HAL open science

Hierarchical Probabilistic Graphical Models and Deep Convolutional Neural Networks for Semantic Segmentation of Remote Sensing Images

Martina Pastorino, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia

To cite this version:

Martina Pastorino, Gabriele Moser, Sebastiano B. Serpico, Josiane Zerubia. Hierarchical Probabilistic Graphical Models and Deep Convolutional Neural Networks for Semantic Segmentation of Remote Sensing Images. Workshop on Machine Learning and Signal Processing on Graphs / Apprentissage Automatique et Traitement du Signal sur Graphes, Nov 2022, Marseille (CIRM, Centre International de Rencontres Mathématiques), France. hal-04269412

HAL Id: hal-04269412

https://inria.hal.science/hal-04269412v1

Submitted on 3 Nov 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Hierarchical Probabilistic Graphical Models and Deep Convolutional Neural Networks for Semantic Segmentation of Remote Sensing Images

Martina Pastorino[1,2], Gabriele Moser[1], Sebastiano B. Serpico[1], and Josiane Zerubia[2]

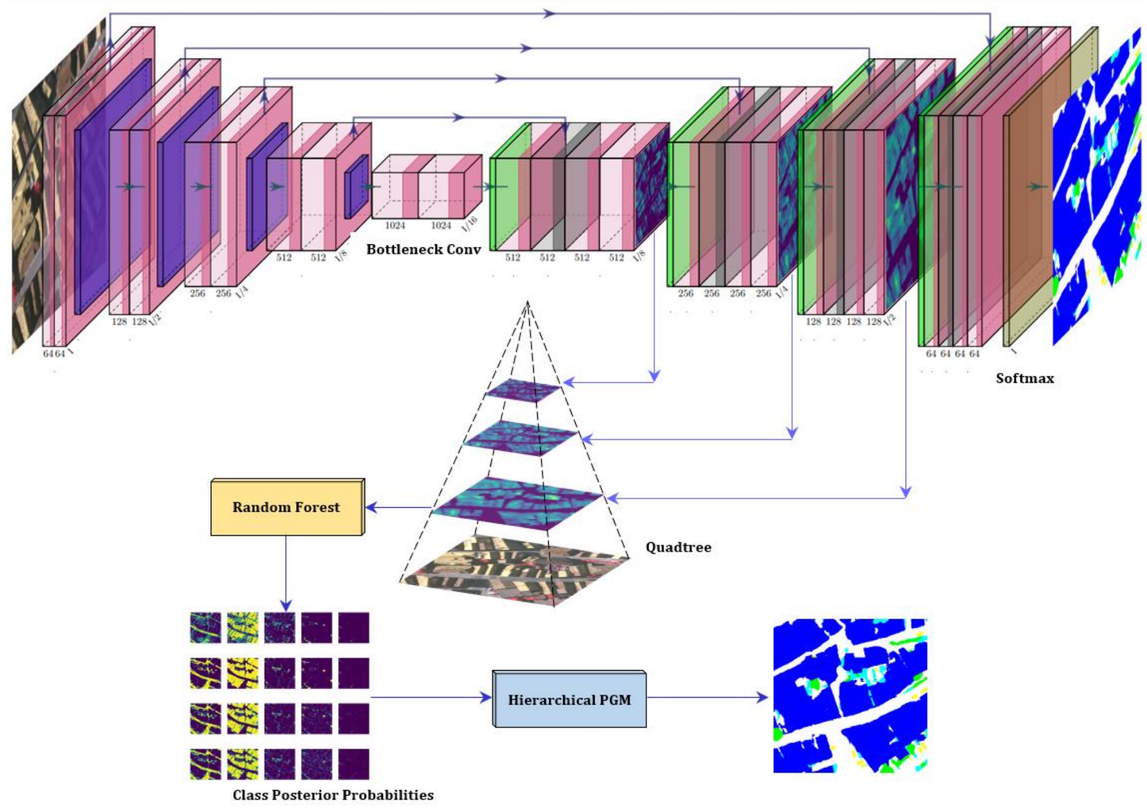
[1]Department of Electrical, Electronic, Telecommunications Engineering, and Naval Architecture
University of Genoa, Italy
[2]INRIA d'Université Cote d'Azur, France

INTRODUCTION

This work deals with the challenge of semantic segmentation based on deep learning methods in the case of realistic scarce ground truth maps. Exhaustive ground truths usually found in benchmark datasets can be used to train deep learning architectures successfully. On the contrary, real-world ground truths are almost never exhaustive, they are spatially sparse and typically they don't represent the spatial borders among the classes, affecting the accuracy of the resulting segmentation maps significantly.

