Ranking Figure-Ground Hypotheses for Object Segmentation

João Carreira, Fuxin Li, Cristian Sminchisescu

Faculty of Mathematics and Natural Science, INS,
University of Bonn
http://sminchisescu.ins.uni-bonn.de/



Principles

- **Avoid early decision making**. Low-level processes should produce plausible segment hypotheses with sufficiently large, non-local spatial scope
- **Exploit mid-level cues**. Some may require the calculation of features over sufficiently large regions (e.g. parallelism, convexity, orientation)
- **Integrate learning and top-down, class information, into bottom-up calculations progressively**. This may still be feasible within a dominantly feed-forward architecture

Mechanism

- **Multiple figure-ground segment hypotheses** are generated by searching for the breakpoints of constrained min-cut energies, solved at multiple scales on image grid (CPMC)
- We **learn to rank segments**. Ranking uses mid-level, class-independent, visual cues
- **Classification** stage sequentially assigns class labels to segments and resolves conflicts among regions with inconsistent labels

Computational pipeline

Generate multiple object segment hypotheses

Rank object hypotheses using mid-level cues (*Class independent scoring)*

...

...

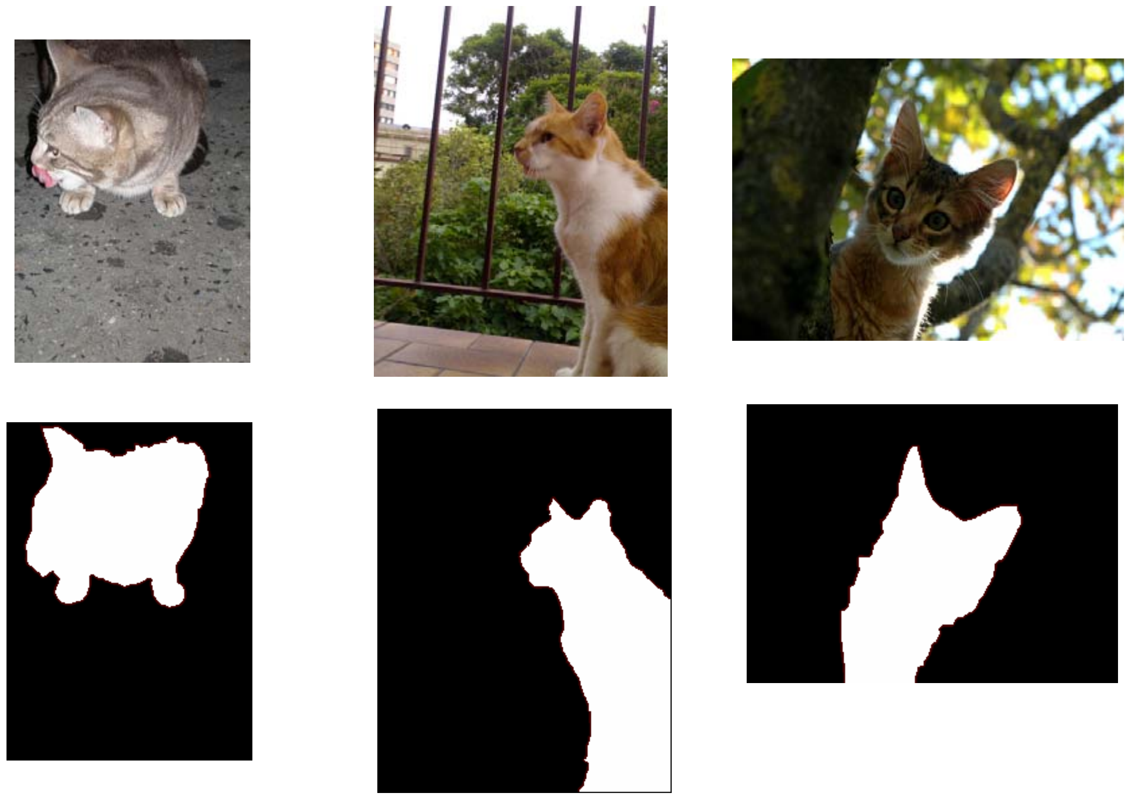
Predict overlap estimate of each segment to specific object class *(1 predictor / class)*

- Select segment/class with highest score
- Consolidate by aggregating multiple high-rank segments with large spatial overlap from the same class
- Add result to final segmentation

Sequentially add segments

Only segmentation labels used for training/testing. No bounding box information/calculation.

