# Hierarchical Image-Region Labeling via Structured Learning

Julian McAuley[1,2], Teofilo de Campos[2,3], Gabriela Csurka[2], Florent Perronnin[2]

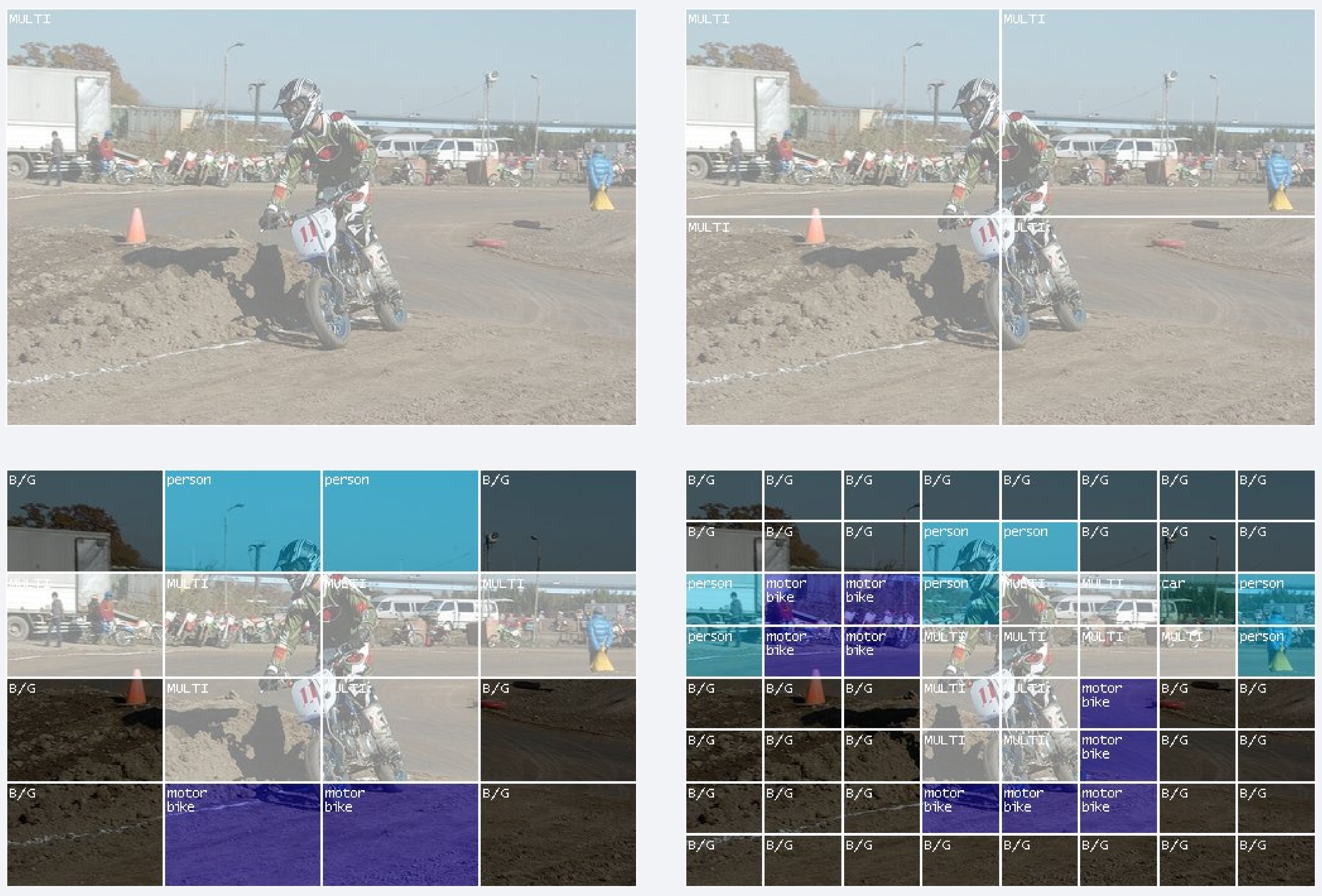
1: Australian National University/NICTA
2: Xerox Research Centre Europe
3: University of Surrey

## Abstract

We present a graphical model that encodes hierarchical constraints for **classifying image regions at multiple scales**. We show that **inference can be performed efficiently and exactly**, rendering it amenable to **structured learning**. Our model is parametrised using the outputs of a series of first-order classifiers, meaning that it **learns which classifiers are useful at different scales**, as well as the relationships between classifiers across scales.

