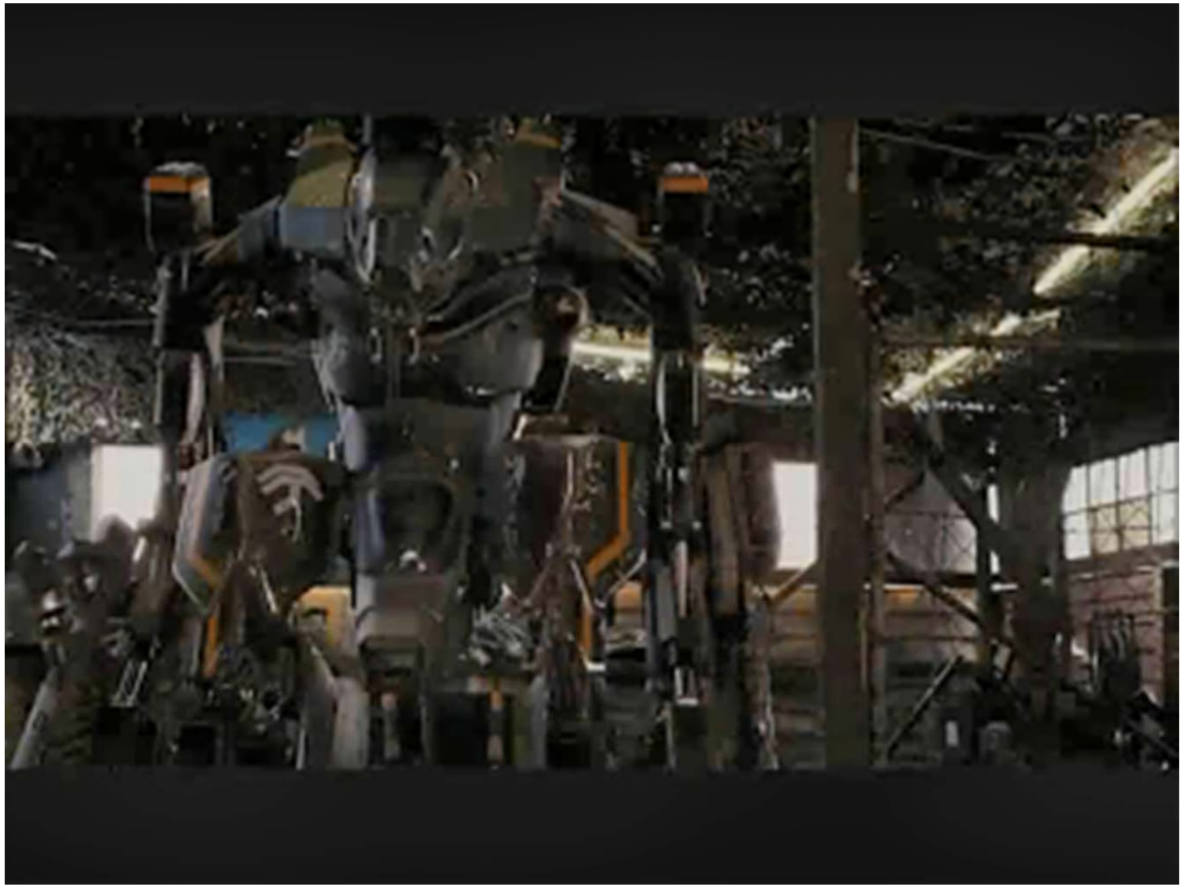
Object Recognition by Sequential Ranking of CPMC Segments

João Carreira[1], Fuxin Li[2,1], Cristian Sminchisescu[1]

[1]University of Bonn, [2]Georgia Institute of Technology

Pascal VOC Challenge 2011 Workshop in Barcelona

Semantic segmentation, alien implementation (from “District 9”)



Principles

- **Avoid early decision making.** Bottom-up processes should produce plausible object segmentation hypotheses with non-local spatial scope, beyond super-pixels
- **Exploit mid-level cues.** Calculate features over sufficiently large regions, e.g. shape, convexity, orientation
- **Take advantage of information in segments covering both entire objects and parts, for recognition**

Mechanism: *Sliding Segments*

- **Multiple figure-ground segmentations** generated by searching breakpoints of constrained min-cut energies, at multiple locations and scales on image grid (CPMC)
- **Plausible object segments** are selected after ranking and diversifying the low-level segmentations based on mid-level, class-independent, visual cues
- **Recognition stage** detects objects from the multiple categories and sequentially resolves inconsistencies

Computational pipeline

Generate multiple object segment hypotheses

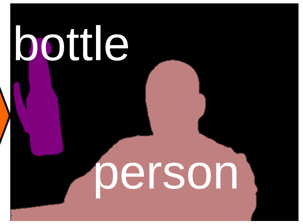
Rank object hypotheses *(Class independent scoring)*

...

..

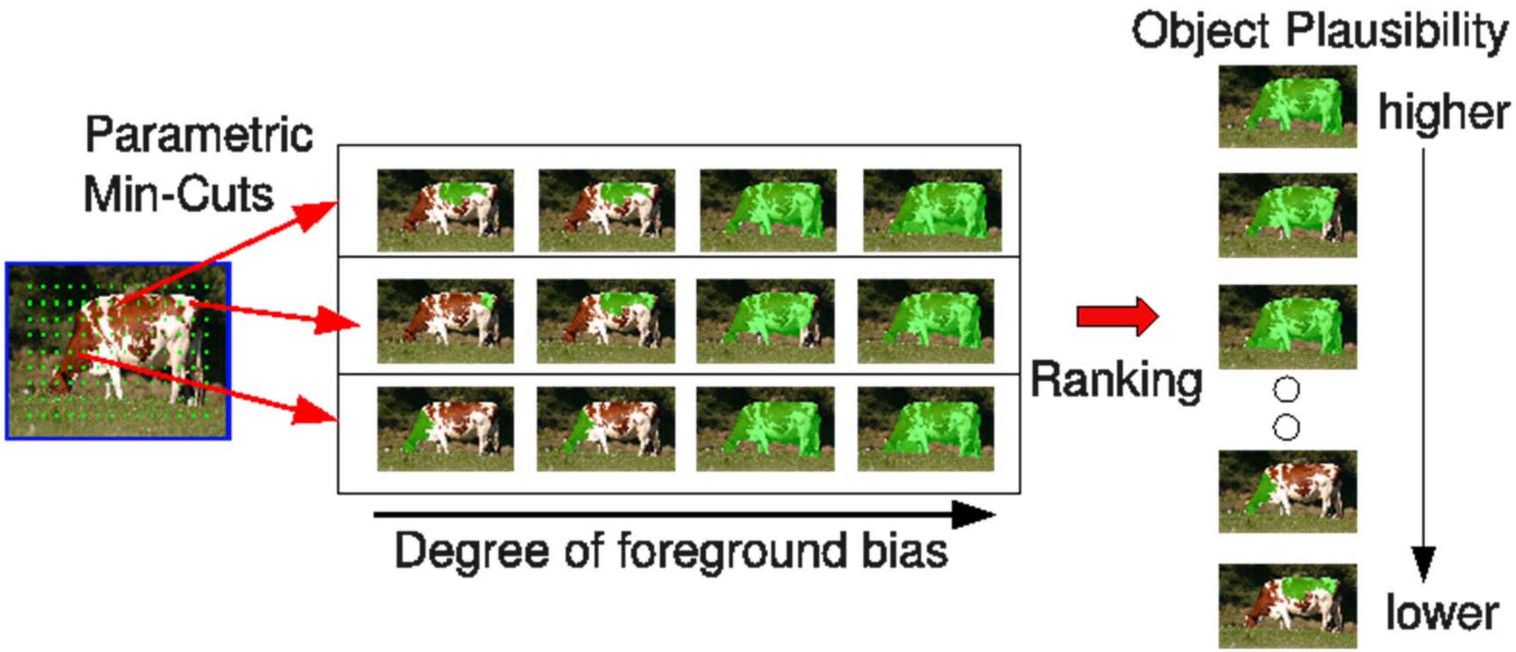
1-against-all class-overlap estimation (instead of classification) of segments