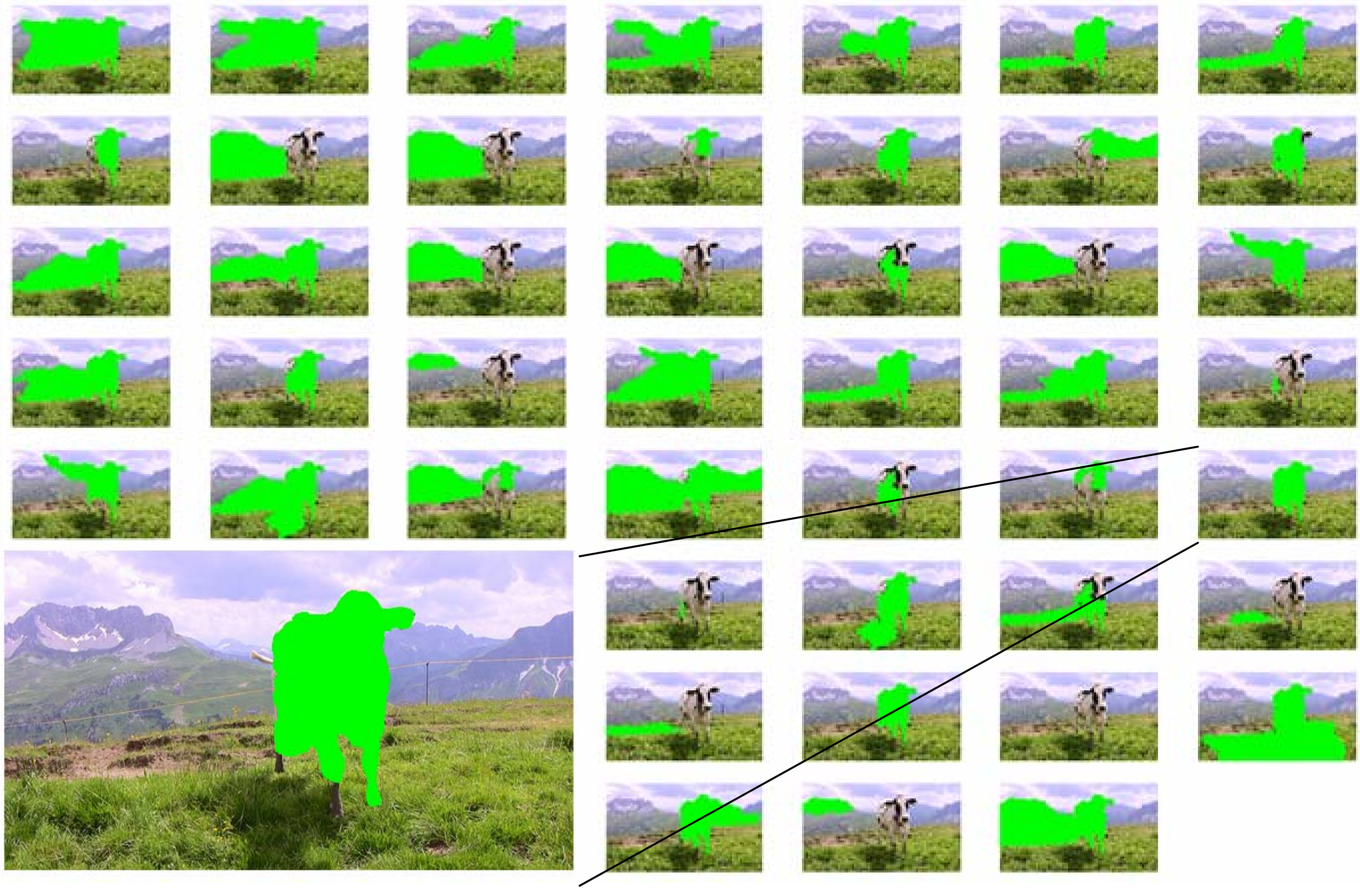
The appeal of bottom-up figure-ground segmentation



Segments obtained by our method

But can accurate results be achieved ?

Instead of committing to one segmentation, generate multiple figure-ground hypotheses

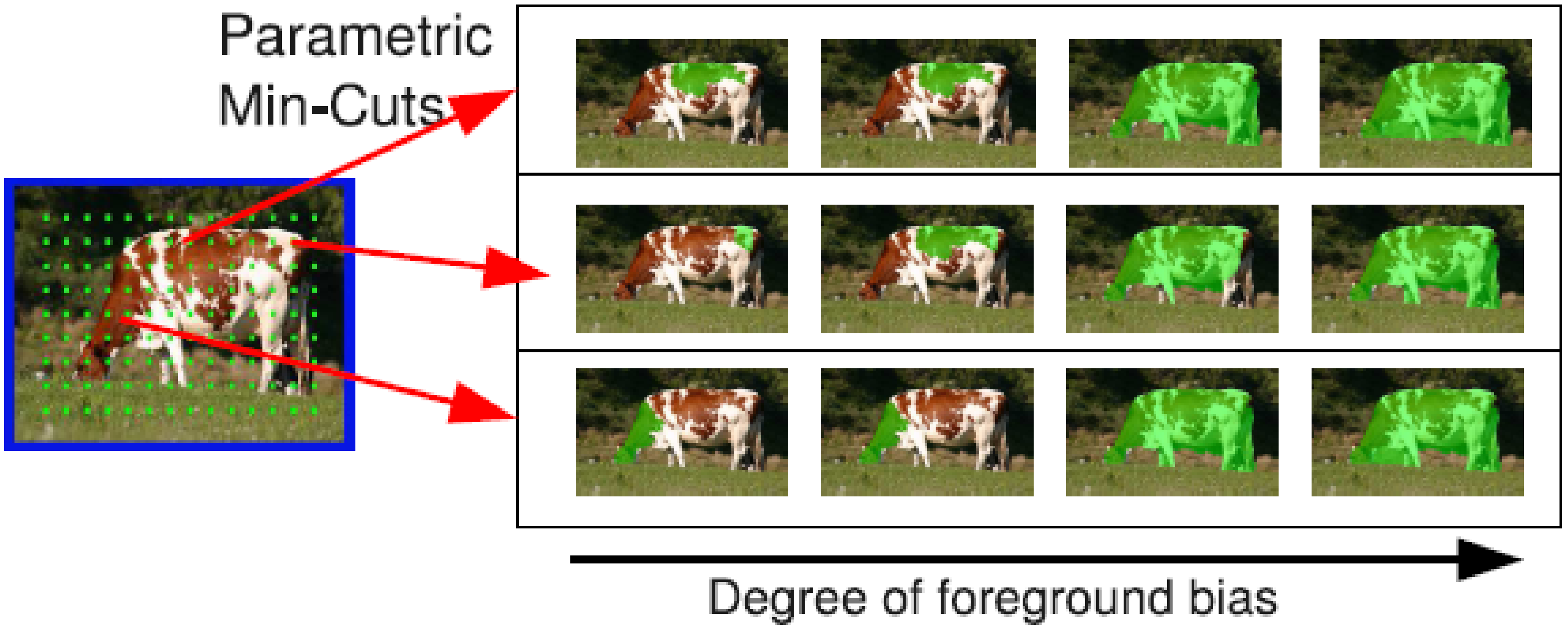


The challenge is to pull out good segments

Constrained Parametric Min-Cuts

$E(A) = \lambda R(A) + B(A)$

Design region term in such a way that λ represents the degree of foreground bias (foreground scale)



• Solve sequence of constrained min-cut problems on regular grid of seeds. Search for all breakpoints using parametric max flow

• Filter solutions with large spatial overlap (small segments co-exist, notice that the method *does not* particularly favor large segments)

J. Carreira and C. Sminchisescu: Image Segmentation with Constrained Parametric Minimum Cuts, Technical Report, University of Bonn, November 2009 (upcoming).

