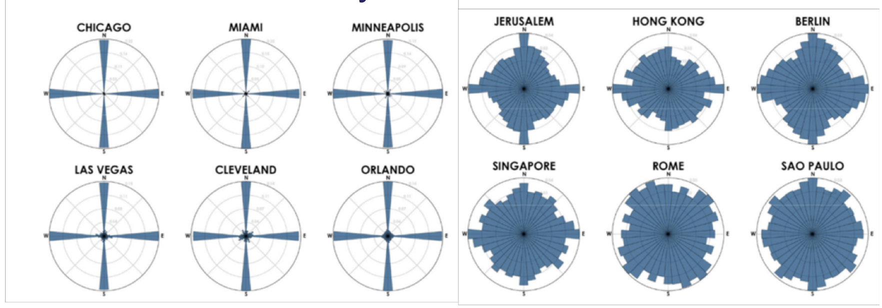
    e.g. [Marcos et al., ICCV 2017] for rotations

- A rotation in the input should not affect the output





Road network analysis



[Boeing, SSRN 2018]

SDSC

Geo-context and spatial smoothness

[Volpi and Ferrari, CVPRW 2015]

- Spatial smoothness
  - Contrasts
- Locally invariant class co-occurrence
  - Geography

SDSC

Learning the context

[Volpi and Ferrari, CVPRW 2015]

- Formulated as a structured output learning problem (over a CRF):
  - Learn mapping from inputs in isolations to class-likelihoods
    - Features locally describing the appearance
  - Learn optimal relationships between local outputs
    - Define a neighborhood system and learn relationships between classes
- Learned using Structured SVM
[Tsochantaridis et al., JMLR 2005]

$$E(\mathbf{x}, \mathbf{y}; \mathbf{w}) =$$

$$\sum_{i \in \mathcal{V}} \varphi_i(y_i, x_i; \mathbf{w}^{\varphi})$$

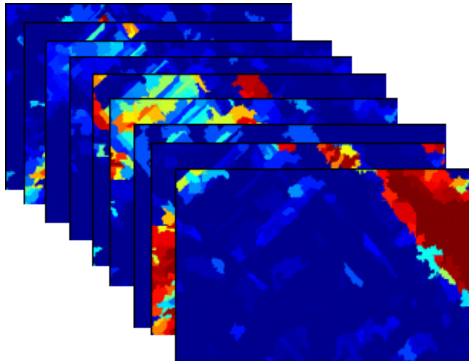
$$+ \sum_{(ij) \in \mathcal{E}} \phi_{ij}(y_i, y_j, x_i, x_j; \mathbf{w}^{\phi})$$

$$= \langle \mathbf{w}, \mathbf{\Phi}(\mathbf{x}, \mathbf{y}) \rangle$$

Learning the context

[Volpi and Ferrari, CVPRW 2015]

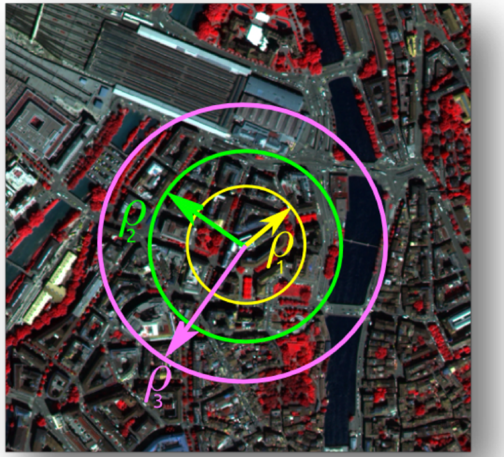
- “Unary” features standard appearance



$$E(\mathbf{x}, \mathbf{y}; \mathbf{w}) =$$

$$\sum_{I \in \mathcal{V}} \varphi_i(y_i, x_i; \mathbf{w}^{\varphi})$$

- Dense “Pairwise” relationships within quantized spatial rings
  - Invariant to rotation
  - Flexible geographical relationships
  - Robust to little and non-dense training data