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



Sequentially add segments

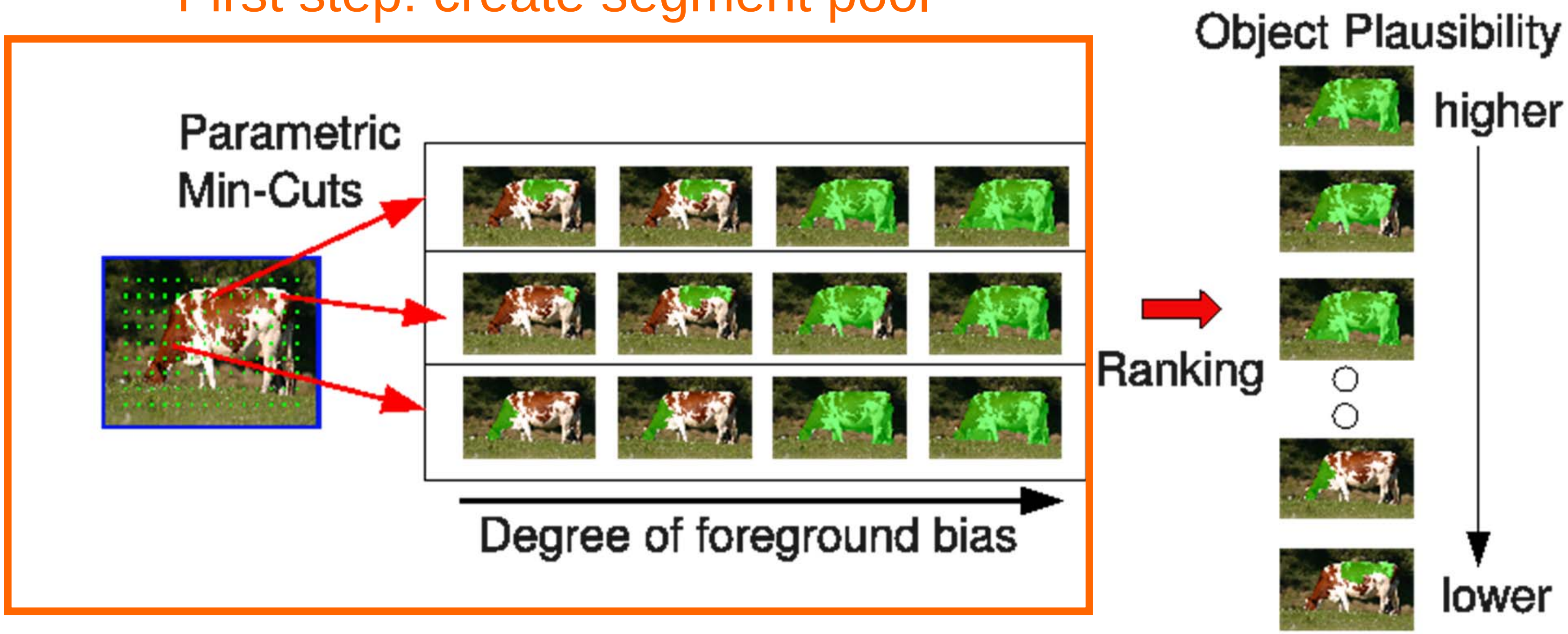
CPMC: Constrained Parametric Min-Cuts for Automatic Object Segmentation



Carreira and Sminchisescu. CVPR 10, PAMI 11

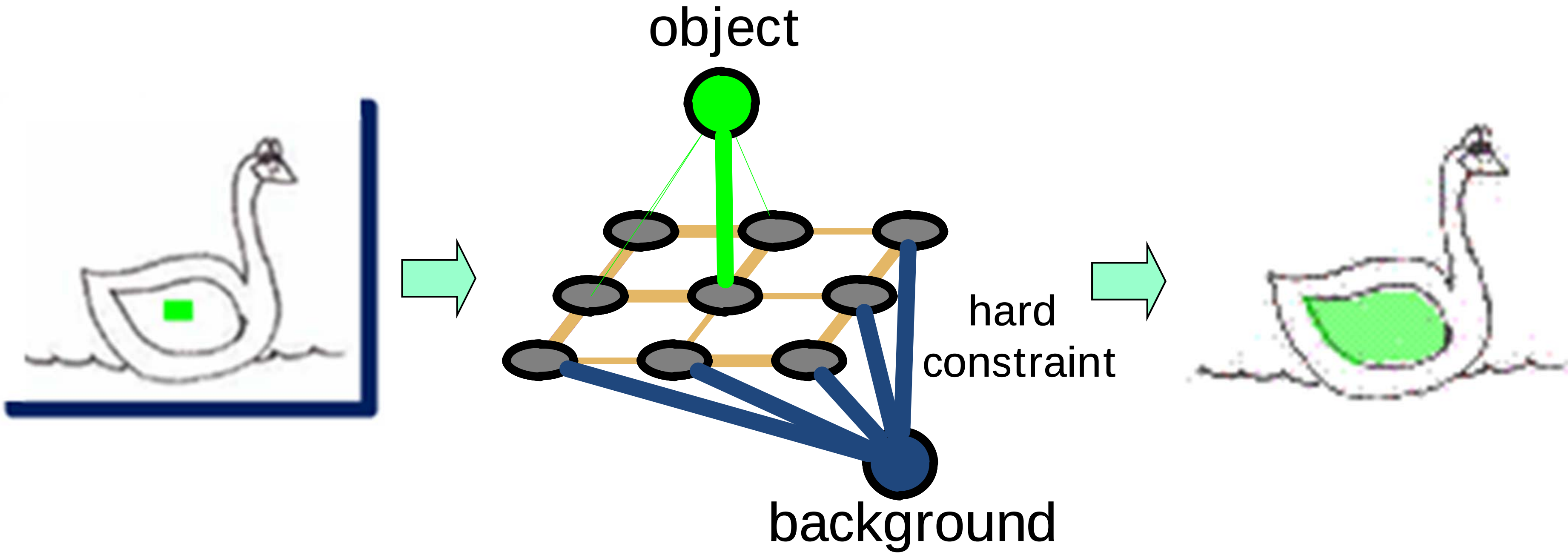
CPMC: Constrained Parametric Min-Cuts for Automatic Object Segmentation