Here, the proposed approach addresses precisely this challenge with a novel combination of hierarchical probabilistic graphical models (PGMs) and deep neural networks. The rationale is to exploit the spatial modeling capabilities of hierarchical PGMs to mitigate the impact of incomplete ground truth and obtain accurate classification results in scenarios where exhaustive ground truth does not exist.

PROPOSED METHOD



The idea is to integrate the multiscale data representation extracted by a fully convolutional network (FCN) with a hierarchical PGM through a quadtree topology:

1. An FCN is trained with a dataset of very high-resolution images. Activations at L different blocks are associated with a quadtree topology.
2. An ensemble approach (Random Forest, RF) links the representation extracted by the FCN and the Bayesian inference structure of the PGM.
3. A hierarchical PGM is defined on the quadtree to model multiscale and long-range spatial dependencies over the pixel lattice.

A neighborhood relation $\precsim$ is introduced to ensure causality over the pixel lattice. The proposed PGM is causal and the joint distribution of all labels and feature vectors in the quadtree is entirely defined by

- the transition probabilities across the scales $P(x_s|x_{s^-})$
- the past neighbor transition probabilities within each layer $P(x_s|x_r, r \precsim s)$
- the pixelwise class-conditional likelihoods $P(y_s|x_s)$

and the inference is addressed through the marginal posterior mode criterion:

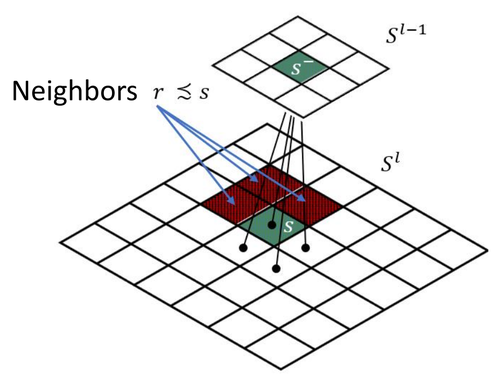
$$P(x_s) = \sum_{x_{s^-}} P(x_s|x_{s^-})P(x_{s^-})$$

$$P(x_s|y_s^d) \propto P(x_s|y_s) \prod_{t \in s^+} \sum_{x_t} \frac{P(x_t|y_t^d)P(x_t|x_s)}{P(x_t)}$$

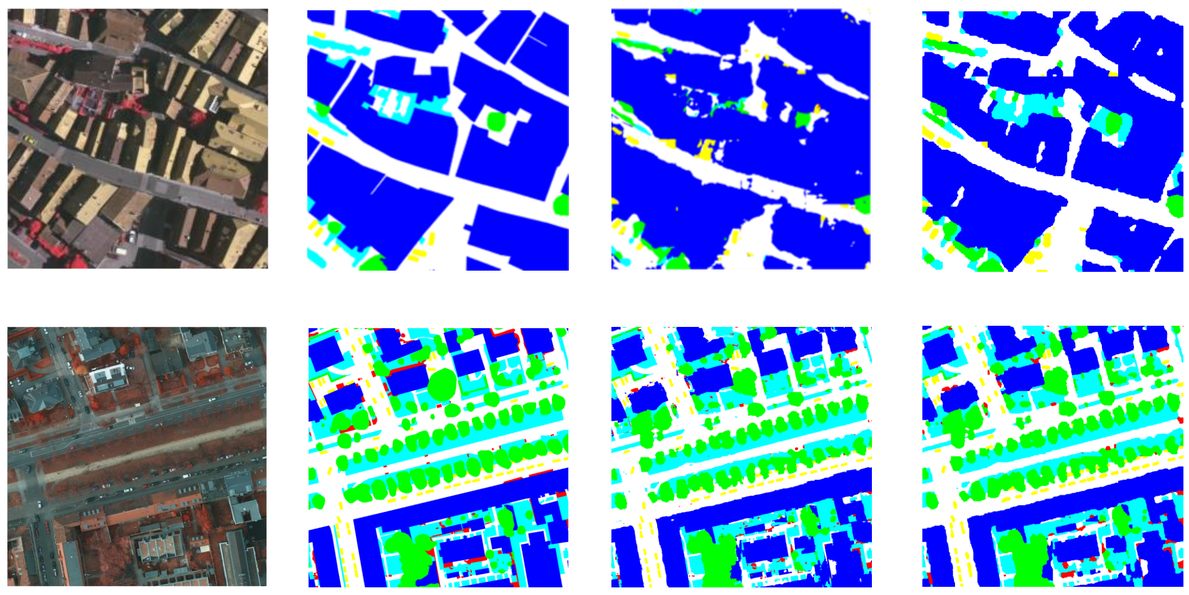
$$P(x_s|x_s^c, y_s^d) \propto P(x_s|y_s^d)P(x_s|x_{s^-})P(x_{s^-})P(x_s)^{-n_s}$$