Computational pipeline

Generate multiple object segment hypotheses

Rank object hypotheses using mid-level cues (*Class independent scoring)*

...

...

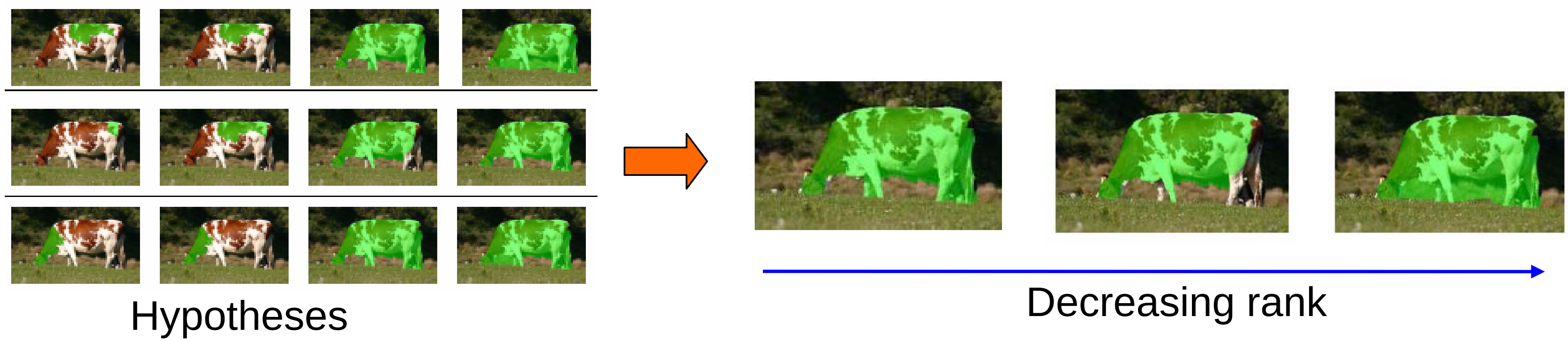
Predict overlap estimate of each segment to specific object class *(1 predictor / class)*

- Select segment/class with highest score
- Consolidate by aggregating multiple high-rank segments with large spatial overlap from the same class
- Add result to final segmentation

Sequentially add segments

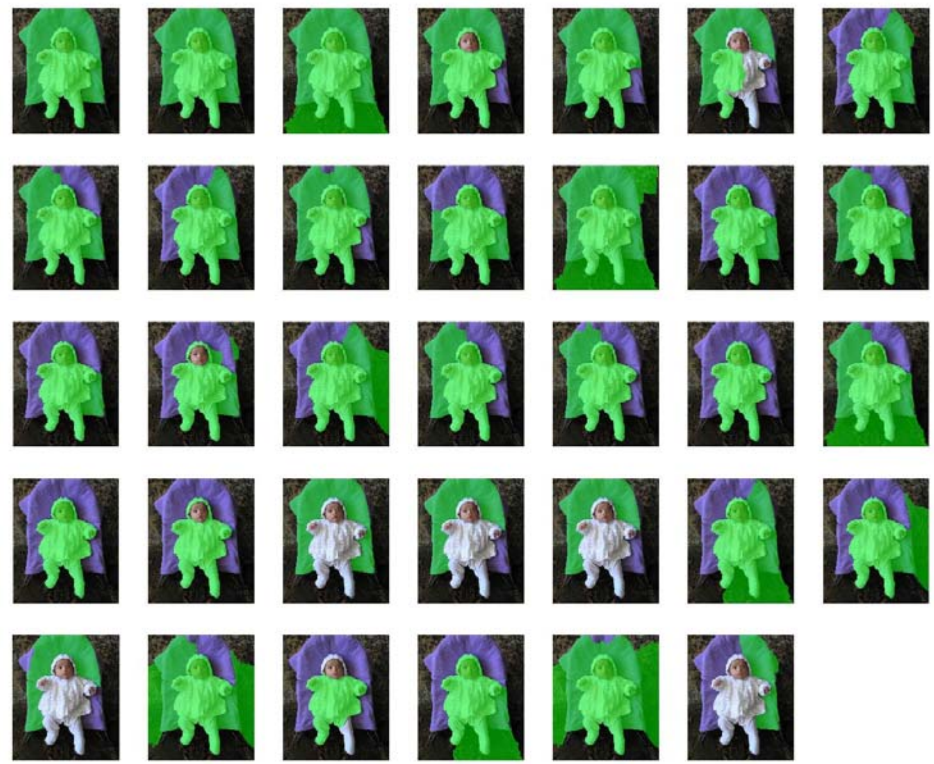
Only segmentation labeling used for training/testing. No bounding box information/calculation.