## Our model

The 'nodes' of our graphical model correspond to overlapping image regions:




Edges are formed by connecting nodes at different scales: we connect two nodes precisely when the corresponding image regions overlap at adjacent scales, so that our graphical model forms a quad-tree.

## First-order (node) features

Our image features are based on those from [2], in which image-level, region-level, and patch-level classifiers are proposed.

We use **all classifiers** at **all scales** ($P_{r,label}$ is the probability that the region $r$ is labeled *label*):

$$\phi^{\text{nodes}}(r, label) = \underbrace{(0,\dots,P^1_{r,label},\dots,0)}_{\text{features for first classifier}}\dots\underbrace{(0,\dots,P^n_{r,label},\dots,0)}_{\text{features for } n^{th} \text{ classifier}}$$

Thus we **learn which classifiers are useful at which scales**.

## Hierarchical constraints

We want to **ban inconsistent assignments** at different scales:

multiple

aeroplane bicycle ... train tvmonitor

background

The label of a child region must be 'below' the label of its parent region.

## Second-order (edge) features

Regions with the same label should have similar features:

$$\phi^{\text{edges}}(r_p, r_c; \underbrace{label_p, label_c}_{\text{parent and child labels}}) = -\underbrace{H(label_p, label_c)}_{\text{hierarchical constraint}}\left|P_{r_p,label_p} - P_{r_c,label_c}\right|^2$$

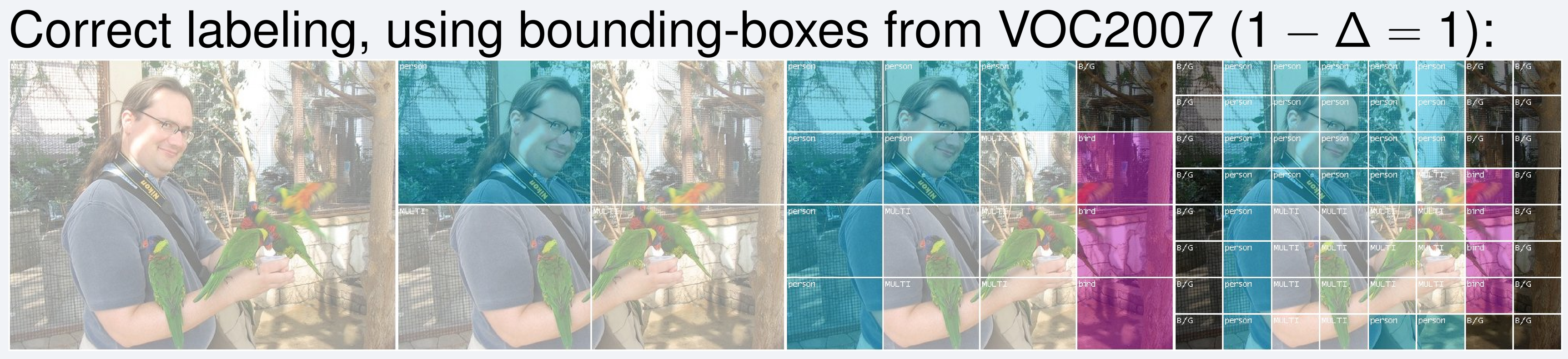
Thus we **learn which classifiers are consistent** across scales

