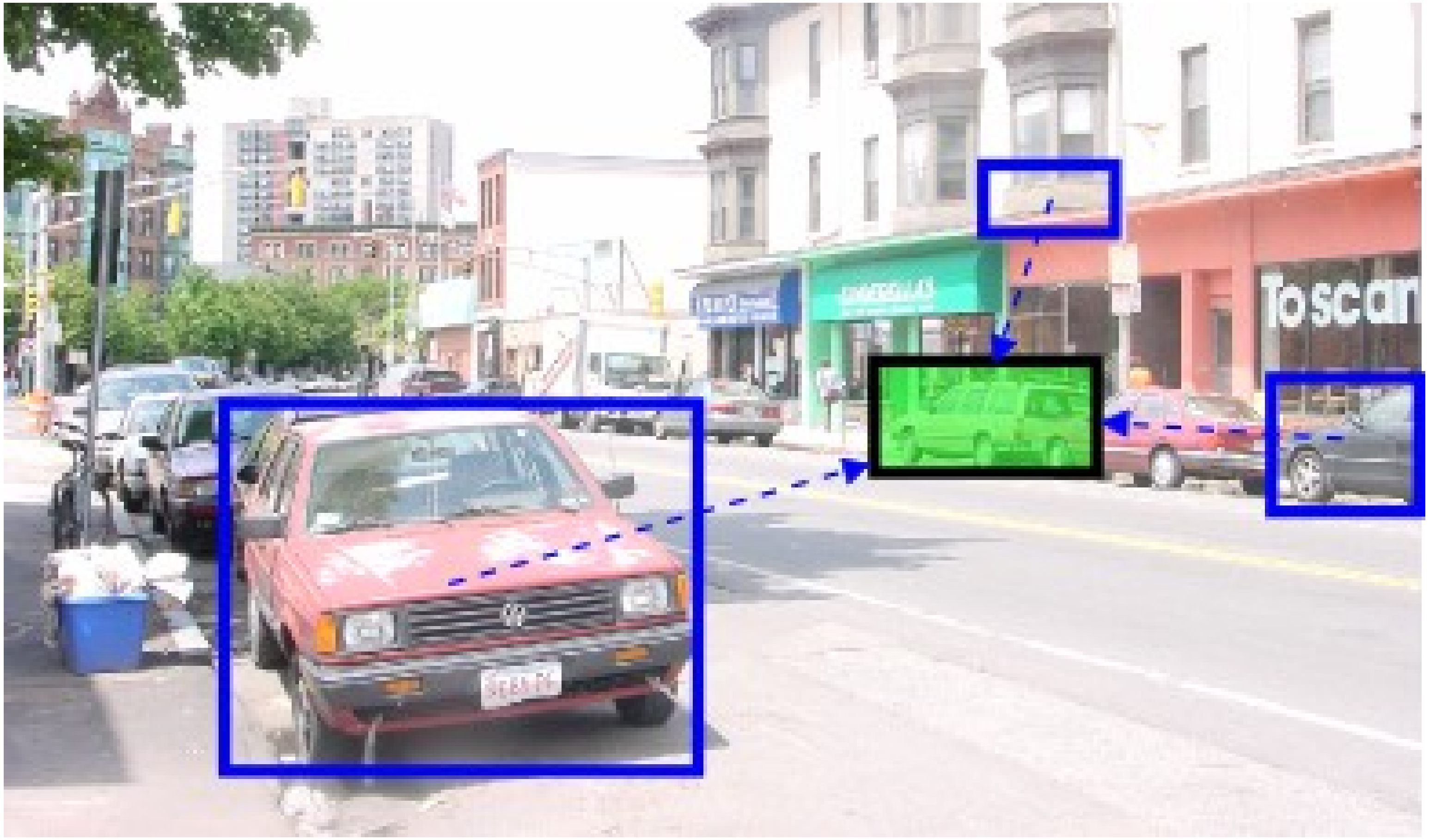
Towards Cautious Collective Inference for Object Verification.