Ranking figure-ground hypotheses



- Hypothesized segments ranked using regression
- Ranking is class–independent (mid-level)
- Features (~2500)
  - *Boundary* – cut, ratio cut, normalized cut
  - *Region* – location, perimeter, area, Euler number, orientation
  - *Gestalt* – convexity, smoothness, symmetry
  - *Appearance/Shape* –BOW, HOG

Highest ranked segment hypotheses



We learn to discard homogenous `non object-like’ segments

Computational pipeline

→ Generate multiple object segment hypotheses

→ Rank object hypotheses using mid-level cues (*Class independent scoring)*

→ …

… Predict overlap estimate of each segment to specific object class *(1 predictor / class)*

→
- Select segment/class with highest score
- Consolidate by aggregating multiple high-rank segments with large spatial overlap from the same class
- Add result to final segmentation

Sequentially add segments

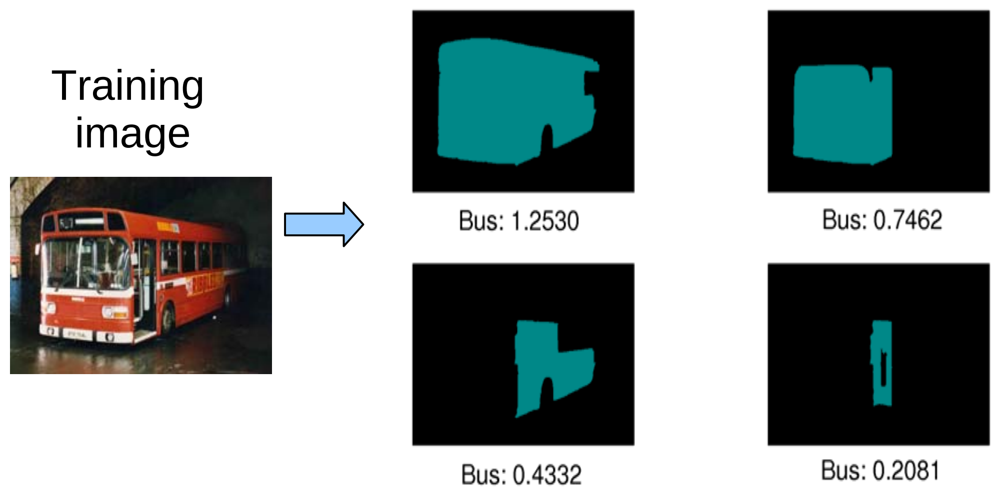
→

Only segmentation labels used for training/testing. No bounding box information/calculation.

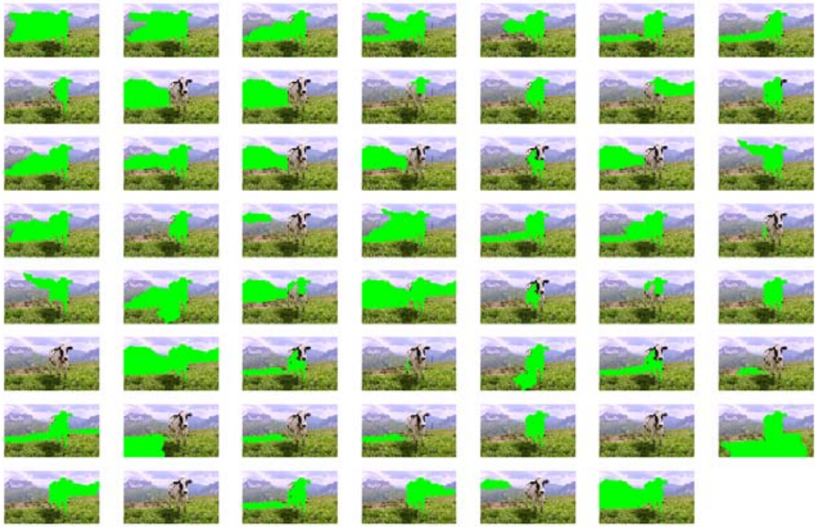
Prediction of segment overlap to class-object

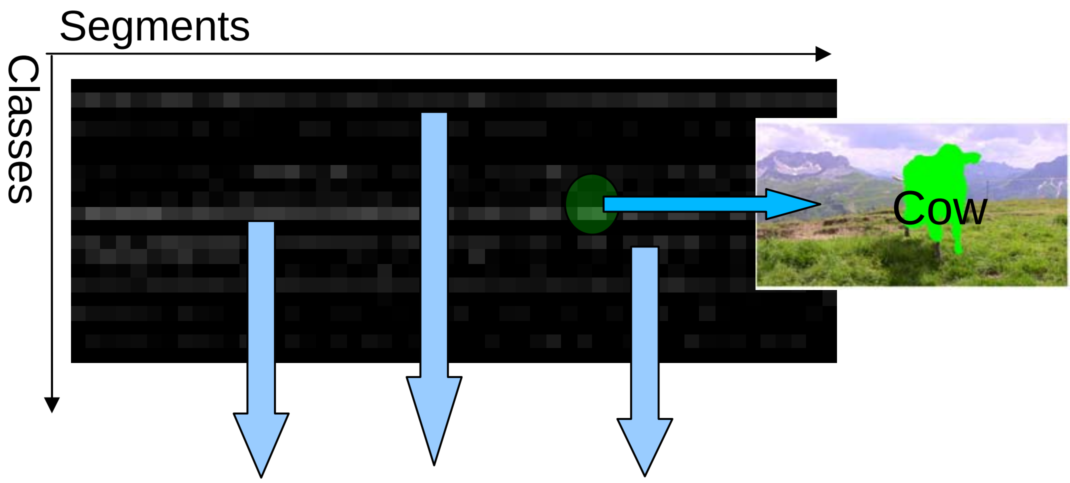
- Support Vector Regression framework
- Each regressor is trained to estimate overlap of foreground segment with object of its class
- Trained with all generated figure segments. For each class, the output training scores of segments that only overlap with other classes (or background) are set to 0


Dataset has plenty of partial views



Prediction of segment overlap to each object class





- One predictor for all visual aspects of each class, evaluates only regions that can plausibly contain objects of interest
- Shape and appearance features on both segment/foreground and background (contour and internal BOW SC, HOG,CSIFT)
- MKL regression framework (8 kernels)