## Structured learning

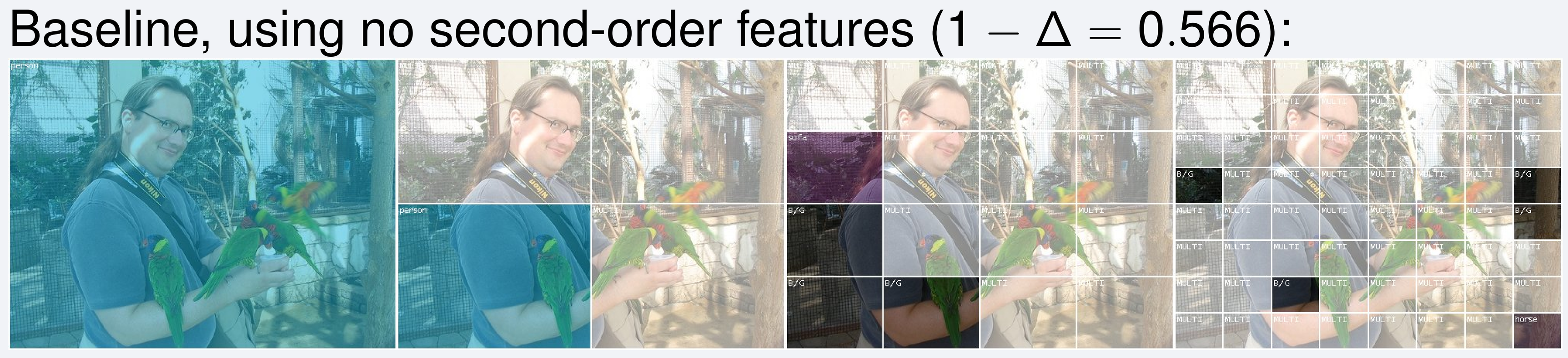
We train our method using *structured learning* [3]. This requires only that we are able to solve the inference problem, and that our *loss function* ($\Delta$) decomposes over the edges in our model, which is certainly true of the Hamming loss.

## Example results

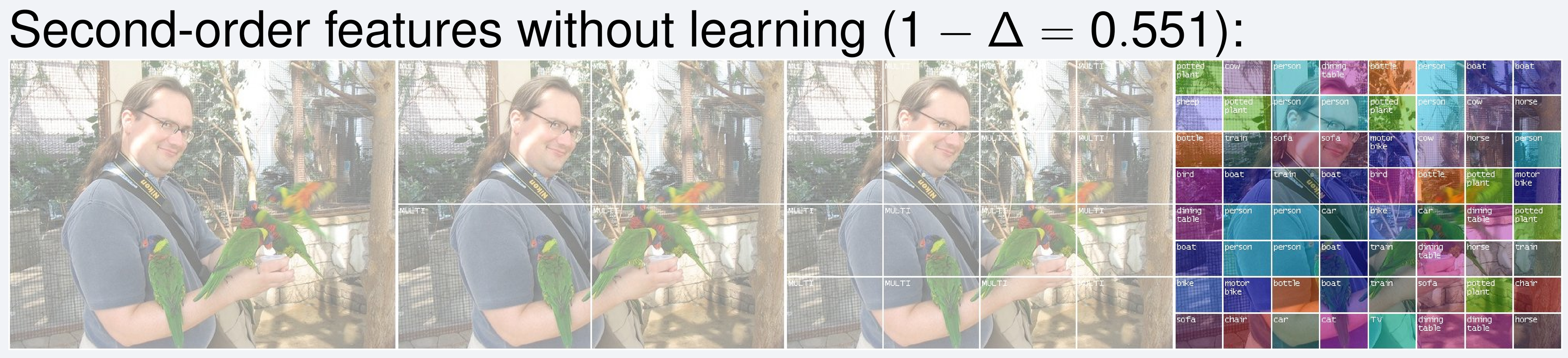
Correct labeling, using bounding-boxes from VOC2007 ($1-\Delta = 1$):



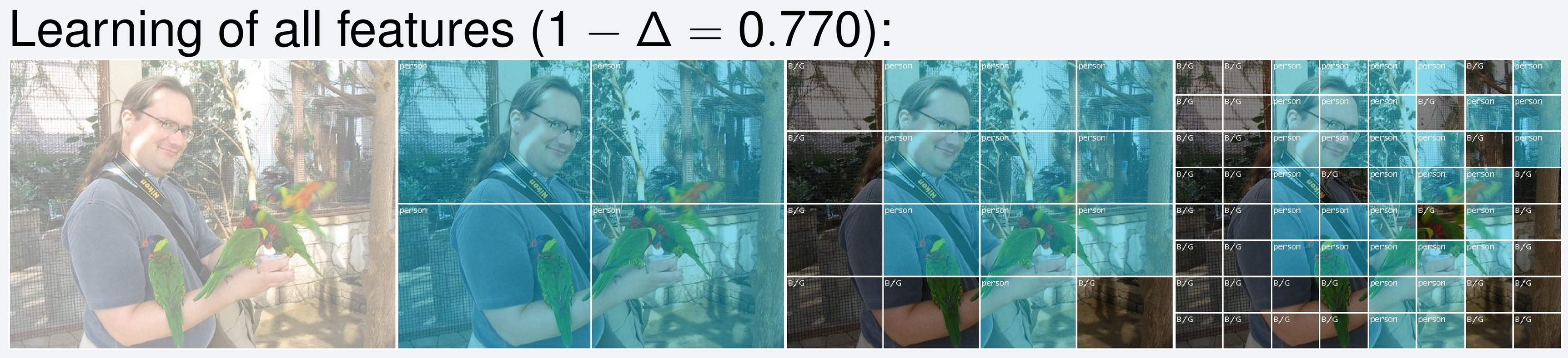
Baseline, using no second-order features ($1-\Delta = 0.566$):