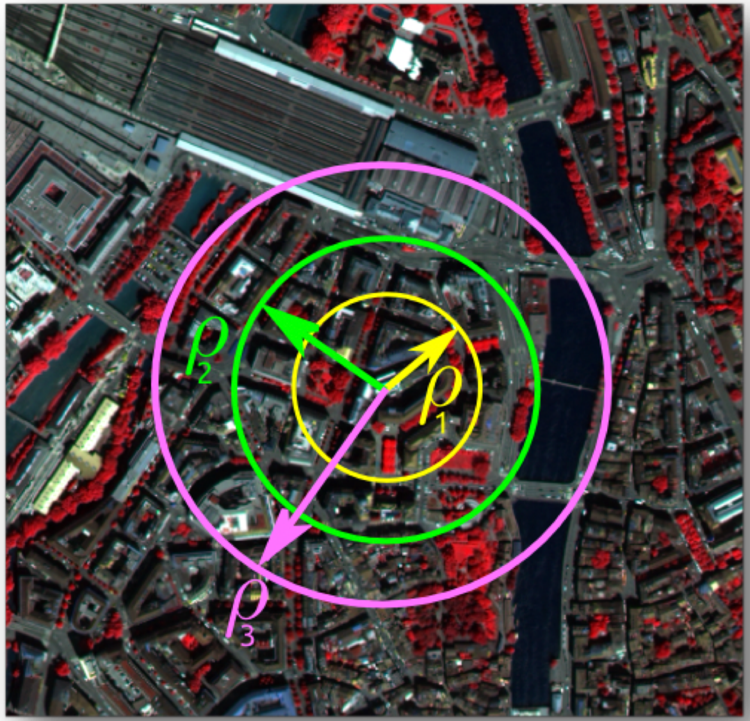


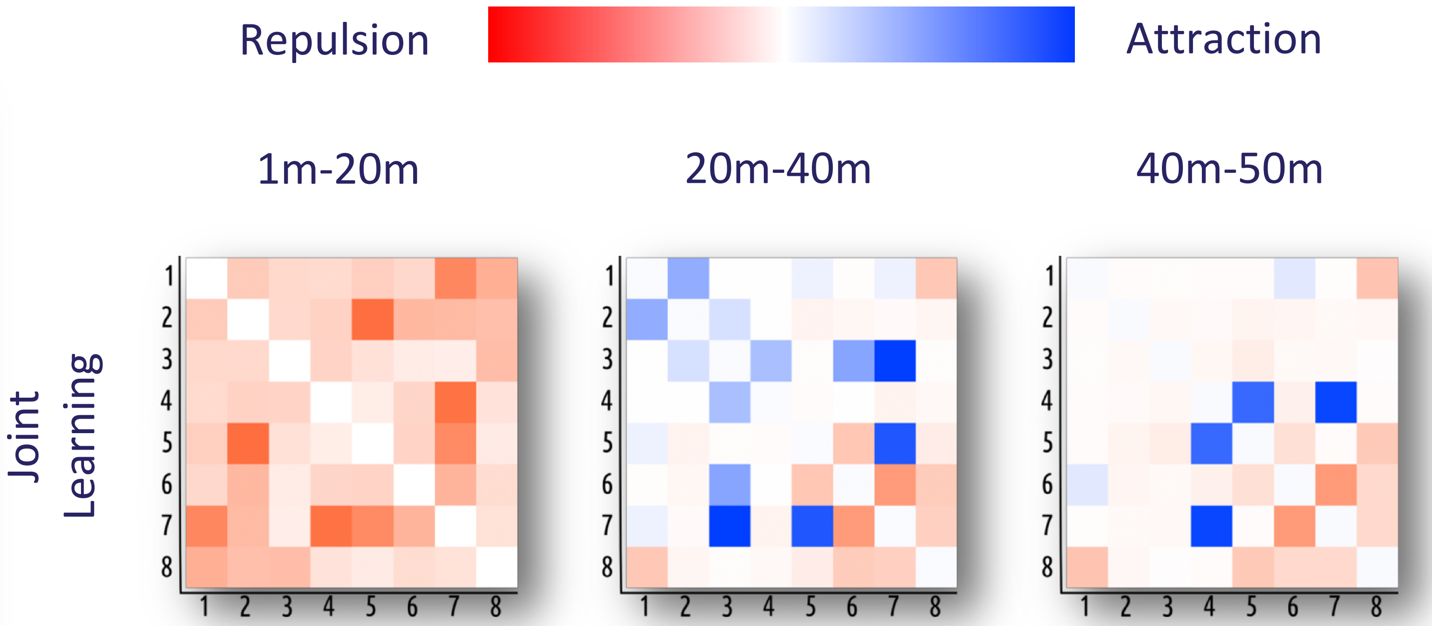
$$+ \sum_{(ij) \in \mathcal{E}} \phi_{ij}(y_i, y_j, x_i, x_j; \mathbf{w}^{\phi})$$

$$= \langle \mathbf{w}, \mathbf{\Phi}(\mathbf{x}, \mathbf{y}) \rangle$$

The learned context

[Volpi and Ferrari, CVPRW 2015]