José Oramas M., Luc De Raedt, Tinne Tuytelaars
KU Leuven
March 24th 2014

The task at hand

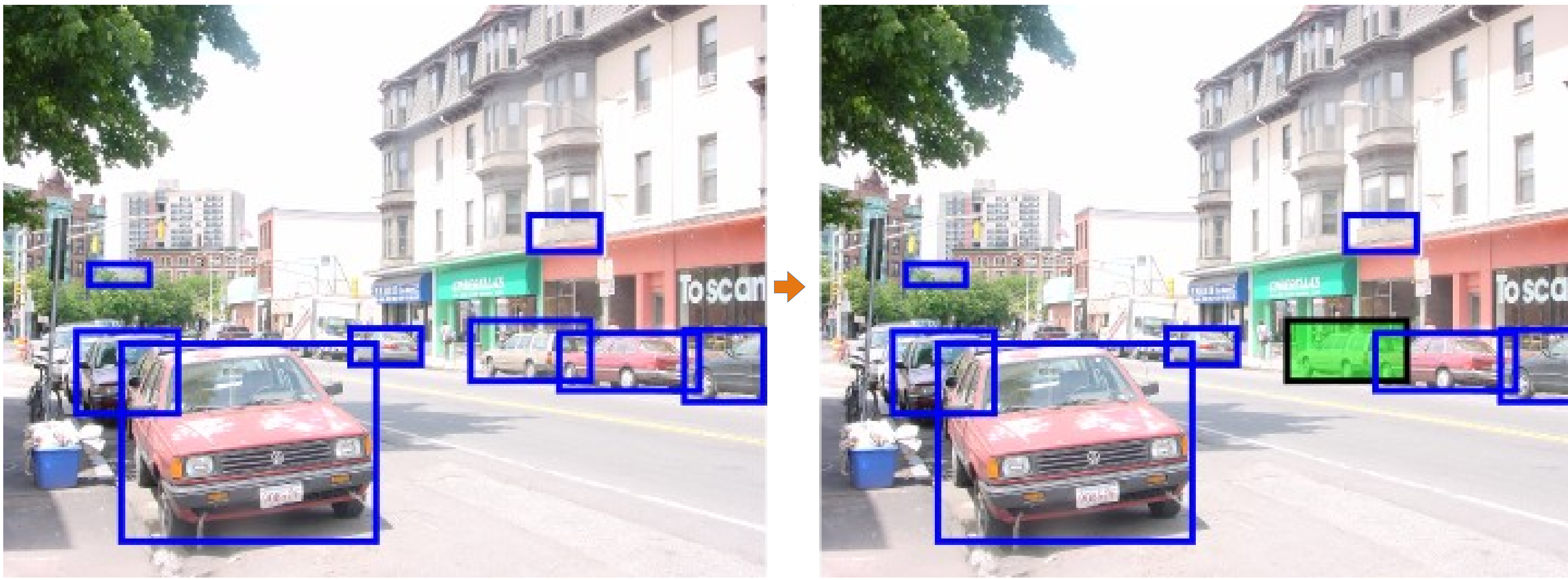
Object Verification

(Context-based Object Detection)