First step: create segment pool



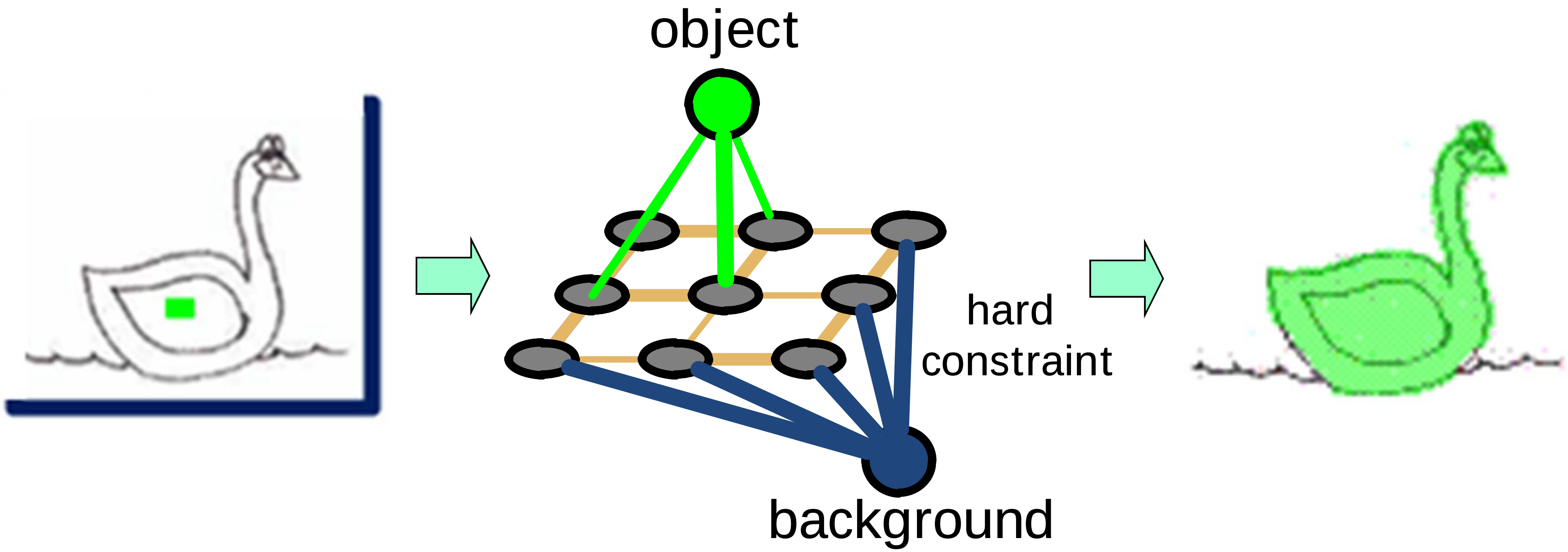
Generating a segment pool: constrained **parametric** min-cut

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



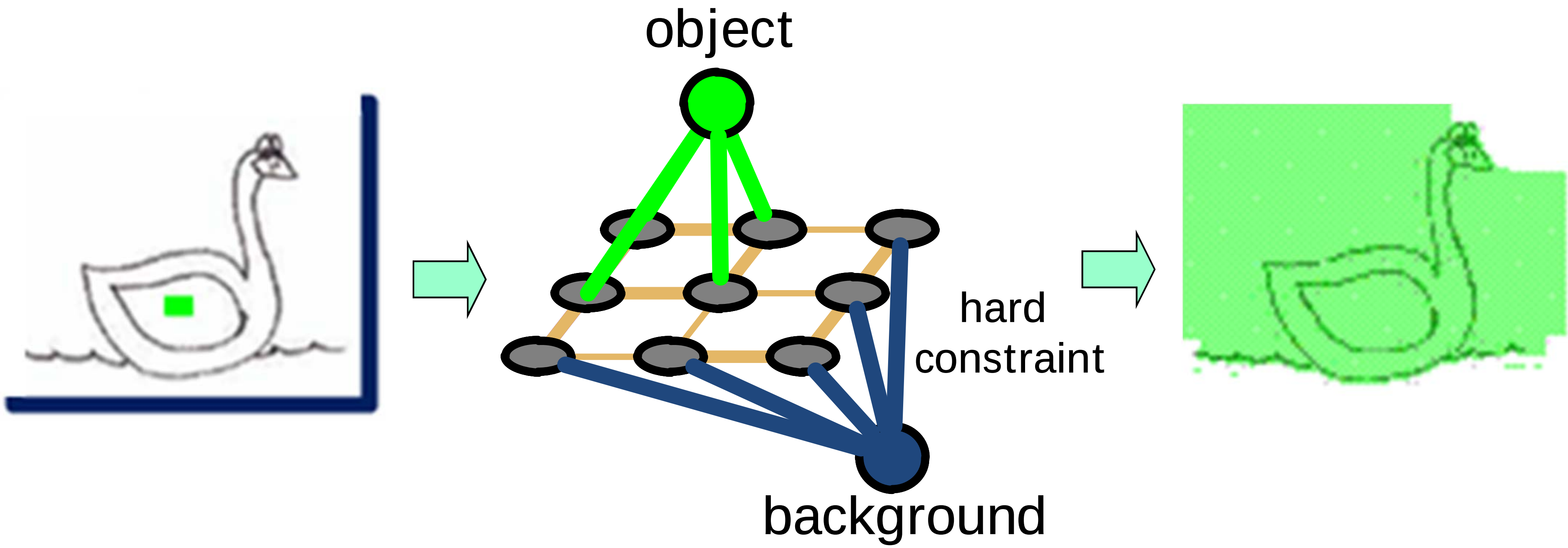
Generating a segment pool: constrained **parametric** min-cut