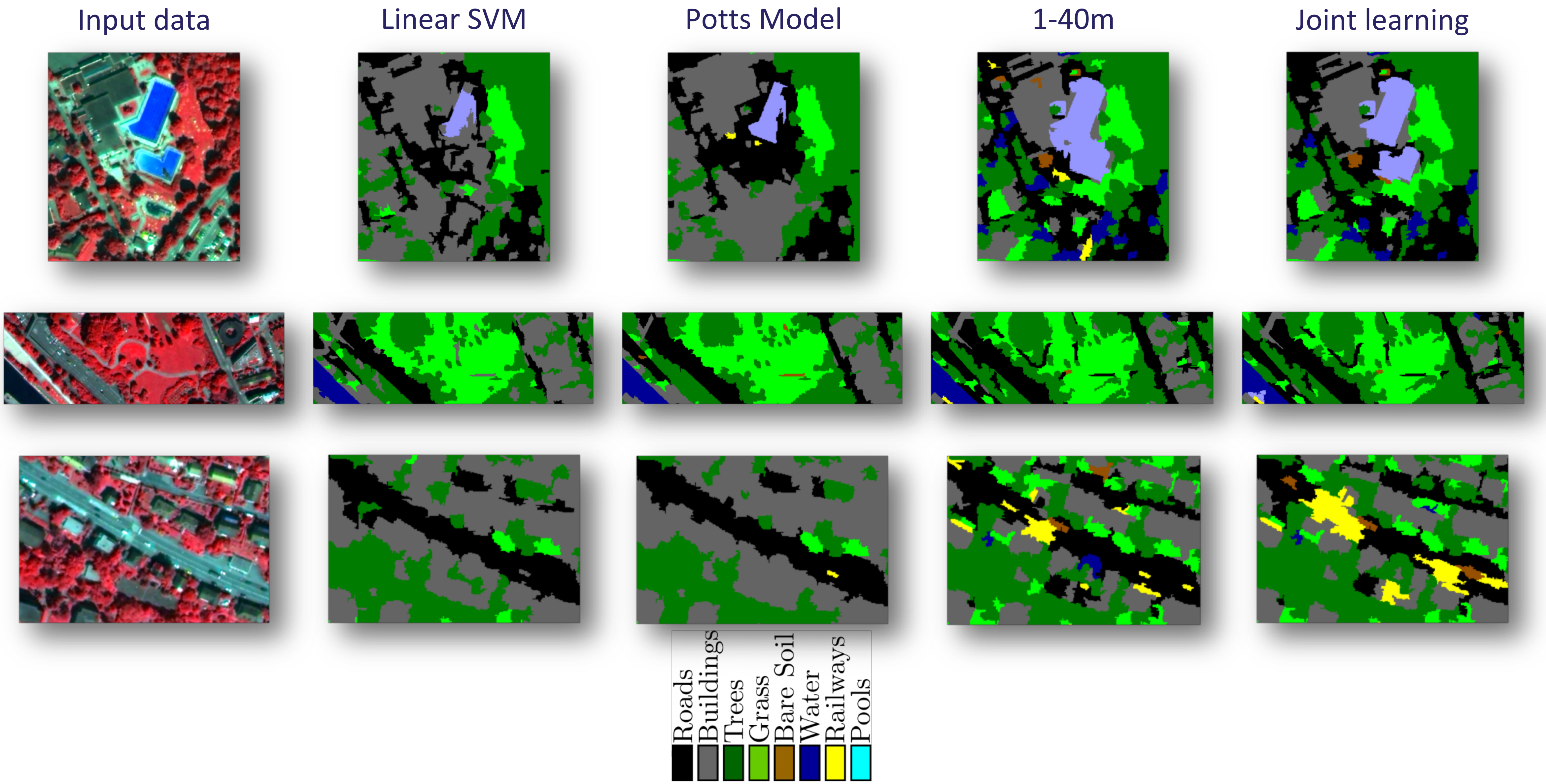


1: Roads 2: Build. 3: Trees 4: Grass 5: Soil 6: Water 7: Rails 8: Pools

SDSC

The learned context

[Volpi and Ferrari, CVPRW 2015]



Does the context help?

- Results improve, but still not optimal: why?
  - Lots of visual ambiguity between classes, linearity of relationships
- Enforces prior beliefs about the problem, very useful for small training data
- More data! More Learning! More weights! More nonlinearities!

Does the context help?

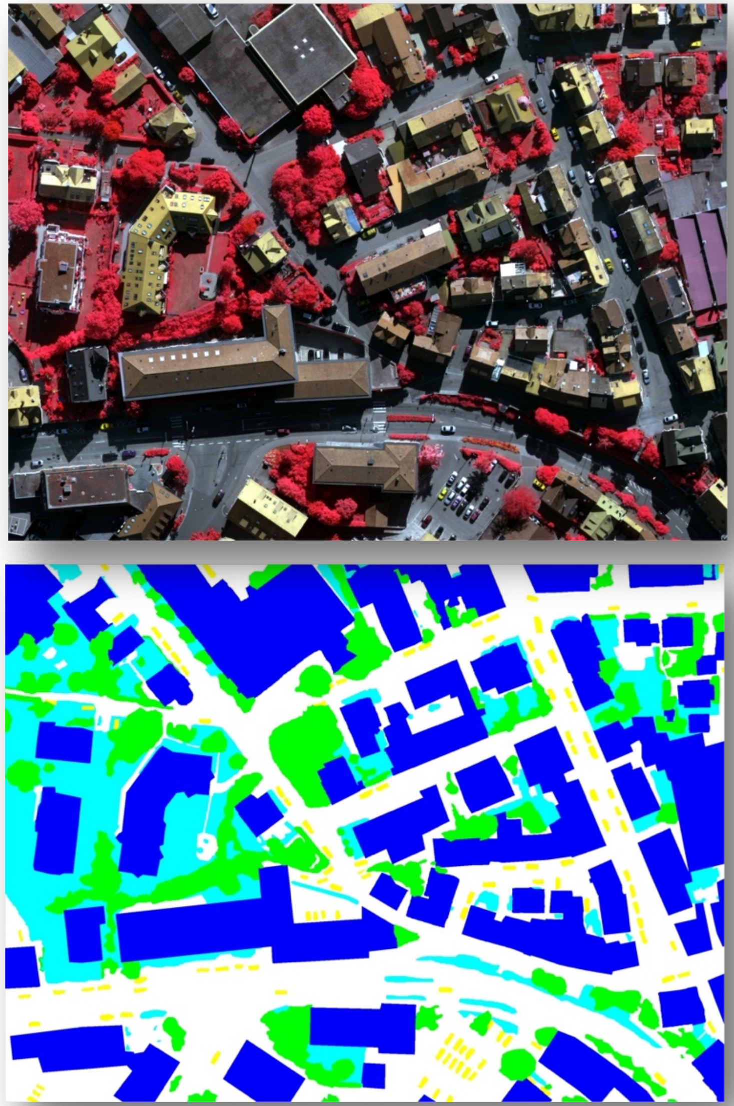
- Results improve, but still not optimal: why?
    - Lots of visual ambiguity between classes, linearity of relationships
- Enforces prior beliefs about the problem, very useful for small training data
- More data! More Learning! More weights! More nonlinearities!

[Campos-Taberner et al., IEEE JSTARS 2016]

SDSC

Does the context help?

- Results improve, but still not optimal: why?
  - Lots of visual ambiguity between classes, linearity of relationships
- Enforces prior beliefs about the problem, very useful for small training data
- More data! More Learning! More weights! More nonlinearities!