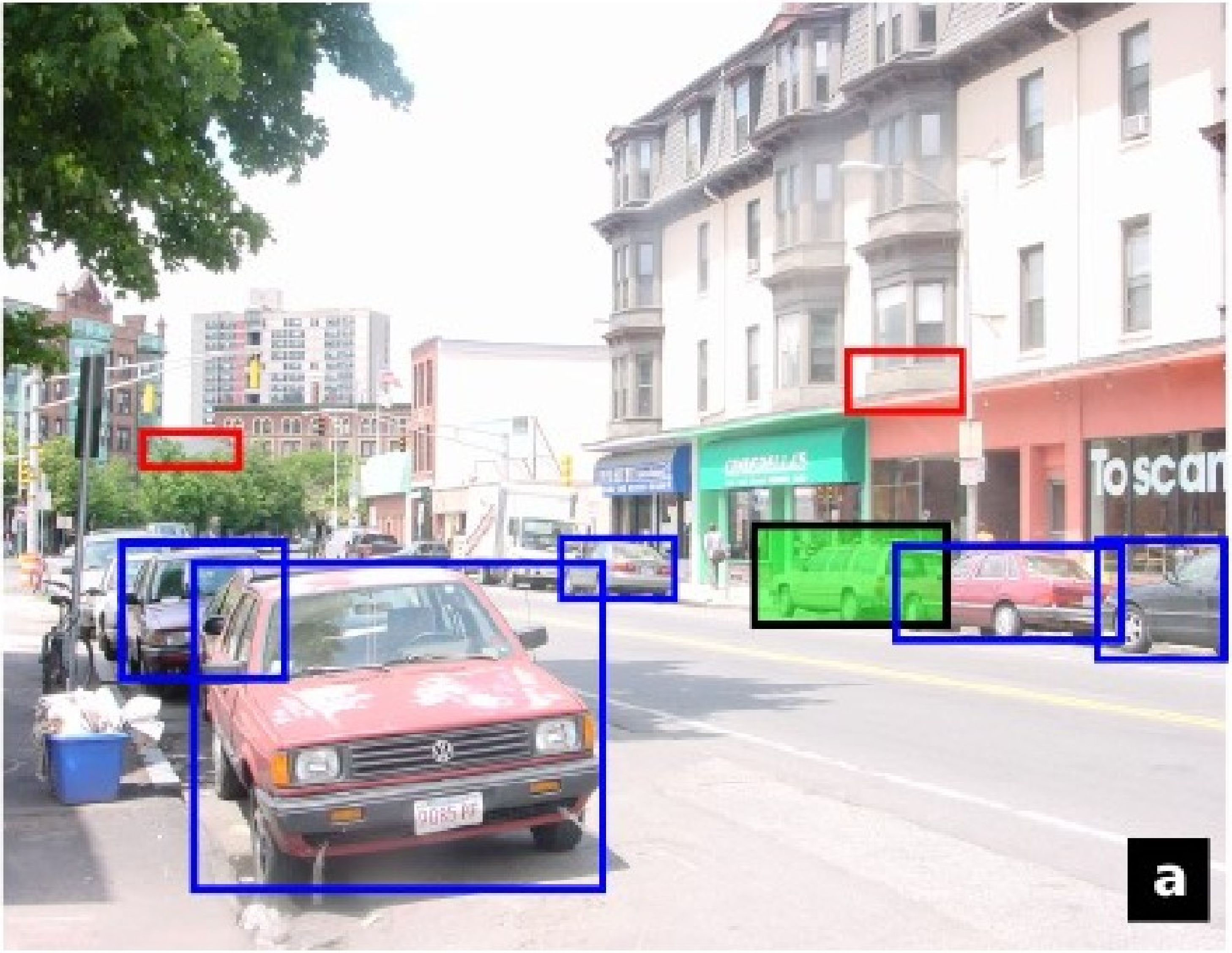
Object Verification

A traditional pipeline