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

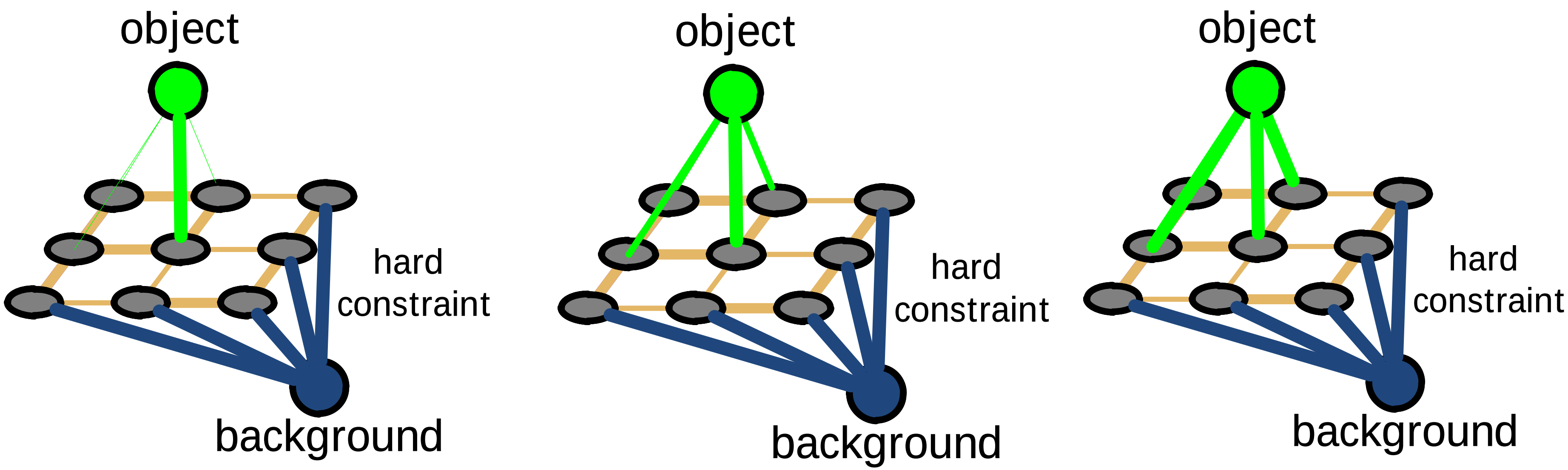


Generating a segment pool: constrained **parametric** min-cut

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



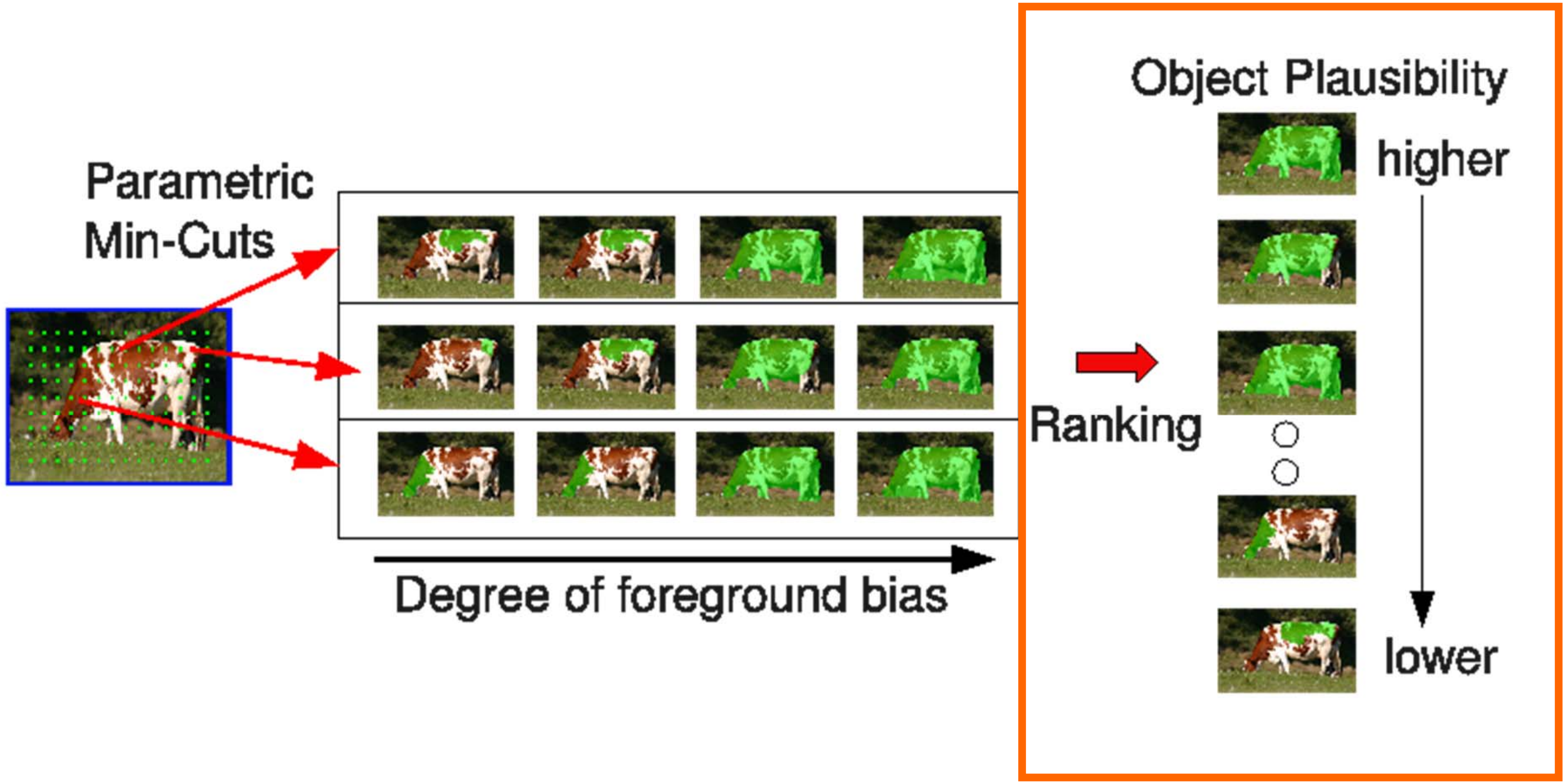
Generating a segment pool: constrained **parametric** min-cut



Can solve for all values of object bias in the same time complexity of solving a single min-cut using a **parametric max-flow solver**

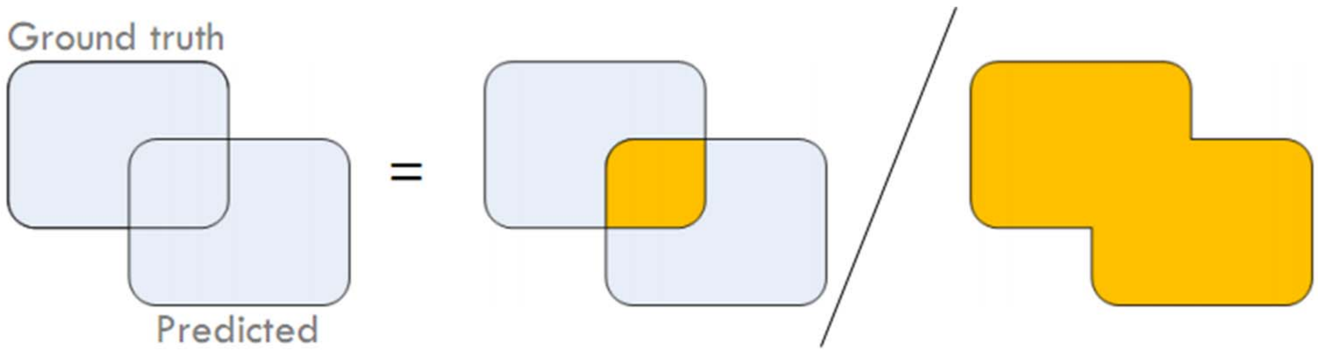
CPMC: Constrained Parametric Min-Cuts for Automatic Object Segmentation

Second step: ranking object hypotheses



How to model segment quality ?

Best **overlap** with a ground truth object computed by intersection-over-union



Ranking object hypotheses

- Aims to handle full object segments and fragments
- Modeled as regression on overlap
- Features
  - *Boundary* – normalized boundary energy
  - *Region* – location, perimeter, area, Euler number, orientation, contrast with background
  - *Gestalt* – convexity, smoothness

Good

High boundary energy

Smooth.

Euler number = 0

Bad

Low boundary energy

Non smooth.

High Euler number

Diversifying the Ranking

Segment Ranking using Maximum Marginal Relevance

$$MMR = \underset{H_i \in H \setminus H_p}{\operatorname{argmax}} \left[\theta \cdot s(H_i) - (1-\theta) \cdot \max_{H_j \in H_p} o(H_i, H_j))\right]$$

