Efficient Convolutional Patch Networks for Scene Understanding

since 1558

Code Available

Clemens-Alexander Brust, Sven Sickert, Marcel Simon, Erik Rodner and Joachim Denzler

Computer Vision Group, Friedrich Schiller University Jena, Germany

http://www.inf-cv.uni-jena.de

Friedrich Schiller University Jena
Computer Vision Group

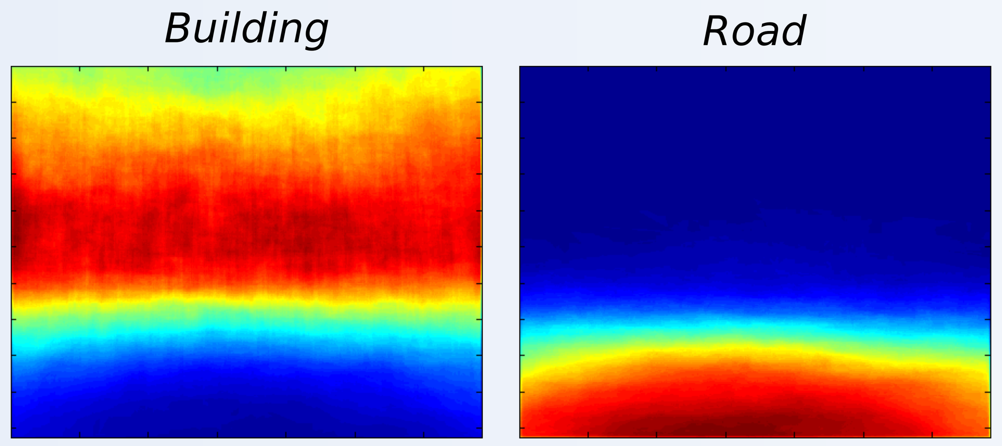
CN24 Open Source Deep Learning Framework

CNN Framework especially designed for scene understanding [1]

- Graph-based network structure can be freely specified
- **OpenCL**, MKL, ACML and dependency-free reference backends
- **Hybrid patch-wise and fully convolutional design** for fast prediction and efficient training
- 3-clause **BSD license** suitable for research and commercial applications

cvjena.github.io/cn24/

Spatial prior for

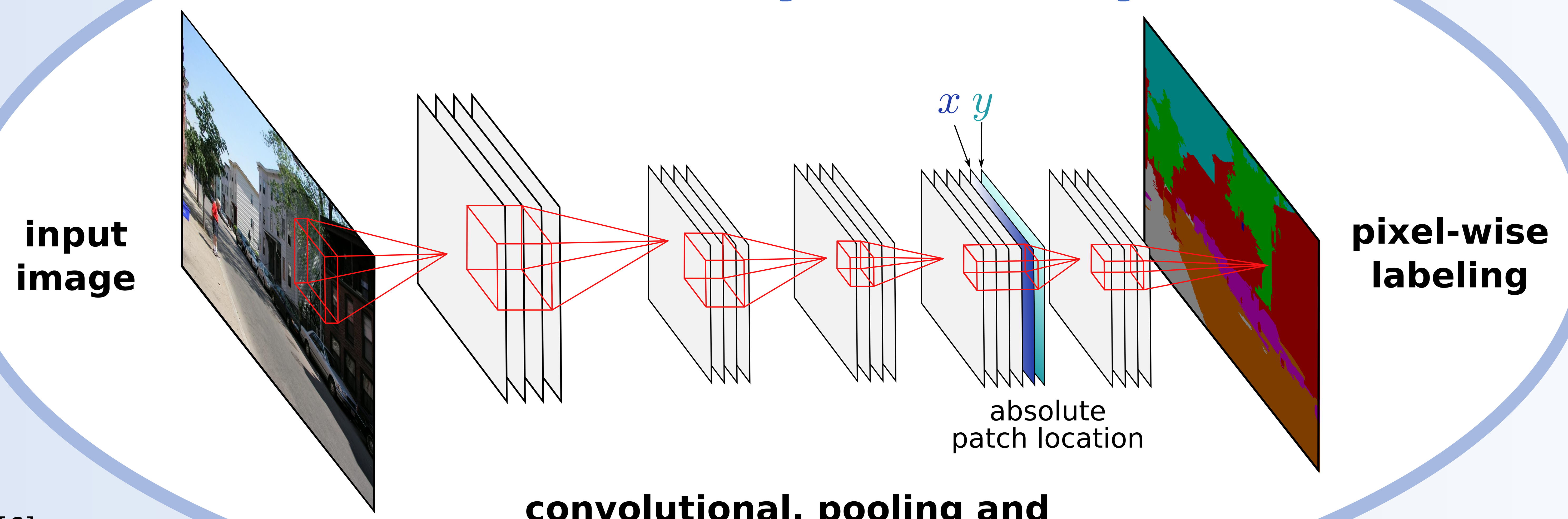


Incorporating Spatial Priors

- Adding the location of a patch to the network's input **increases classification performance significantly**
- Simple solution: We add additional input layers with pixel coordinates

Visit our project page!

www.inf-cv.uni-jena.de/semseg



Qualitative Results

Method Details

- **Fully convolutional networks (FCN)** [6] for prediction speed-up (up to 100x) over patch-wise approach
- **Patch-wise training** leads to faster optimization
- Real-time segmentation of VGA-sized inputs
- Incorporating position information as **spatial priors**
- Post-processing: **quality enhancement** of network outputs using unsupervised segmentation [5]
- Weighting: optimization with **inverse class frequency weights** accounts for imbalanced training set

Experiments

- Experiments on LabelMeFacade [2,3]
- 945 images, 8 classes
- Performance evaluated with average and overall recognition rate (ARR and ORR)

Quantitative Results

Method	ORR	ARR
RDF+SIFT [2]	49.06%	44.08%
ICF [3]	67.33%	56.61%
RDF-MAP [4]	71.28%	-
Our Approach		
CNN outputs (FCN training)	58.17%	29.48%
CNN outputs (patch-wise training)	67.87%	42.89%
+ spatial prior	72.21%	47.74%
+ post processing	**74.33**%	47.77%
+ weighting	63.41%	**58.98**%

[1] Brust et al.: Convolutional Patch Networks with Spatial Prior for Road Detection and Urban Scene Understanding, In VISAPP, 2015, 510-517
[2] Fröhlich, Rodner, Denzler: A Fast Approach for Pixelwise Labeling of Facade Images, In ICPR, 2010, 3029-3032
[3] Fröhlich, Rodner, Denzler: Semantic Segmentation with Millions of Features: Integrating Multiple Cues in a Combined Random Forest Approach, In ACCV, 2012, 218-231
[4] Nowozin: Optimal Decisions from Probabilistic Models: The Intersection-over-Union Case, In CVPR, 2014, 548-555
[5] Felzenszwalb, Huttonlocher: Efficient Graph-Based Image Segmentation, In IJCV 59(2), 2004, 167-181
[6] Long, Shelhamer, Darrell: Fully Convolutional Networks for Semantic Segmentation In CVPR, 2015