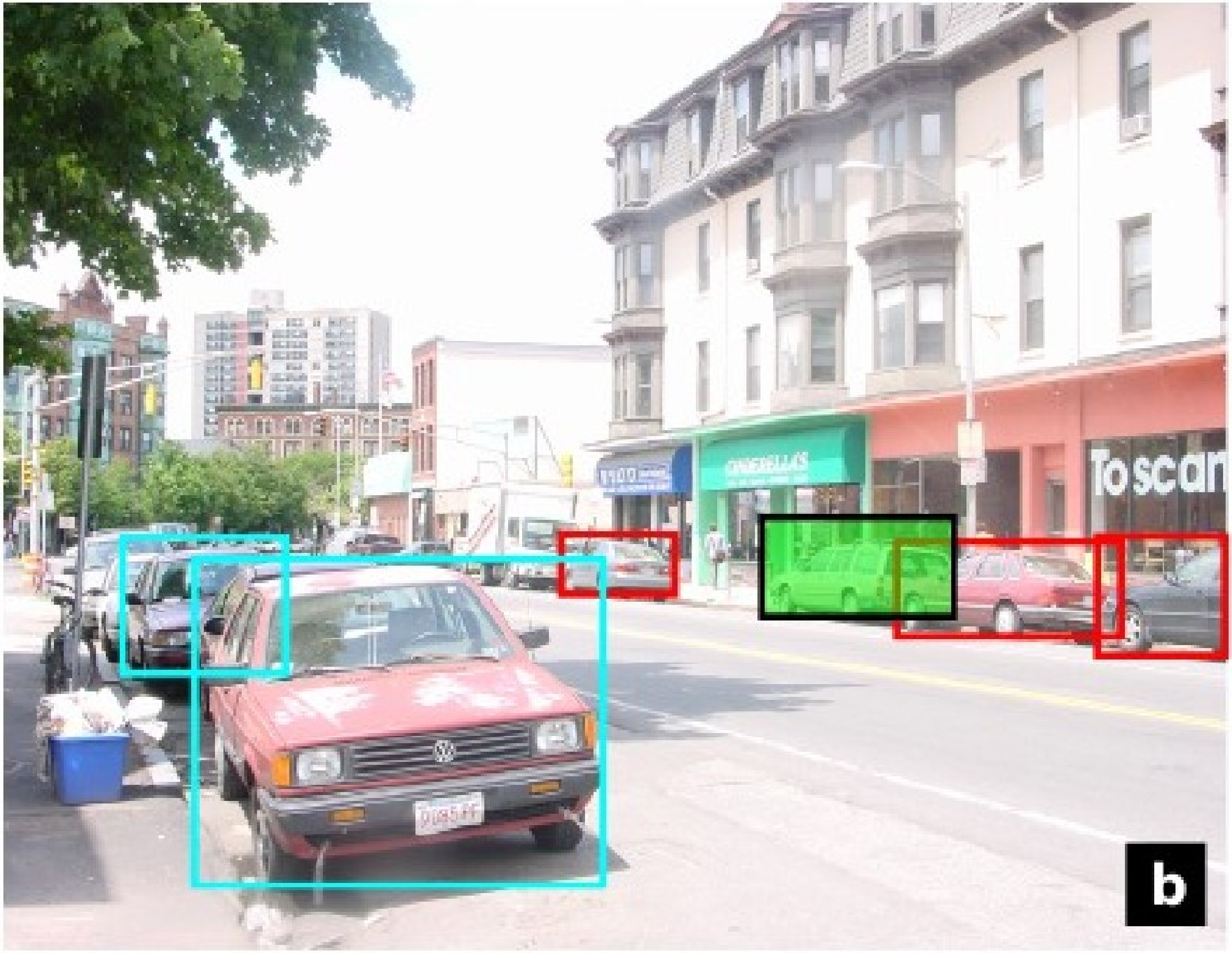
Problem Statement

(Q) How to properly use relations between objects?