Segment score

Pairwise segment overlap

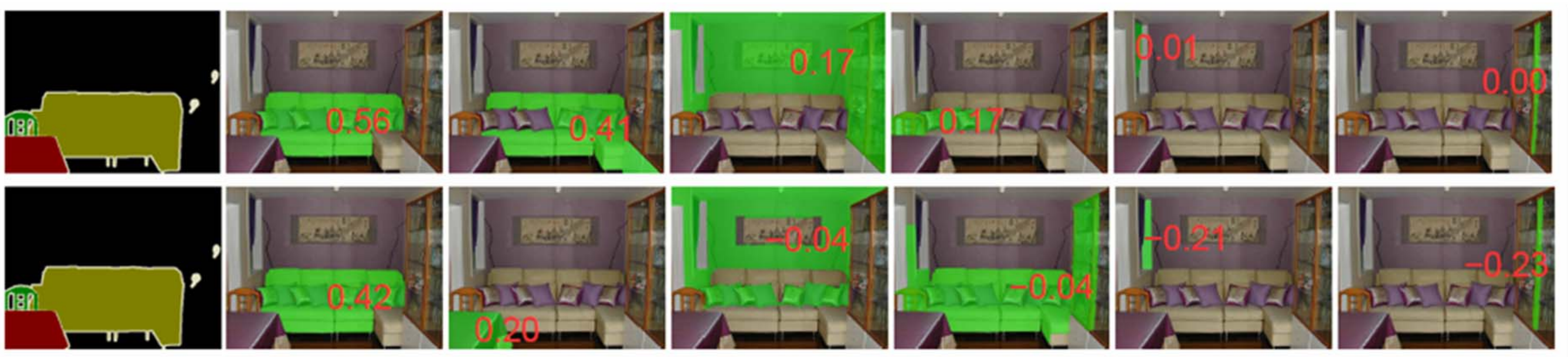
Best two hypotheses

Middle two hypotheses

Worst two hypotheses

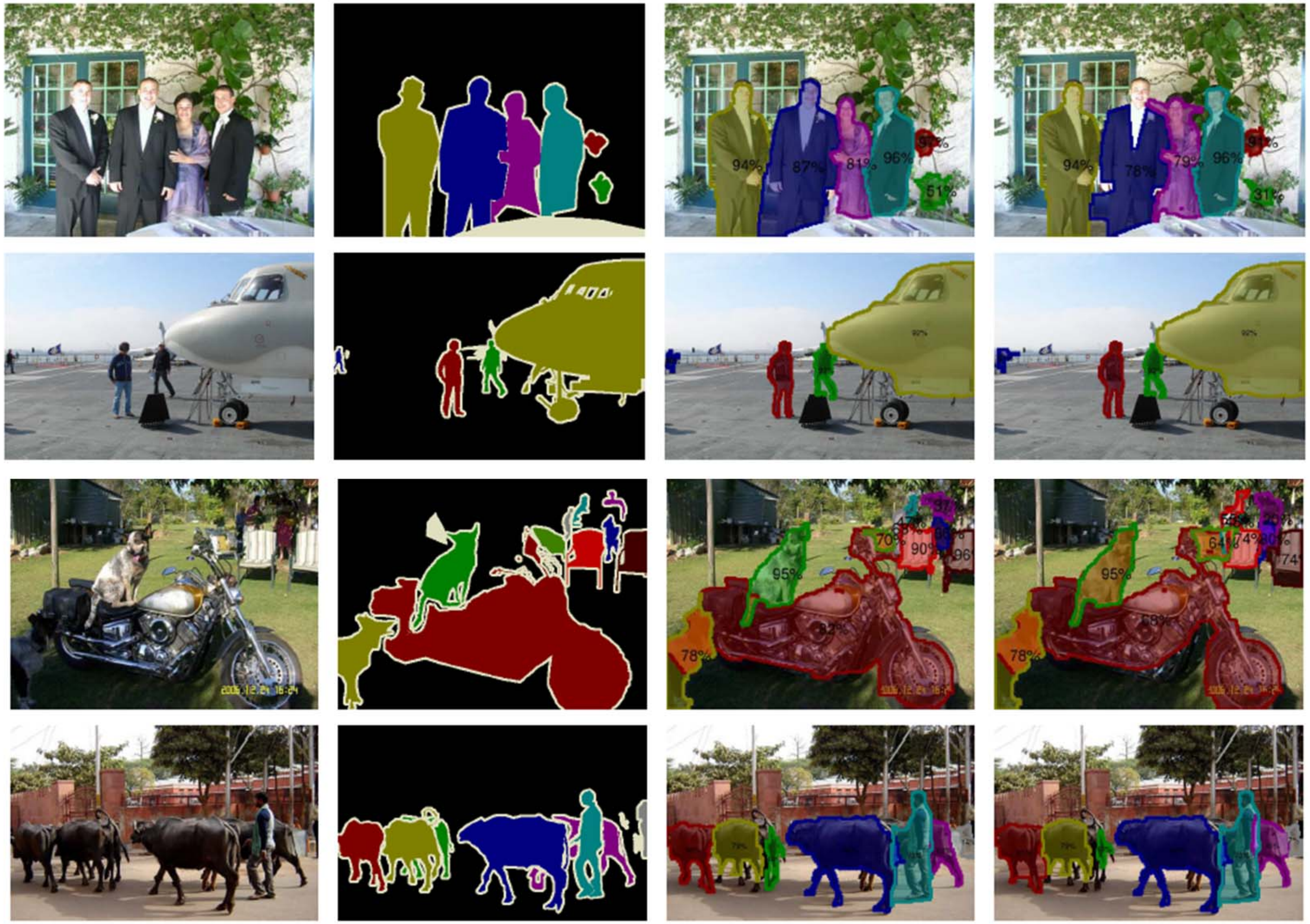
Original

Diversified



Segmentation Examples

Image	Ground-truth Objects	Best in segment pool	Best in top-200



CPMC Segment Generation (video)

Code at http://sminchisescu.ins.uni-bonn.de/cpmc

CPMC Segment Generation (video)

Code at http://sminchisescu.ins.uni-bonn.de/cpmc

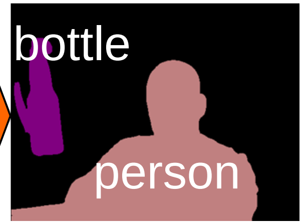
Computational pipeline

Generate multiple object segment hypotheses

Rank object hypotheses *(Class independent scoring)* ...

.. 1-against-all class-overlap estimation (instead of classification) of segments

