Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FACULTY OF MATHEMATICS, INFORMATICS AND NATURAL SCIENCES
DEPARTMENT OF INFORMATICS
COMPUTER VISION GROUP

Christian Wilms and Simone Frintrop
{wilms,frintrop}@informatik.uni-hamburg.de

AttentionMask: Attentive, Efficient Object Proposal Generation Focusing on Small Objects

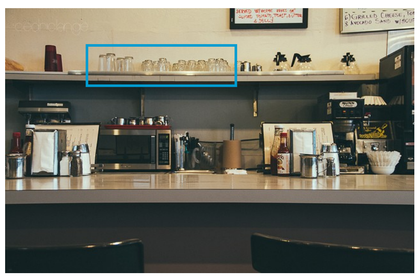
Motivation

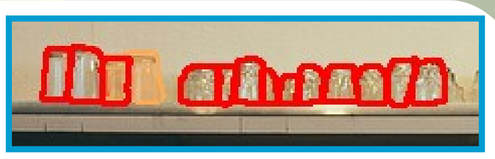
Task: Object Proposal Generation
Goal: Generating class-agnostic object candidates
Problem 1: State-of-the-art systems often miss small objects
Problem 2: Simply adding a module to detect small objects is impossible due to inefficient use of resources
Our idea: Starting from [1], we focus processing on relevant parts of the image to save resources
→ use those resources to better detect small objects

Original

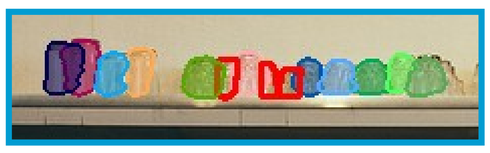
FastMask [1]

The problem with small objects...





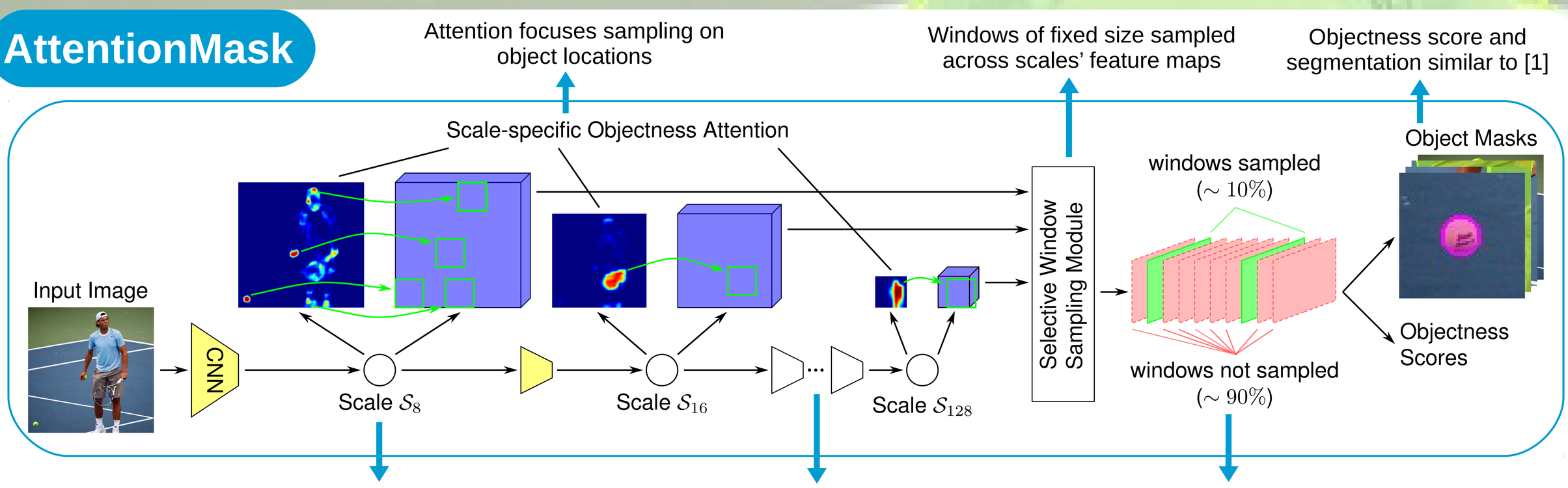
SharpMaskZoom [2]



AttentionMask (ours)

Red non-filled shapes denote missed objects. Colored filled shapes denote found objects.