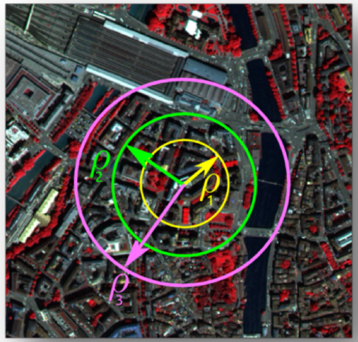
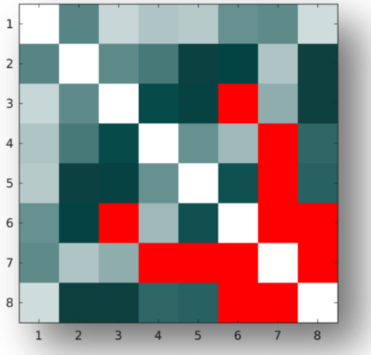


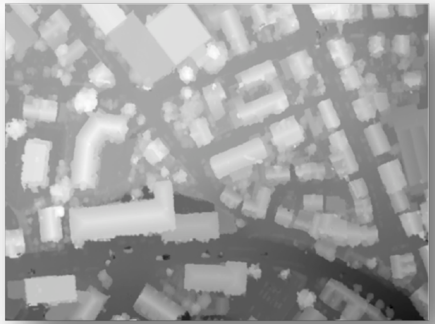
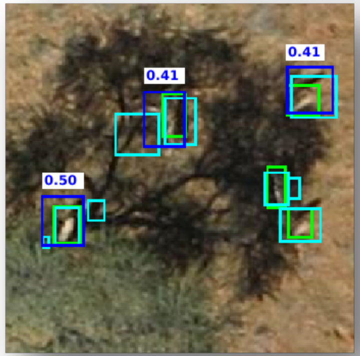
[Rottensteiner et al., JISPRS SI 2014]

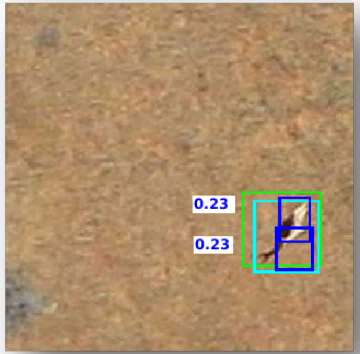
SDSC

Side information

- What can we use to inform models?
  - Additional priors related to the problem (e.g. spatial and geographic context)
  - Additional inputs (e.g. DSM, time series, image features, …)
  - Additional outputs (e.g. object detection, semantic boundaries, …)



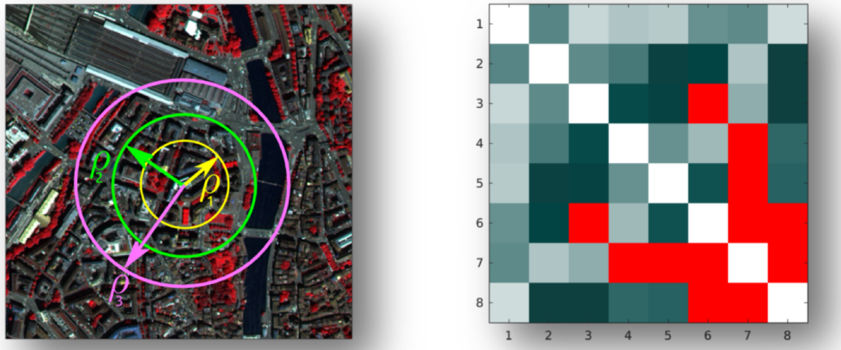
[Kellenberger et al., IEEE IGARSS 2017]

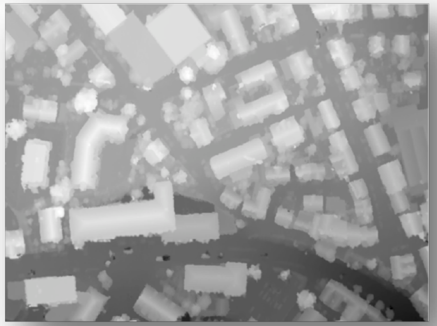
SDSC

Side information

[Volpi and Tuia, JISPRS 2018]

- What can we use to inform models?
  - Additional priors related to the problem (e.g. spatial and geographic context)
  - **Additional inputs** (e.g. DSM, time series, image features, …)
  - **Additional outputs** (e.g. object detection, semantic boundaries, …)