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

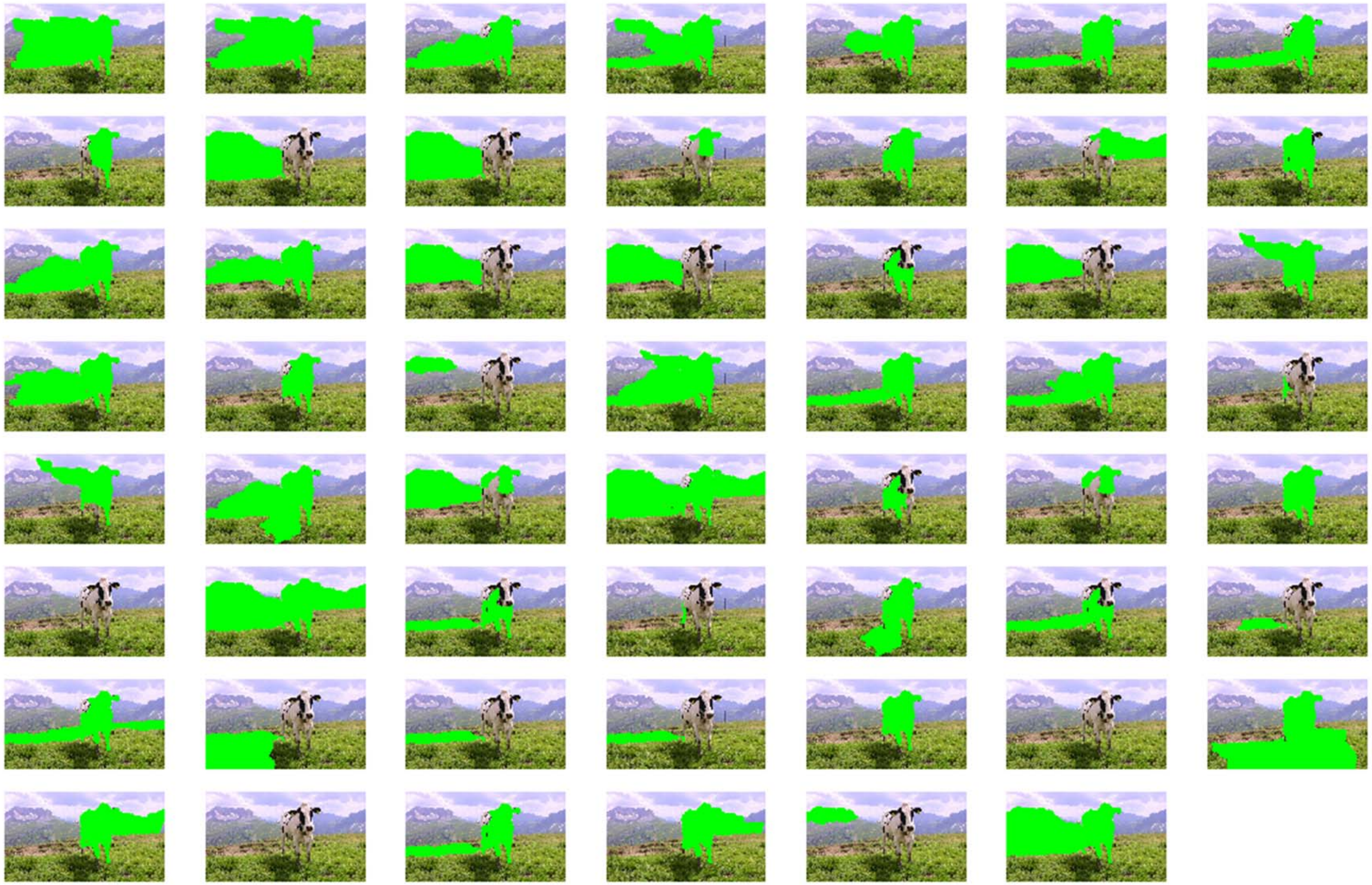


Sequentially add segments

Sliding-Segment Recognition

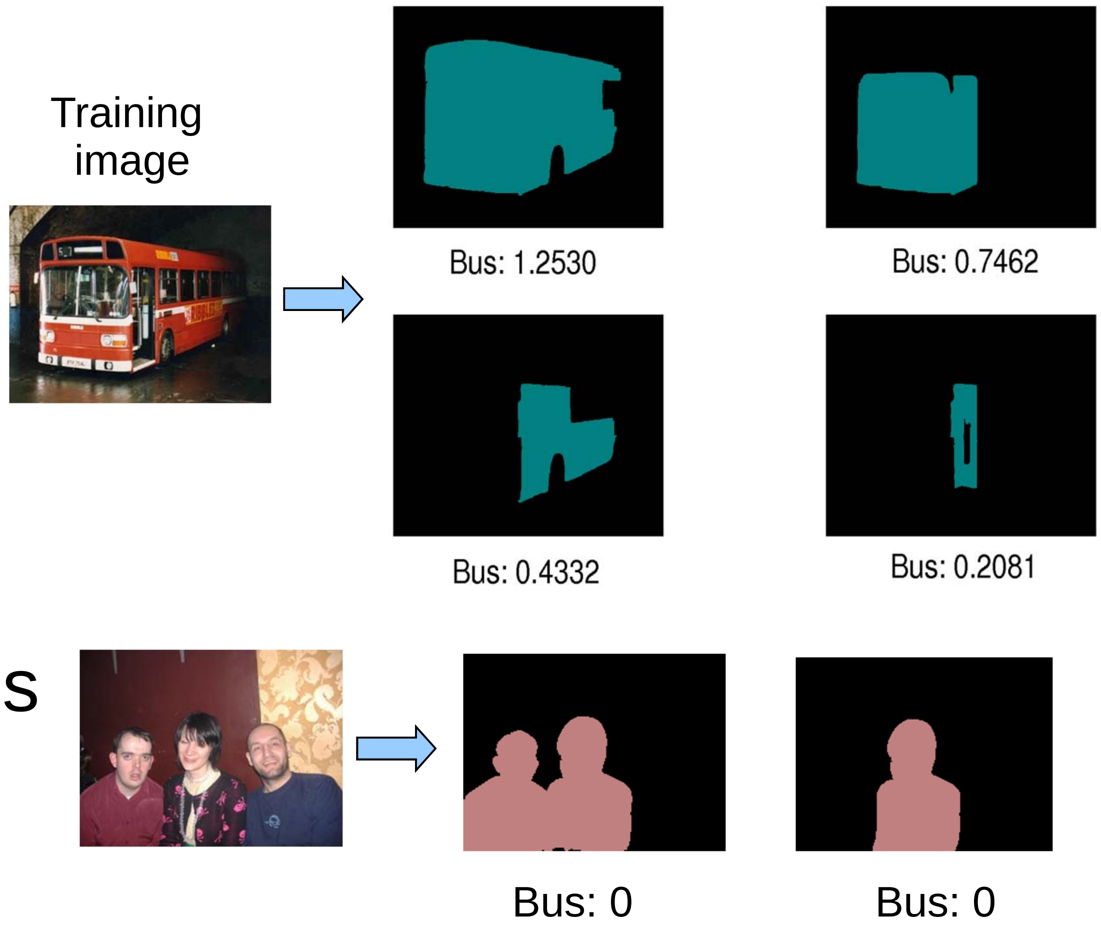
- Each segment categorized by individual class predictors
- Sequential strategy removes inconsistencies and consolidates masks (both spatial support and labels)

Sliding Segments



Prediction of Class-Specific Overlap

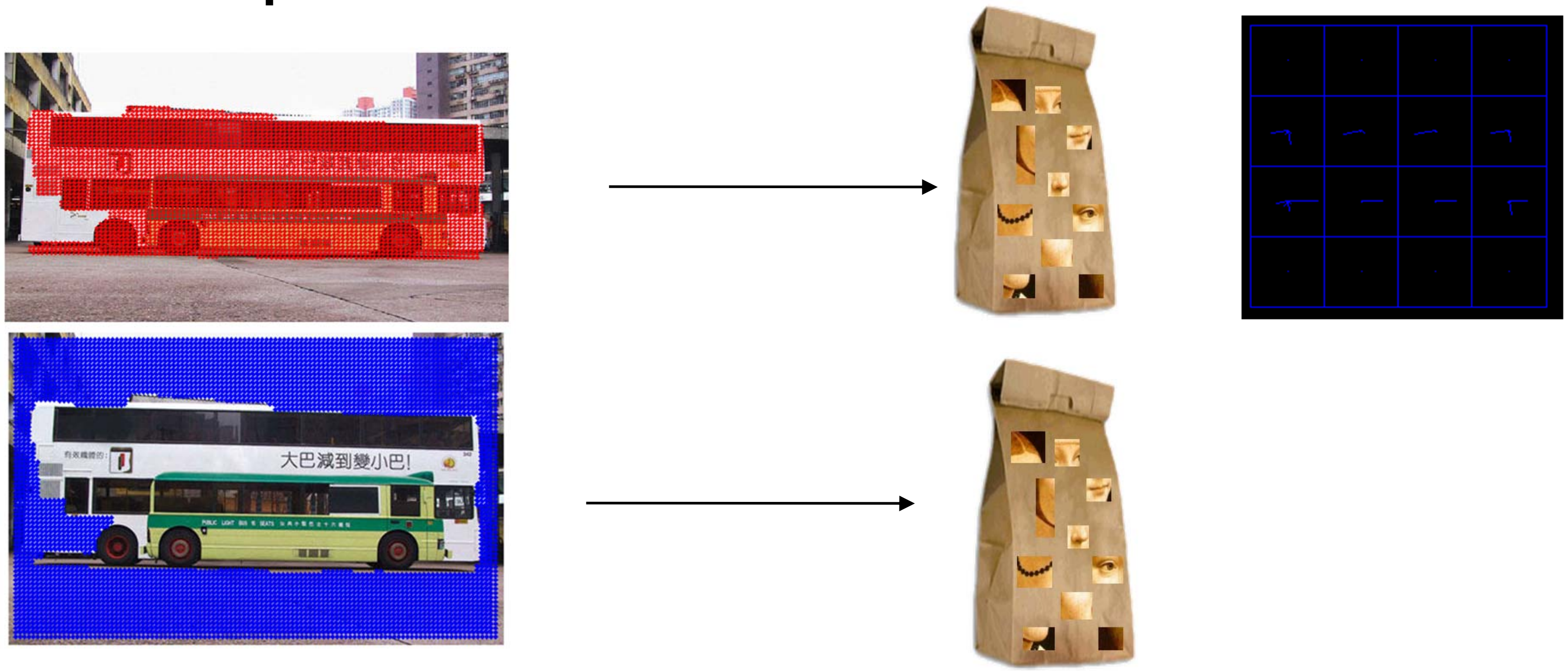
- Support Vector Regression on class-specific overlap
- 1 Regressor for each class (one against all)
- Negative examples are assigned overlap zero