Computational pipeline

→ Generate multiple object segment hypotheses

→ Rank object hypotheses using mid-level cues *(Class independent scoring)* → …

… Predict overlap estimate of each segment to specific object class *(1 predictor / class)*

→

- Select segment/class with highest score
- Consolidate by aggregating multiple high-rank segments with large spatial overlap from the same class
- Add result to final segmentation

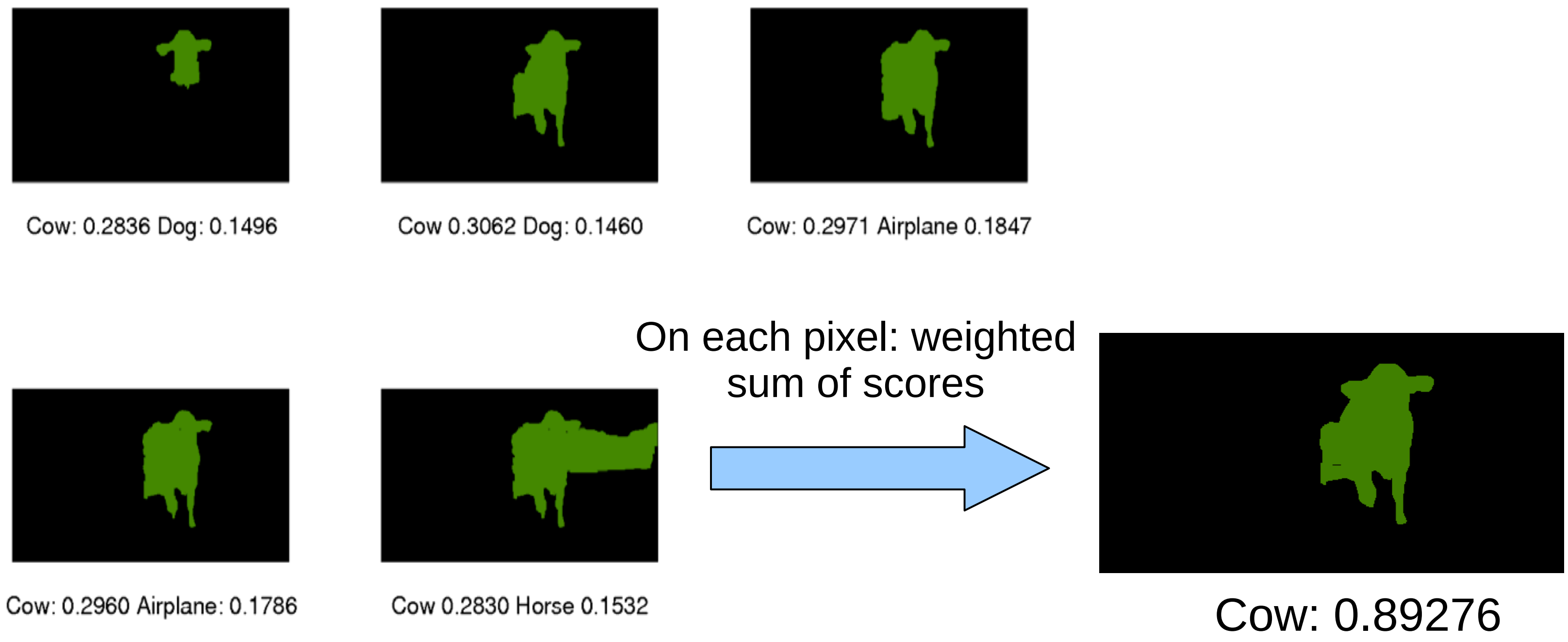
Sequentially add segments

→

Only segmentation labels used for training/testing. No bounding box information/calculation.

Increasing segmentation robustness

- Find k other segments with largest overlap, from the class of top-scoring segment
- Generate final solution by weighted combination