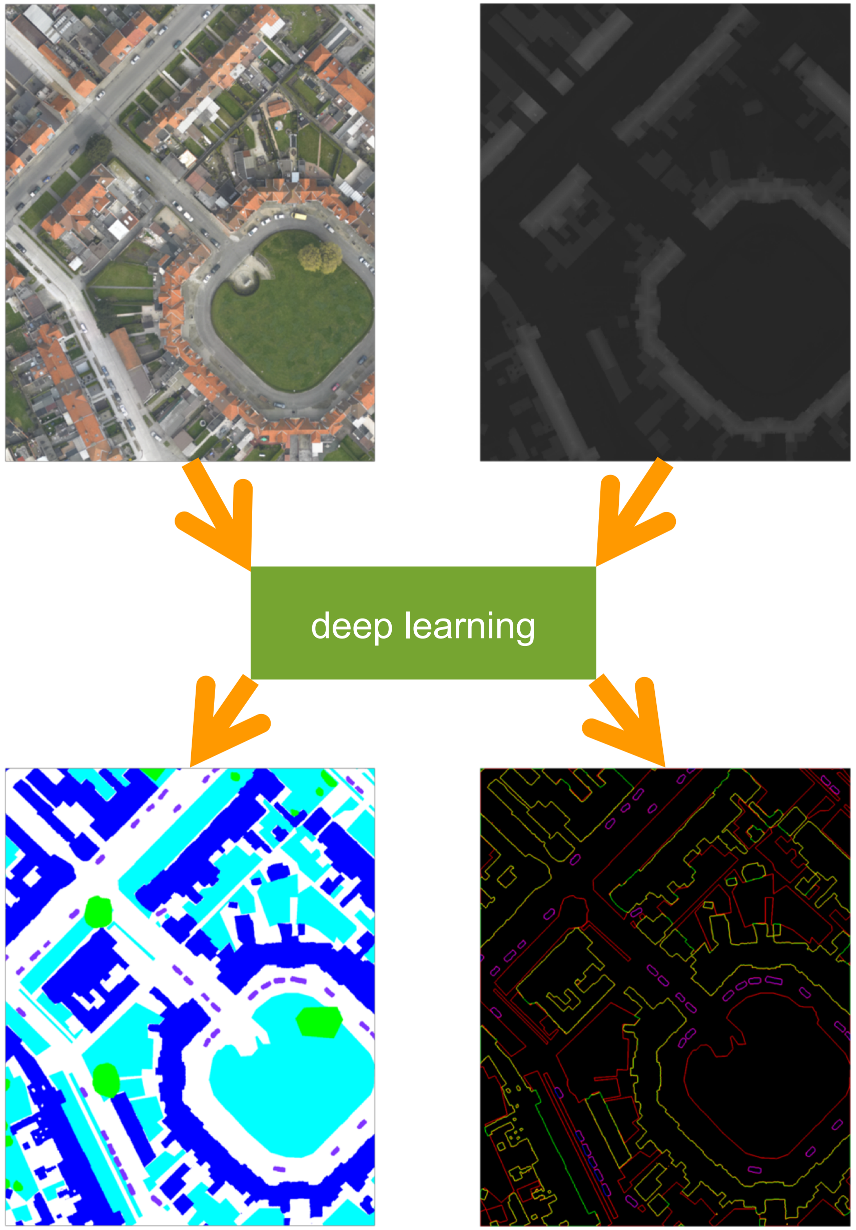
[Kellenberger et al., IEEE IGARSS 2017]

Multi-task segmentation

[Volpi and Tuia, JISPRS 2018]

- Multi-task learning
  - Give same inputs, learn a model able to predict several outputs at the same time

- **Jointly** learn semantic segmentation and semantic edges
  - Related tasks, mutually informative [Marmanis et al., JISPRS 2018]



SDSC

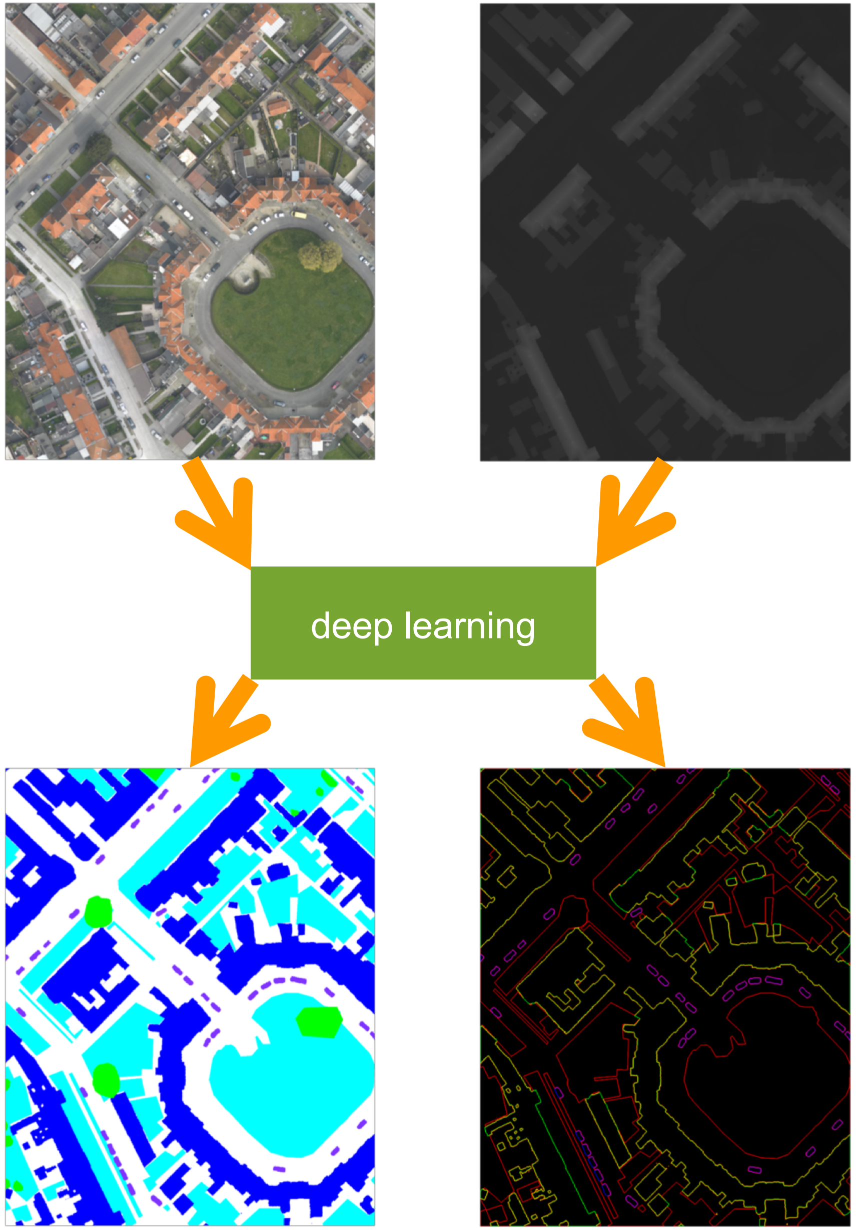
Multi-task learning

[Volpi and Tuia, JISPRS 2018]