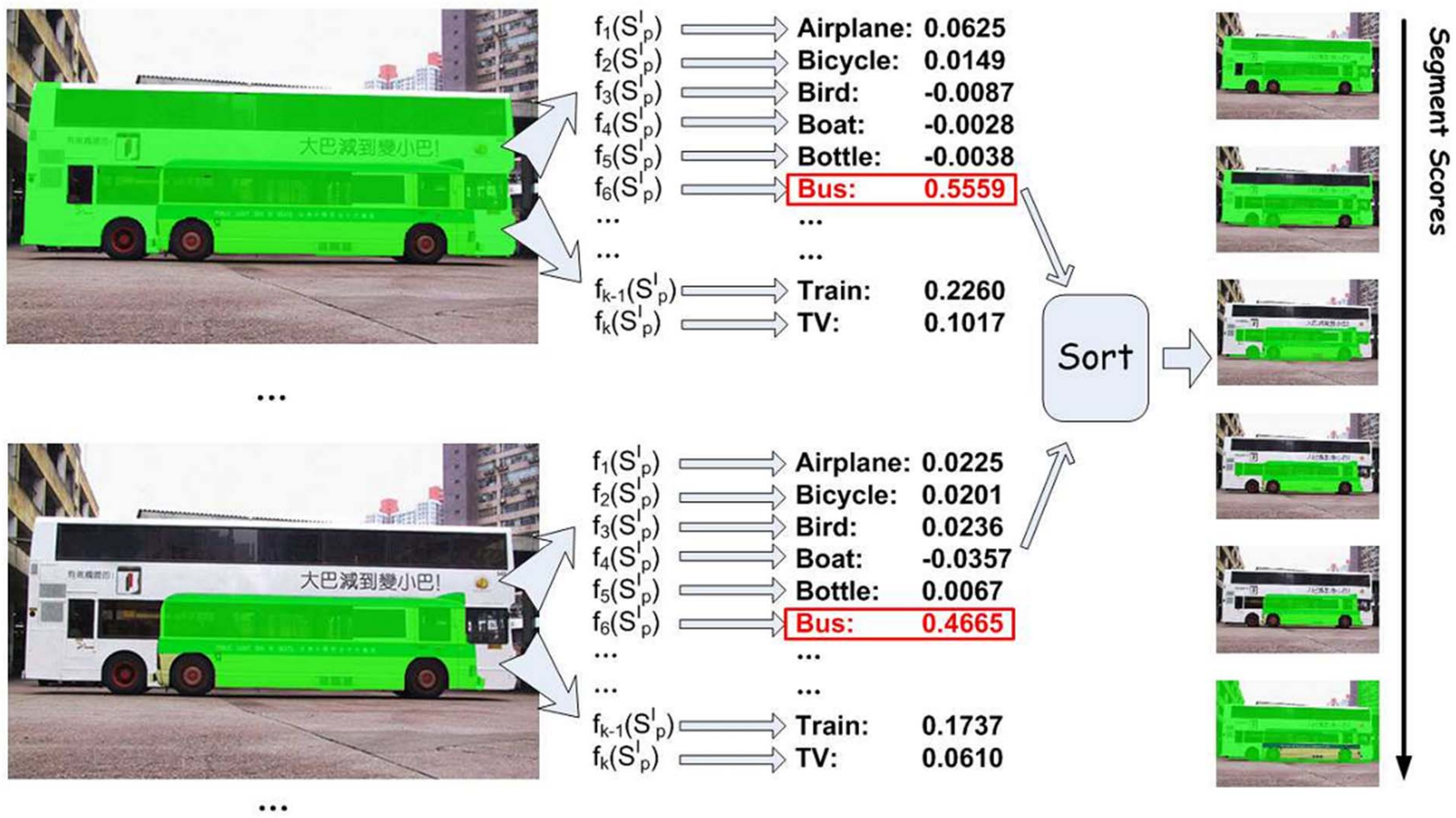
Li, Carreira and Sminchisescu, CVPR 10, IJCV 11

Features

- Extract on both segment foreground and background
- BOW-SIFT, BOW-CSIFT, HOG
- Multiple kernels combined

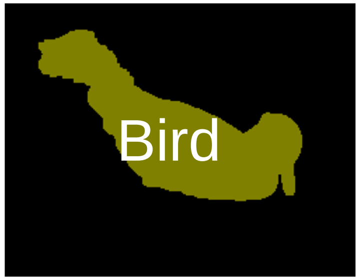


Test on sliding segments

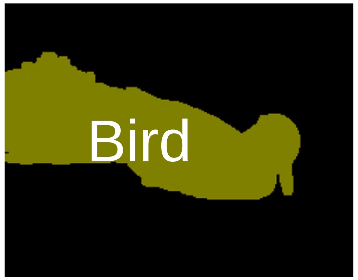


Segment Consolidation

Combine highly-overlapping segments.



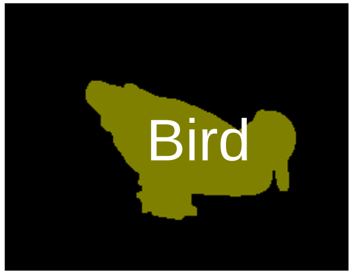
Bird: 0.393



Bird: 0.368

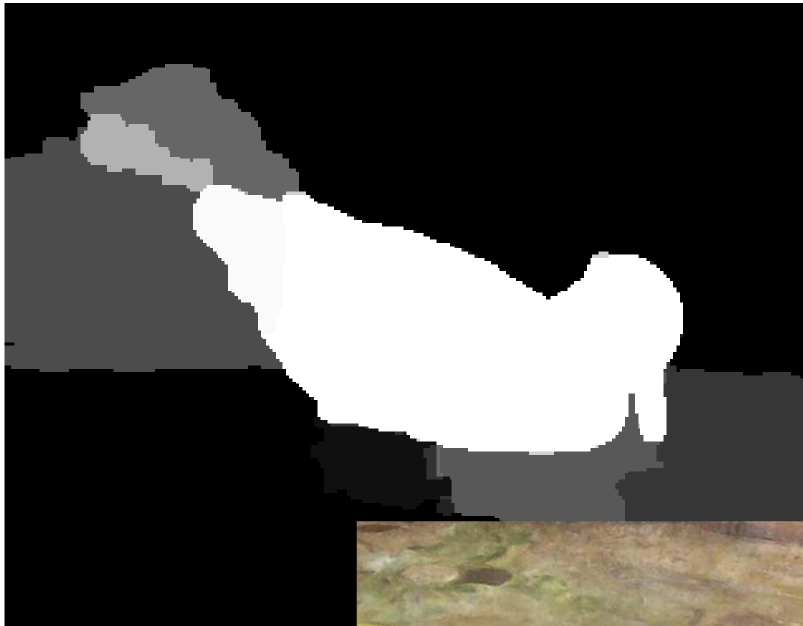
Bird: 0.351

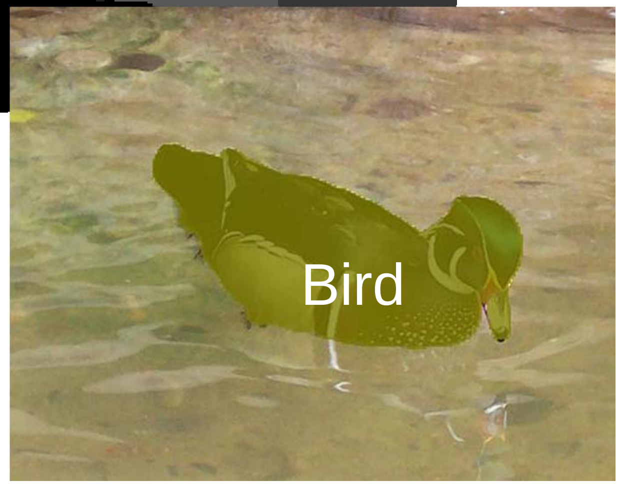
Bird: 0.332



Bird: 0.329

More robust because multiple segments agree on the object region!