$$P(x_s|\mathcal{Y}) = \sum_{x_s^c} P(x_s|x_s^c, y_s^d)P(x_{s^-}|\mathcal{Y})$$

$$\max_{x_s} P(x_s|\mathcal{Y})$$



RESULTS



(a) (b) (c) (d)

Figure 1. Ground truth and classification maps: (a) IRRG image, (b) ground truth; classification maps obtained by (c) the FCN (U-Net) and (d) the proposed method. Classes: building, road, vegetation, tree, car, clutter.

The ISPRS semantic labeling challenge Potsdam and Vaihingen datasets were used for the experiments. To approach realistic ground truths, several training conditions were considered, e.g., the removal of a percentage of labeled pixels and morphological erosion. The estimation of the posterior probabilities used to link the FCN and the PGM was also performed in different ways, resulting in different formulations of the proposed method.

	Architecture	building	road	vegetation	tree	car	OA	recall	precision	K	F1 score
Full dataset	Standard U-Net	0.92	0.83	0.71	0.92	0.74	0.85	0.83	0.84	0.79	0.83
	Standard SegNet	0.90	0.73	0.73	0.92	0.67	0.83	0.79	0.81	0.76	0.80
	DeepLabV3+	0.91	0.87	0.66	0.90	0.80	0.84	0.83	0.82	0.78	0.82
	Proposed	0.87	0.84	0.71	0.92	0.89	0.82	0.83	0.78	0.75	0.80
Erosion	Standard U-Net	0.91	0.92	0.15	0.62	0.27	0.87	0.57	0.77	0.75	0.66
	Proposed	0.88	0.87	0.51	0.60	0.70	0.86	0.71	0.68	0.72	0.69
70% unlab.	Standard U-Net	0.83	0.92	0.55	0.87	0.80	0.80	0.79	0.74	0.72	0.76
	Proposed	0.82	0.85	0.76	0.86	0.96	0.82	0.85	0.72	0.76	0.78
Connected components	Standard U-Net	0.96	0.65	0.47	0.89	0.48	0.76	0.69	0.81	0.66	0.75
	Multiscale feature fusion	0.91	0.75	0.44	0.87	0.51	0.76	0.70	0.64	0.67	0.72
	Proposed	0.94	0.68	0.49	0.86	0.74	0.76	0.74	0.76	0.67	0.75

CONCLUSIONS

- Novel technique that combines FCNs and hierarchical probabilistic graphical models to overcome the critical problem of the impact of spatially sparse ground truth on semantic segmentation maps obtained by deep learning methods.
- Very effective in the discrimination of minority classes.
- Advantageous in terms of recall and computational burden.
- These advantages are more remarkable the more the training set approaches the real case scenario of spatially sparse ground truth.

REFERENCES

[1] M. Pastorino, G. Moser, S. B. Serpico, and J. Zerubia, "Semantic Segmentation of Remote-Sensing Images through Fully Convolutional Neural Networks and Hierarchical Probabilistic Graphical Models," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, no. 5407116, pp. 1-16, 2022.

[2] M. Pastorino, A. Montaldo, L. Fronda, I. Hedhli, G. Moser, S. B. Serpico, and J. Zerubia, "Multisensor and Multiresolution Remote Sensing Image Classification through a Causal Hierarchical Markov Framework and Decision Tree Ensembles," *Remote Sensing*, vol. 13, no. 849, Feb. 2021.

[3] M. Pastorino, G. Moser, S. B. Serpico, and J. Zerubia, "Semantic Segmentation of Remote Sensing Images Combining Hierarchical Probabilistic Graphical Models and Deep Convolutional Neural Networks," IGARSS 2021, Jul. 2021, Brussels, Belgium.