- **Jointly** learn semantic segmentation and semantic edges
  - Related tasks, mutually informative
    [Marmanis et al., JISPRS 2018]

- Learn a shared representation between two related tasks

$$\mathcal{L}(y, \hat{y}; x) = \underbrace{\beta_s \mathcal{L}^s(y_s, \hat{y}_s; x)}_{\text{Segmentation loss}} + \underbrace{\beta_e \mathcal{L}^e(y_e, \hat{y}_e; x)}_{\text{Boundary loss}}$$



SDSC

Multi-task learning

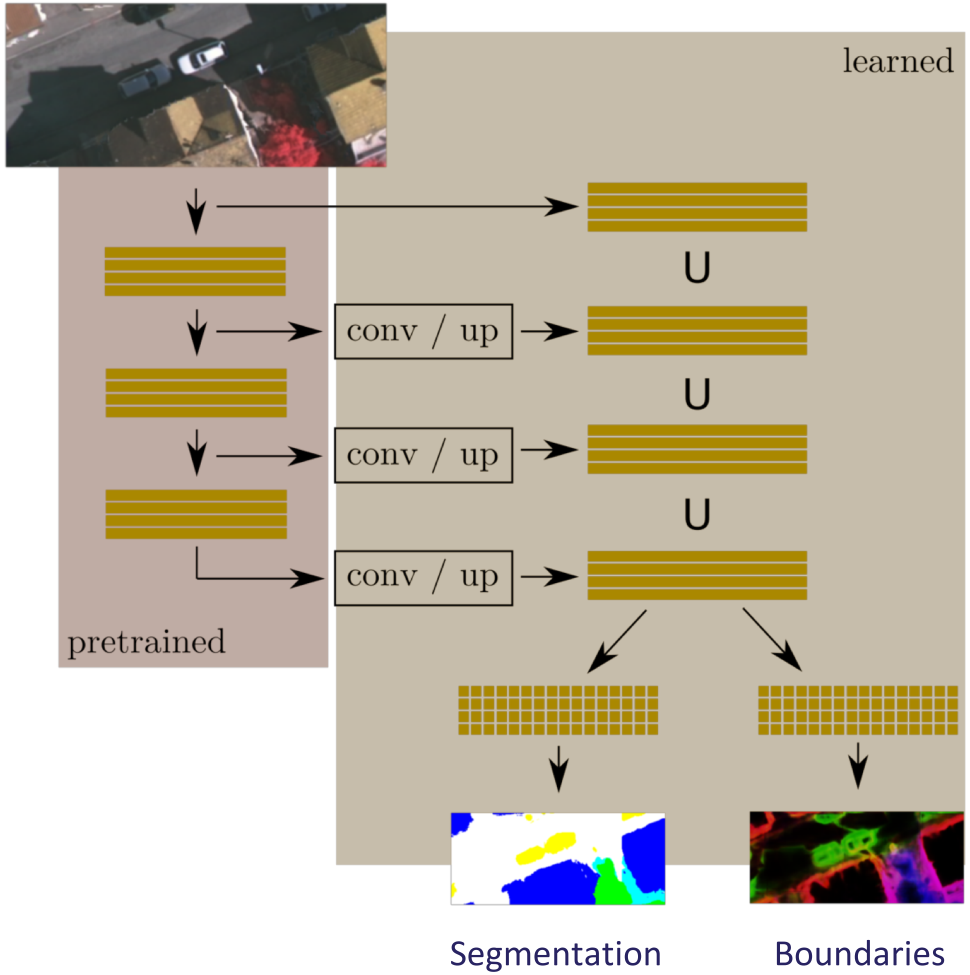
[Volpi and Tuia, JISPRS 2018]

- **Jointly** learn semantic segmentation and boundaries
  - Related tasks, mutually informative
    [Marmanis et al., JISPRS 2018]

- Learn a shared representation between two related tasks

$$\mathcal{L}(y, \hat{y}; x) = \underbrace{\beta_s \mathcal{L}^s(y_s, \hat{y}_s; x)}_{\text{Segmentation loss}} + \underbrace{\beta_e \mathcal{L}^e(y_e, \hat{y}_e; x)}_{\text{Boundary loss}}$$