Second-order features without learning ($1-\Delta = 0.551$):

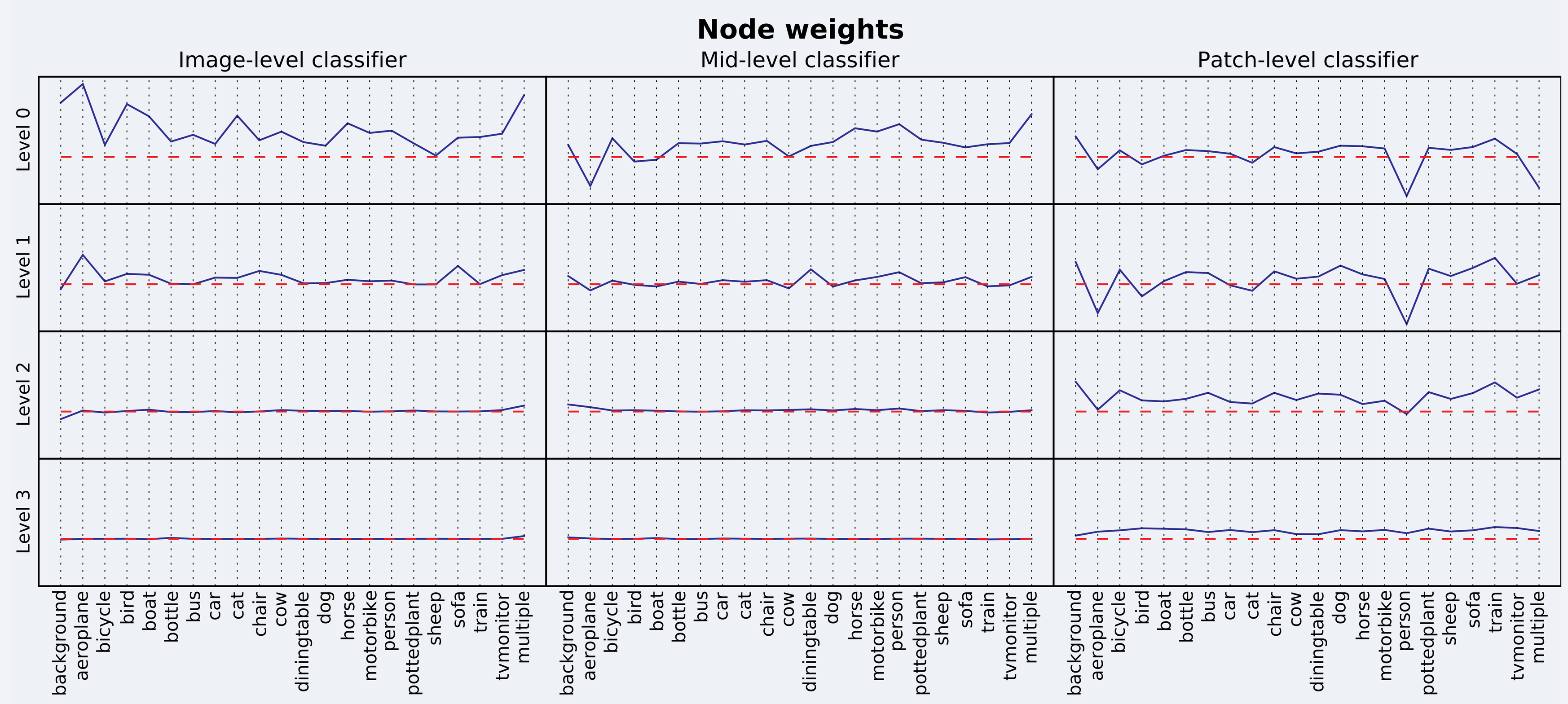


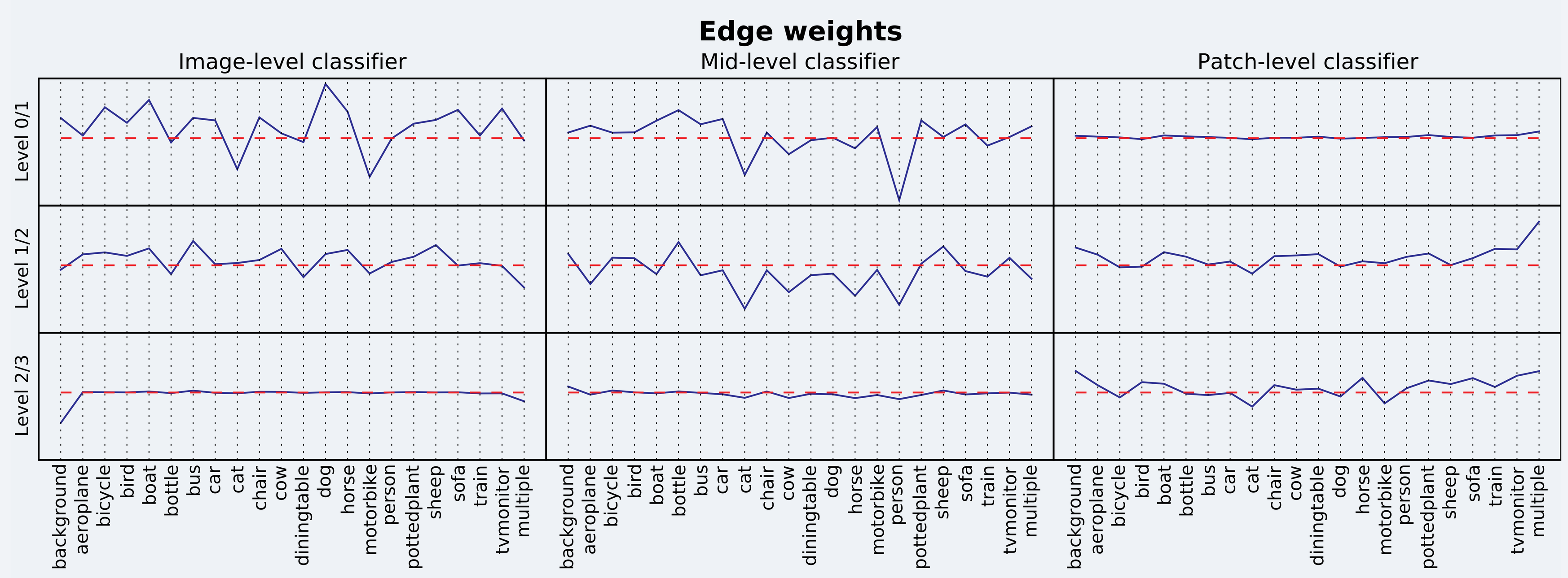
Learning of all features ($1-\Delta = 0.770$):



Colour-code for labels:


## Our learned model




## Bibliography

[1] J. J. McAuley, Teofilo de Campos, Gabriela Csurka, and Florent Perronnin.
Hierarchical image-region labeling via structured learning.
In *BMVC*, 2009.

[2] G. Csurka and F. Perronnin.
A simple high performance approach to semantic segmentation.
In *BMVC*, 2008.

[3] I. Tsochantaridis, T. Hofmann, T. Joachims, and Y. Altun.
Support vector machine learning for interdependent and structured output spaces.
In *Predicting Structured Data*, pages 823–830, 2004.

This poster is available at `http://users.rsise.anu.edu.au/~julianm/`

julian.mcauley@nicta.com.au, t.decampos@st-annes.oxon.org, gabriela.csurka@xrce.xerox.com, florent.perronnin@xrce.xerox.com