Sequential segment classification

Segment #1:
TV/Monitor 1.6037

Segment #2:
TV/Monitor 1.7132

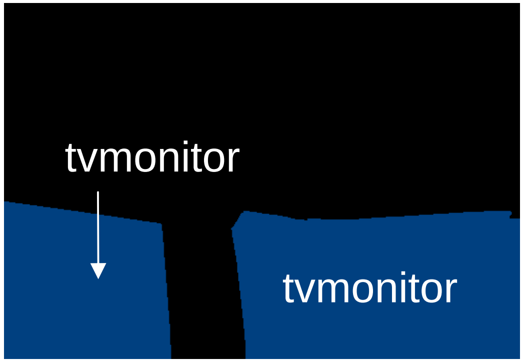
Segment #3:
TV/Monitor 1.3615

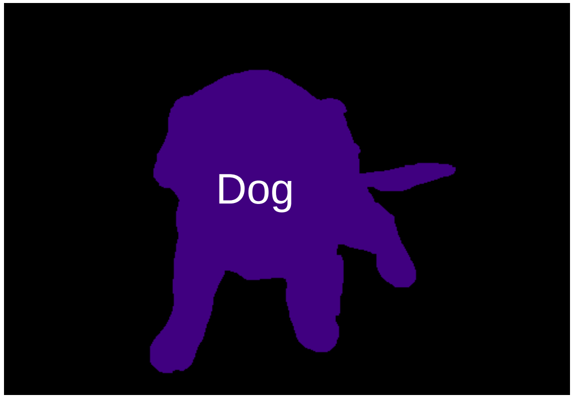
Segment #4:
Chair 0.50904

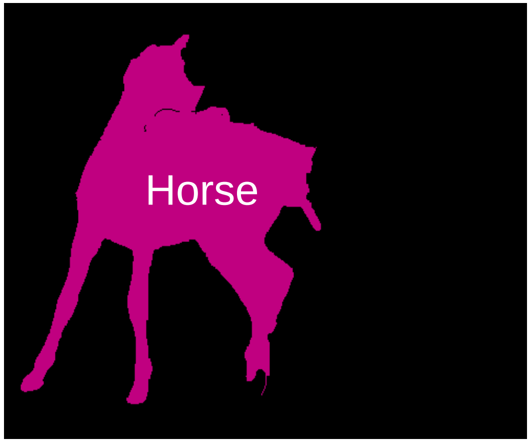
Learned threshold

After combination

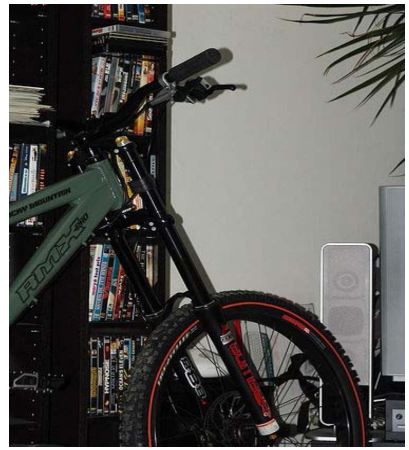
Results: success stories

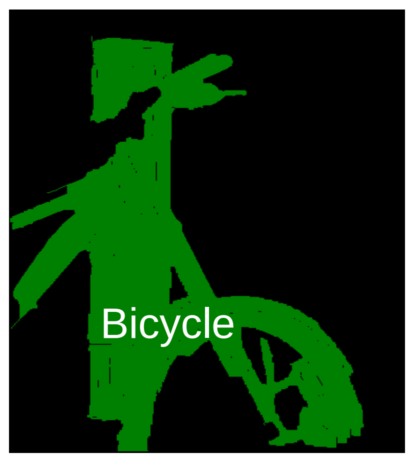


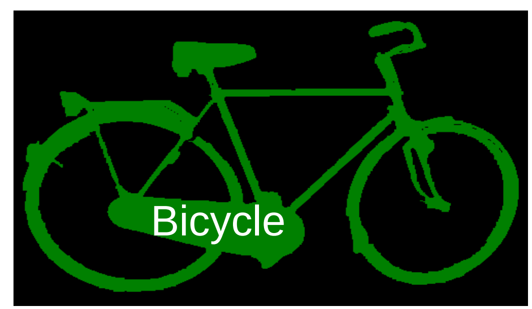




Success stories –
bikes and motorbikes

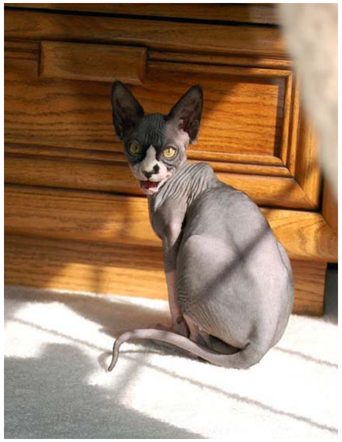
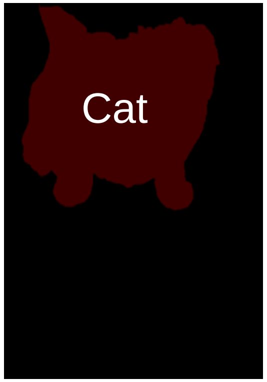