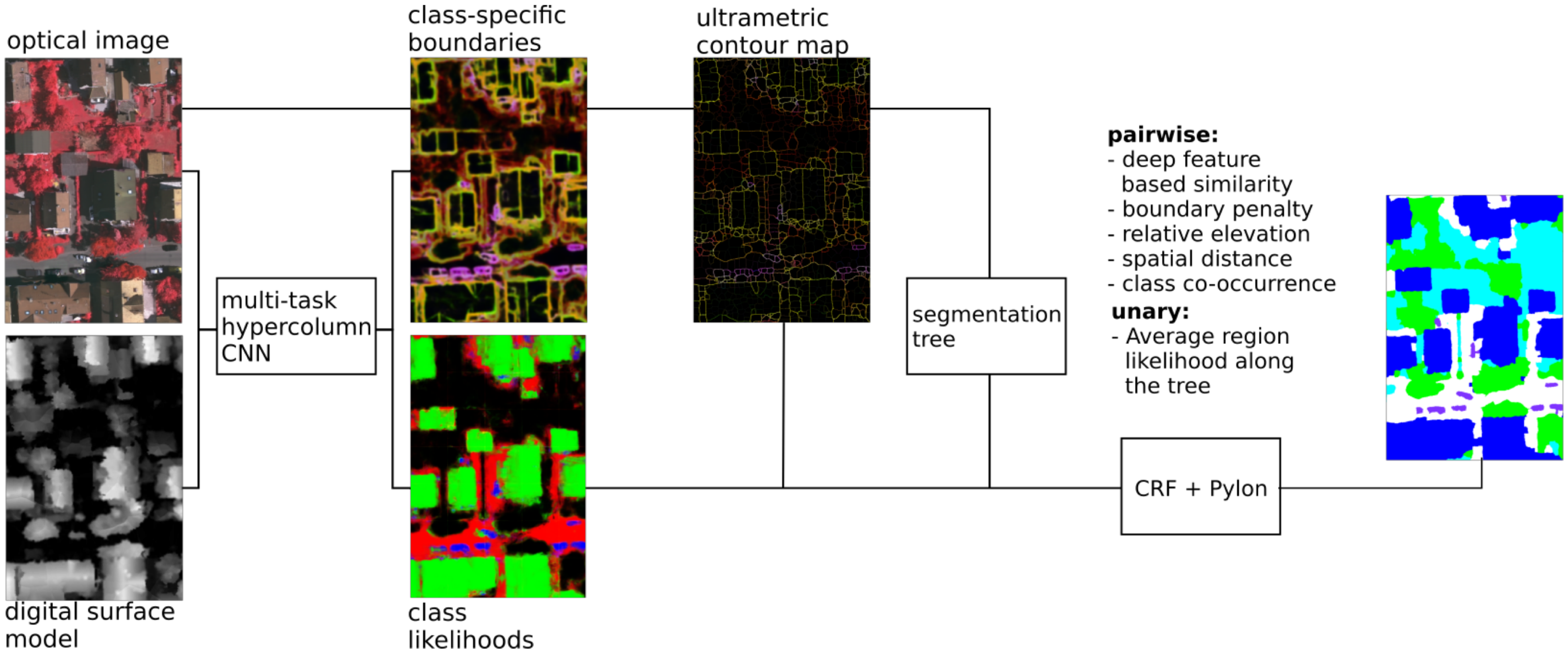
- Tie parameters of a CNN (hypercolumns architecture)



SDSC

Full segmentation pipeline

[Volpi and Tuia, JISPRS 2018]

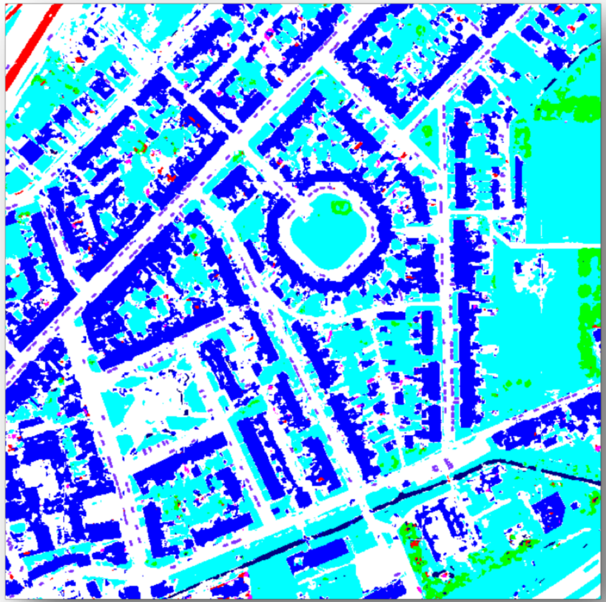


Full segmentation pipeline

[Volpi and Tuia, JISPRS 2018]