Segment Consolidation

- Sequential design: The process is performed for the highest-scored overlapping segments, then lower-scored ones
- Object co-occurrence constraints also added
- More robust than non-maximum suppression
  - 2.5% better on VOC10 validation

Results

- VOC 2011 Result: 43.3%, wins 14 out of 21 classes. 2 methods from Bonn (SVR-SEGM and FGT-SEGM win 20 out of 21 classes (standard challenge)
- Overall with results of challenge 6, we still win 11 classes, despite using less annotations

	mean	bk grnd	aero plane	bike	bird	boat	bottle	bus	car	cat	chair	cow	Dine table	dog	horse	motor bike	person	pot plant	sheep	sofa	train	tv
BONN_FGT_SEGM	41.4	83.4	51.7	23.7	46	33.9	49.4	66.2	56.2	41.7	10.4	41.9	29.6	24.4	49.1	50.5	39.6	19.9	44.9	26.1	40	41.6
BONN_SVR_SEGM	43.3	84.9	54.3	23.9	39.5	35.3	42.6	65.4	53.5	46.1	15	47.4	30.1	33.9	48.8	54.4	46.4	28.8	51.3	26.2	44.9	37.2
BROOKES_STRUCT_DET_CRF	31.3	79.4	36.6	18.6	9.2	11	29.8	59	50.3	25.5	11.8	29	24.8	16	29.1	47.9	41.9	16.1	34	11.6	43.3	31.7
NUS_CONTEXT_SVM	35.1	77.2	40.5	19	28.4	27.8	40.7	56.4	45	33.1	7.2	37.4	17.4	26.8	33.7	46.6	40.6	23.3	33.4	23.9	41.2	38.6
NUS_SEG_DET_MASK_CLS_CRF	37.7	79.8	41.5	20.2	30.4	29.1	47.4	61.2	47.7	35	8.5	38.3	14.5	28.6	36.5	47.8	42.5	28.5	37.8	26.4	43.5	45.8
BERKELEY_REGION_CLASSIFY	39.1	83.3	48.9	20	32.8	28.2	41.1	53.9	48.3	48	6	34.9	27.5	35	47.2	47.3	48.4	20.6	52.7	25	36.6	35.4

Example Images in Each Category

- Almost perfect

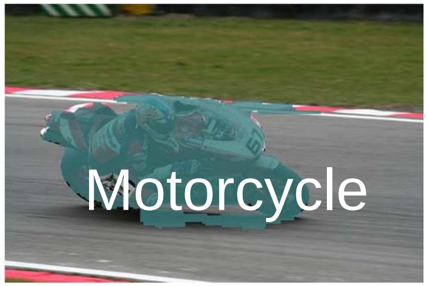
- Correct to an extent

- False Positive

Example Images in Each Category

- Misclass-ifications

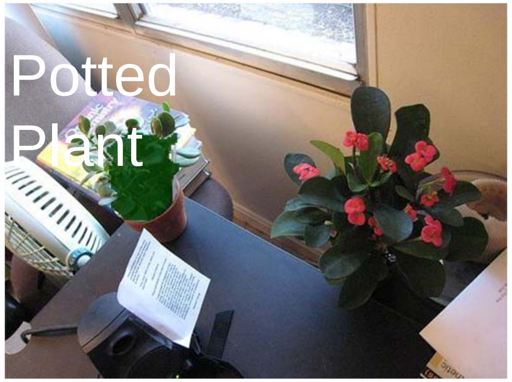
- Misses
  - Objects in interaction



  - Detected but not retained

Example Images in Each Category

Other misses:



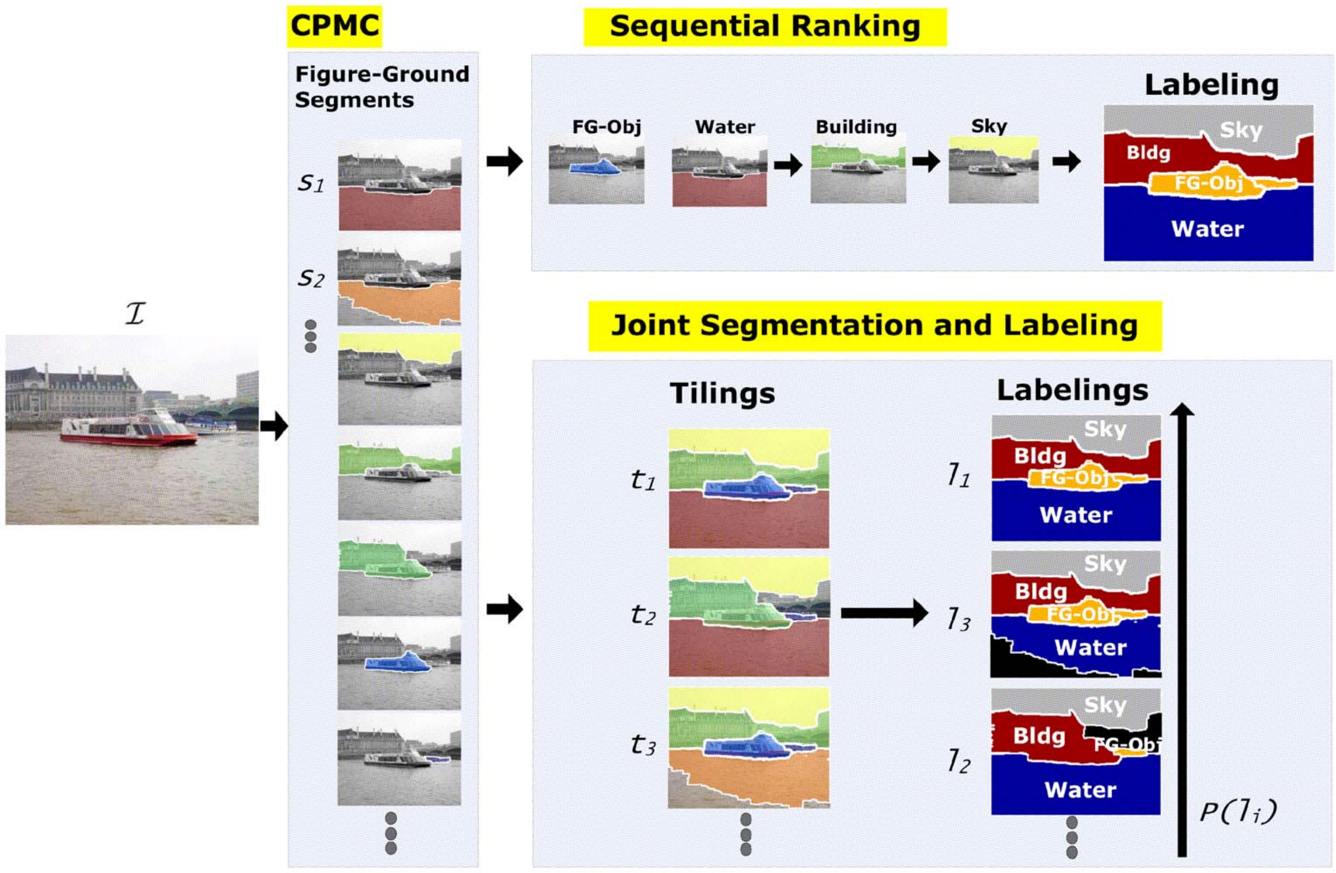
Conclusions

Sequential recognition framework

- Operates with figure-ground segmentations generated by combinatorial optimization (parametric max-flow) with diversified ranking on mid-level features (CPMC)
- Regression on overlap exploits information in both full object segments and fragments
- Sliding segment-based recognition approach with consolidation improves over non-maximum suppression

See also our second entry for Joint Segmentation and Labeling

BONN-FGT-SEGM



Ion, Carreira, Sminchisescu, ICCV 11, NIPS 11

Thank you!

CPMC segmentation code is available online at
http://sminchisescu.ins.uni-bonn.de/cpmc