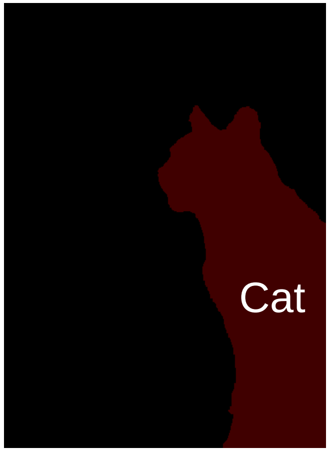


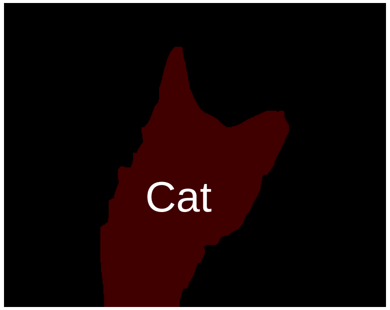


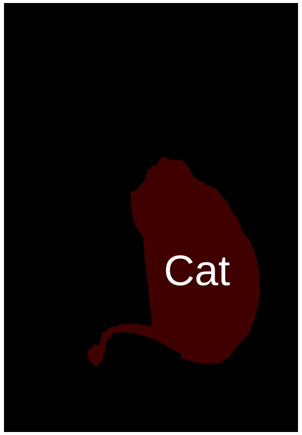
Success stories - cats



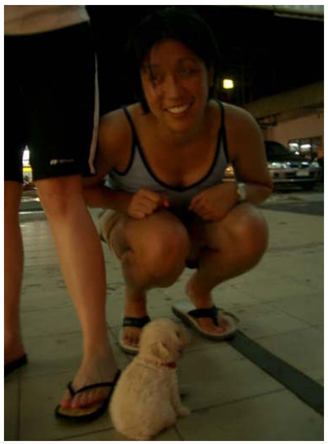




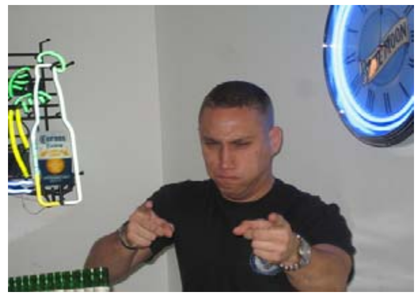


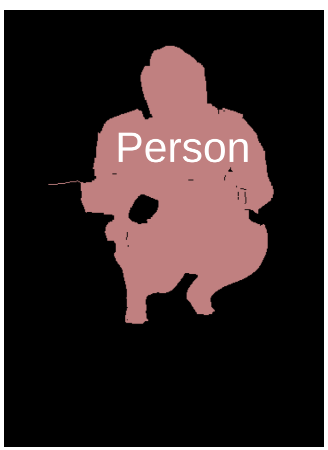


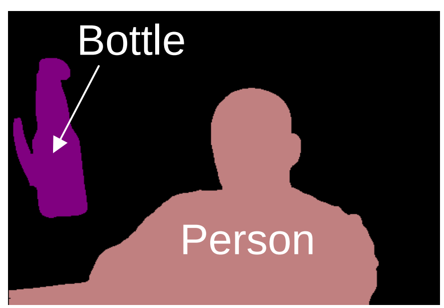


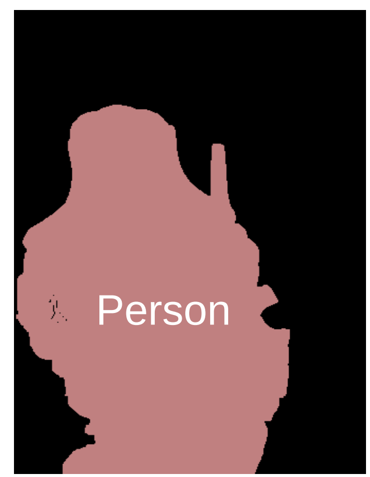
Success stories - people



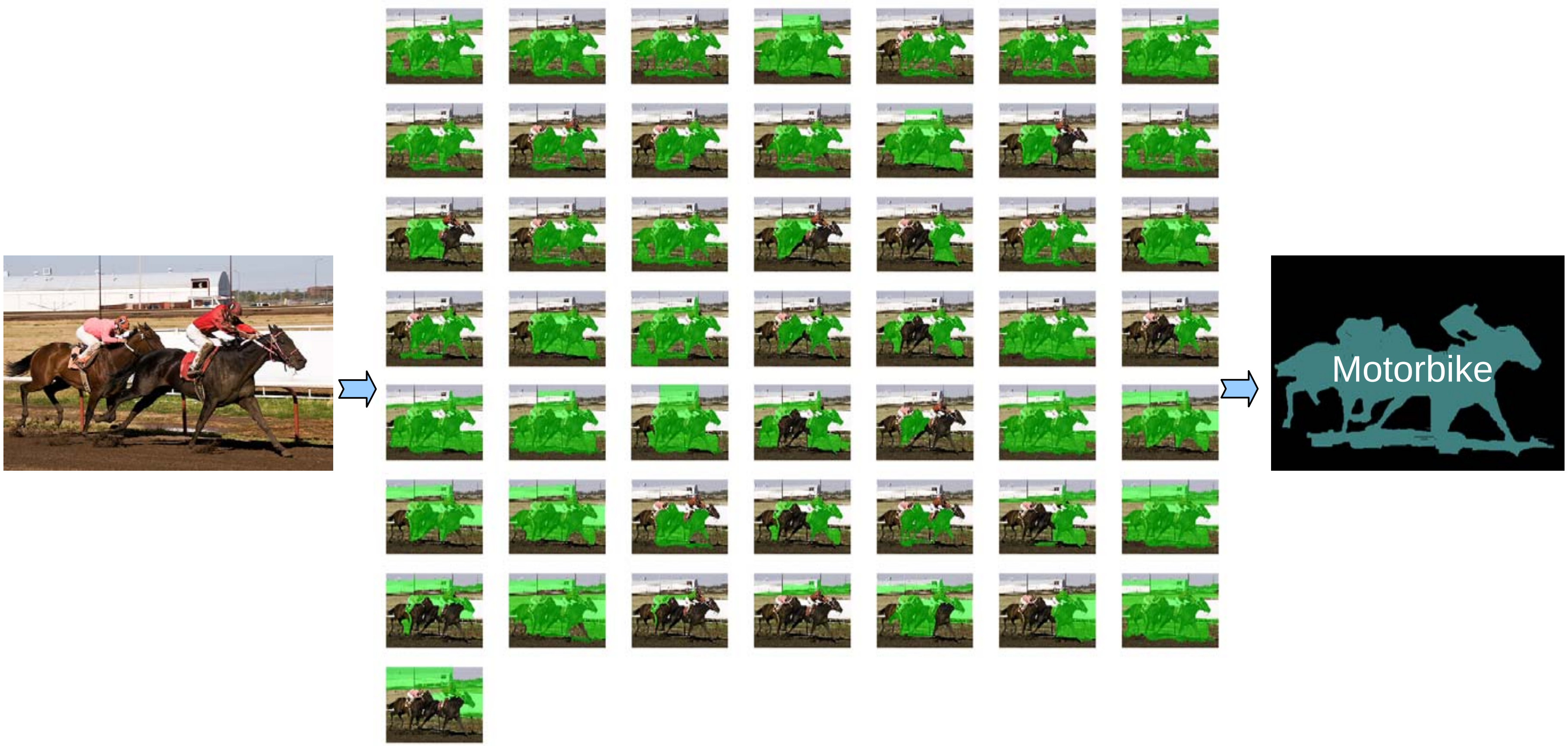




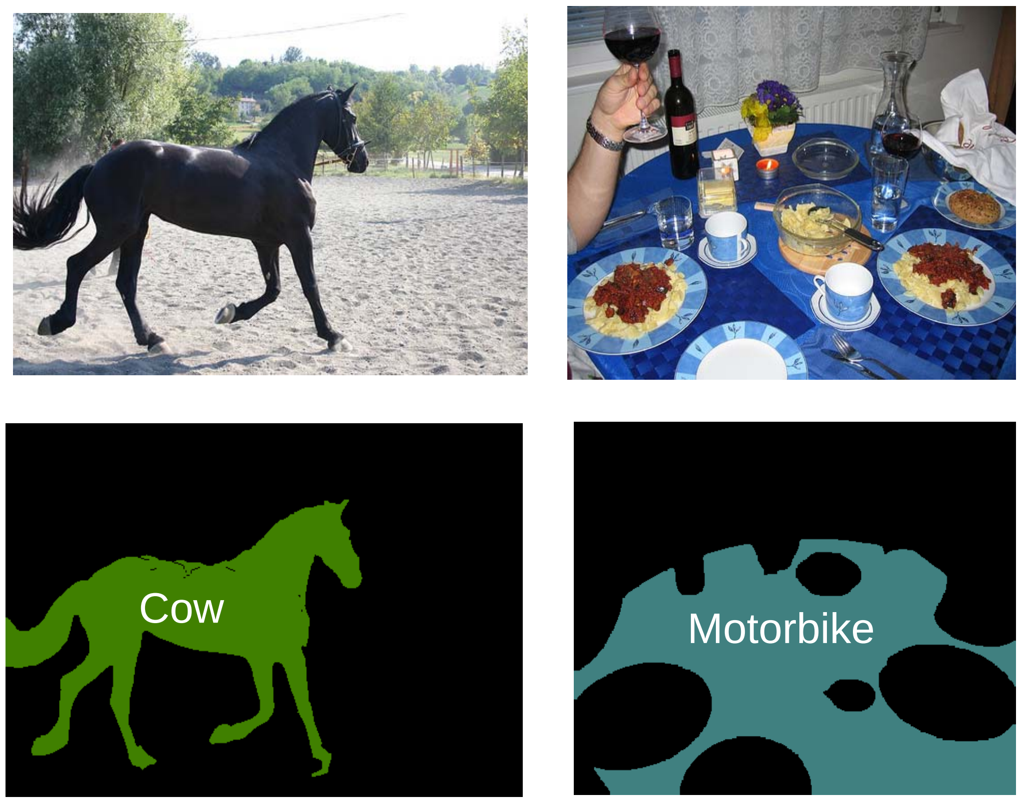




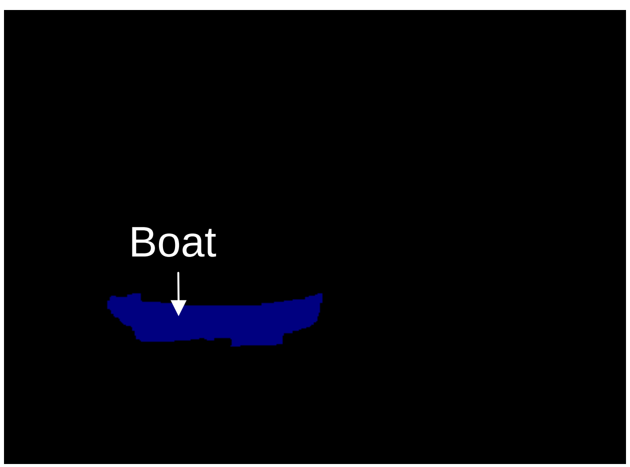
Failure modes - wrong segments

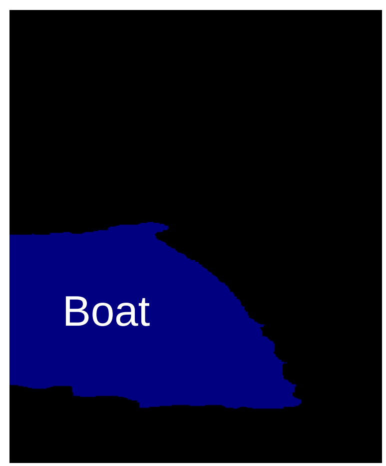


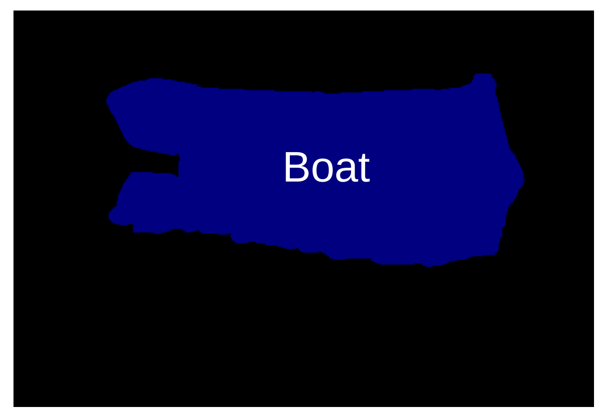
Failure modes - wrong classification



A peculiarity - reflections







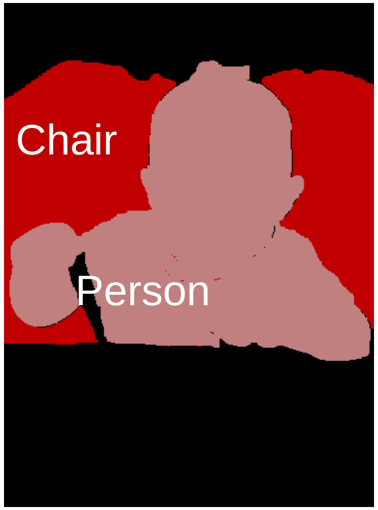
Discussion

Pros

- Small number of classification decisions
- Can use global object shape features
- Context as additional feature
- One regressor for all visual aspects of each category
- Learn partial object views from full views

Cons

- Reliant on reasonable segmentability
- May learn "intertwined" object classes – people on horses and bikes. Depending on the goal, this can be a feature…



Conclusions

- Segmentation-recognition pipeline
  - Ranking and sequential classification of multiple segmentation hypotheses, generated using Constrained Parametric Minimum Cuts (CPMC)
- Winning segmentation entry in VOC 2009
- Future work
  - Integrate information from bounding box detectors
  - Closer integration of learning in the segmentation-recognition loop