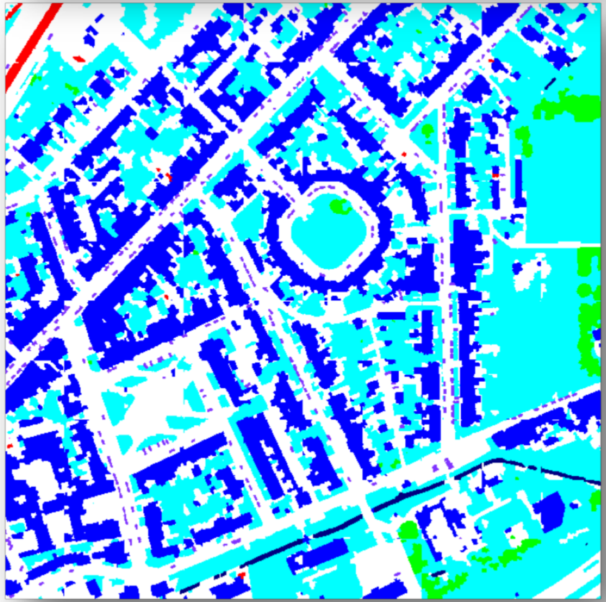
Input data RGB

Only CNN



Ground Truth

Full pipeline

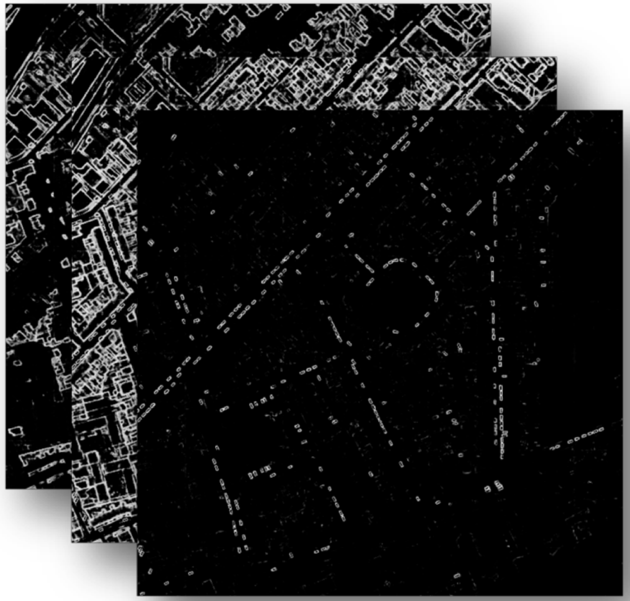


SDSC

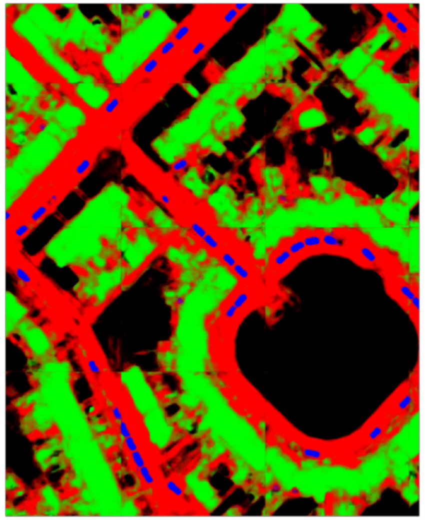
“Side products”

[Volpi and Tuia, JISPRS 2018]

Semantic boundaries



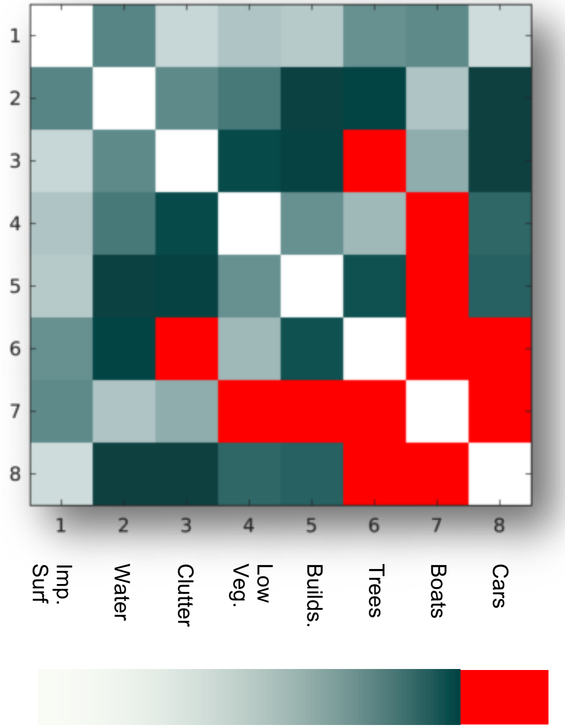
∝ “Semantic” Probabilities



Class likelihoods

Edge-based hierarchy (UCM)

Co-occurrences (hierarchy)



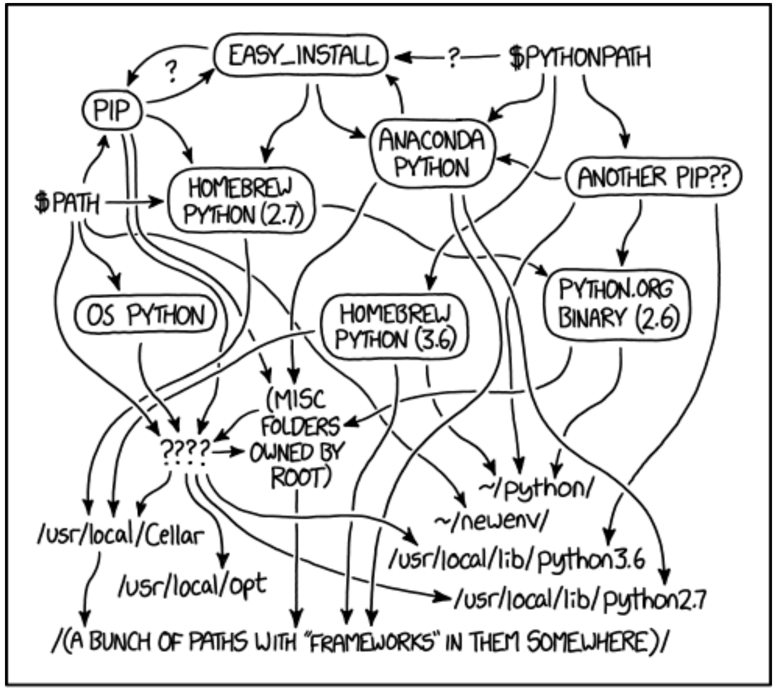
SDSC

All nice but…

- Models *and* data become increasingly complex:

  - No way to **plug-and-reuse** different pipeline parts
  - Full reproducibility **does not** rely only on source code availability
    - data updates, package versions, environment definition, dependencies, etc.
  - Same for **collaboration**

Relevant XKCD: xkcd.com/1987/



MY PYTHON ENVIRONMENT HAS BECOME SO DEGRADED THAT MY LAPTOP HAS BEEN DECLARED A SUPERFUND SITE.

Try RENKU!

- SDSC contribution: RENKU 連句
  - Personalized environments in the cloud (docker)
  - Jointly version data, code, outputs and modularly relate them
  - Versions independent “runs”
  - Reuse, trace the use, allow full reproducibility
  - Open source

-> datascience.ch/solutions
-> renkulab.io

Summing up

- Prior knowledge helps for low data regimes
  - Expert and domain knowledge
- Additional data provides additional evidence as long as models can ingest it
  - Reusable information available
- Sharing is caring
  - Reproducible and reusable developments

-> datascience.ch/solutions

-> renkulab.io

carvingthroughdata.ch

Carving Through Data

A different introductory course to Machine Learning

11 -15 March, 2019 Laax, Switzerland

SDSC

ETH zürich

EPFL
ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

PAUL SCHERRER INSTITUT
PSI

WSL
Eidg. Forschungsanstalt für Wald
Schnee und Landschaft WSL

Empa
Materials Science and Technology

eawag
aquatic research