AttentionMask



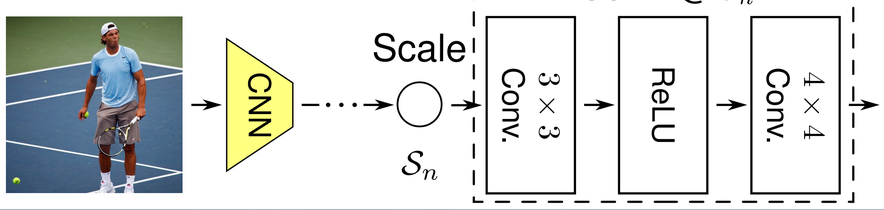
- New scale for small objects
- **Only possible because of efficient use of resources due to attention**

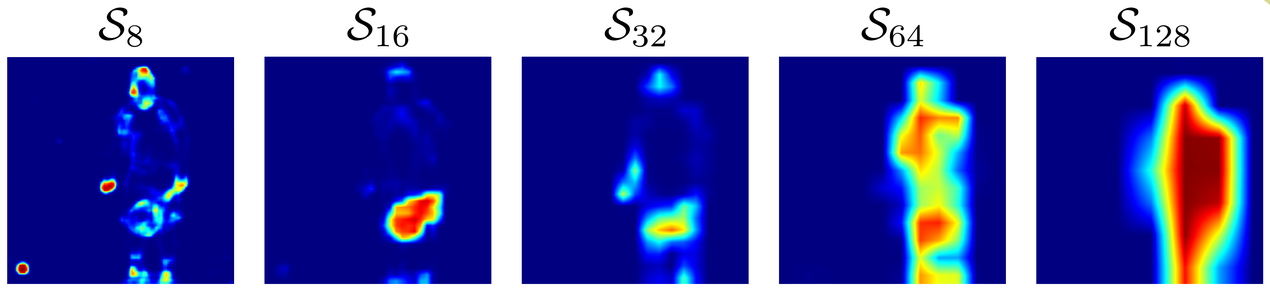
- Feature pyramid with 5-9 scales
- Downscaled by factor 8 to 196

91.7% of the windows are pruned due to low attention
→ **resources saved**

Scale-specific Objectness Attention

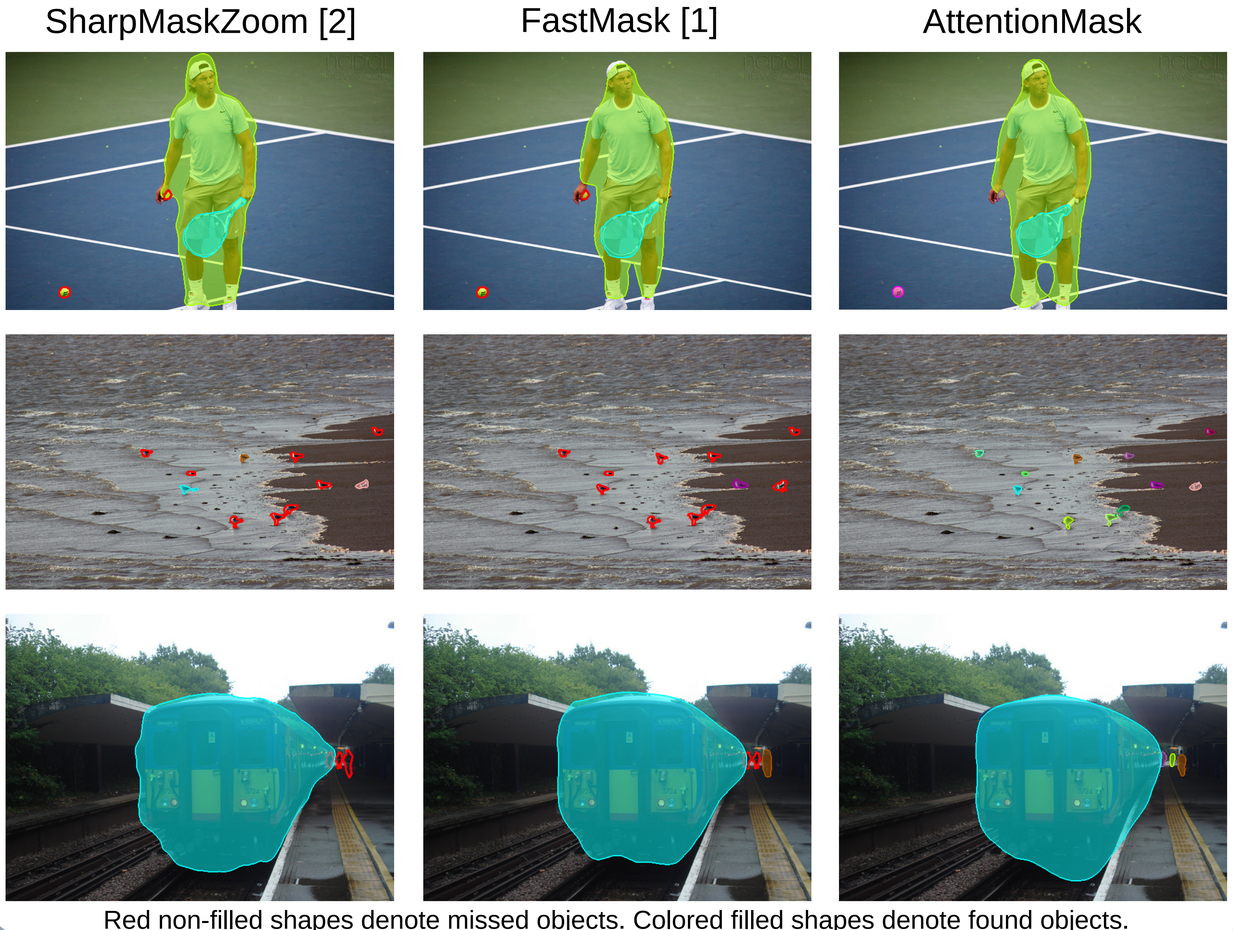
Attention modules (SOAMs) are very fast and lightweight components





Each scale has its own attention, thus focusing on objects of different sizes

Results (for more results scan the QR code)



Red non-filled shapes denote missed objects. Colored filled shapes denote found objects.

Evaluation on MS COCO

- Evaluation against several state-of-the-art systems
- Average recall (AR@# of proposals) is used as evaluation measure

	Across all scales			Small / Medium / Large Objects			
Method	AR@10	AR@100	AR@1k	AR^S@100	AR^M@100	AR^L@100	Time
MCG [4]	0.077	0.186	0.299	-	-	-	$45s$
DeepMaskZoom [2]	0.151	0.286	0.371	0.093	0.389	0.466	$1.35s$
SharpMask [2]	0.154	0.278	0.360	0.035	0.399	0.513	$1.03s$
SharpMaskZoom [2]	0.156	0.304	0.401	0.099	0.412	0.495	$2.02s$
InstanceFCN [3]	0.166	0.317	0.392	-	-	-	$1.50s$
FastMask [1]	0.169	0.313	0.406	0.106	0.406	0.517	$0.33s$
AttentionMask	**0.180**	**0.349**	**0.444**	**0.162**	**0.421**	**0.560**	**0.22s**

AttentionMask beats all state-of-the-art methods across all categories including runtime!

Paper + Code

References

[1] Hu, H., Lan, S., Jiang, Y. Cao, Z. Sha, F. : FastMask: Segment Multi-scale Object Candidates in One Shot. In: CVPR (2017)
[2] Pinheiro, P., Lin, T., Collobert, R., Dollár, P..: Learning to refine object segments. In ECCV (2016)
[3] Dai, J., He, K., Li, Y., Ren, S., Sun, J.: Instance-sensitive fully convolutional networks. In: ECCV (2016)
[4] Pont-Tuset, J., Arbelaez, P., Barron, J., Marques, F, Malik, J..: Multiscale combinatorial grouping for image segmentation and object proposal generation. TPAMI 39(1), 128-140 (2017)