Problem Statement

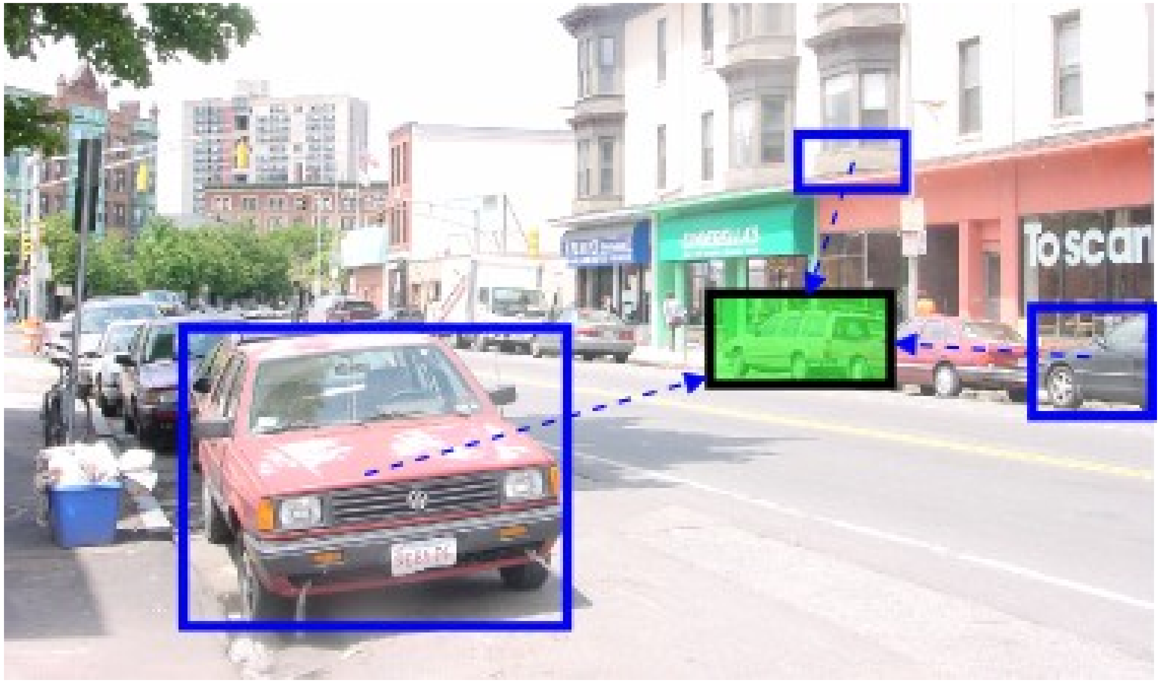
(Q) How to properly use relations between objects?



Proposed Method

Defining Pairwise Relations between Objects

Derived from the features of the bounding boxes



Relations :

$$r_{ij}^{(RF1)} = (\Delta x_{ij}, \Delta y_{ij}, \Delta\theta_{ij})$$

$$r_{ij}^{(RF2)} = (rx_{ij}, ry_{ij}, r\rho_{ij}, ra_{ij})$$ (based on Li et al's, CVPR 2012)

$$r_{ij}^{(RF3)} = (\Delta x_{ij}, \Delta y_{ij})$$ (based on Perko et al's, CVIU 2010)

Proposed Method

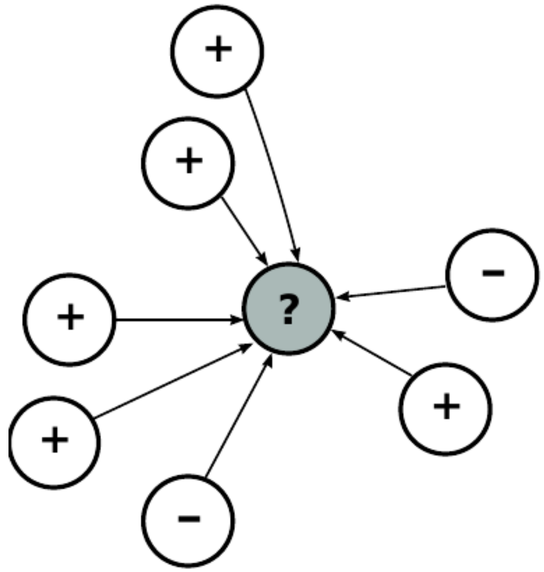
Aggressive Inference

$$wvRN(o_i|N_i) = \frac{1}{z}\sum_{o_j \in N_i} p(o_i|o_j).w_{ij}$$

(Mackassy et al., JMLR 2007)

$$p(o_i|o_j) = p(o_i|r_{ij}) = \frac{p(r_{ij}|o_i)p(o_i)}{p(r_{ij}|o_i)p(o_i) + p(r_{ij}|\neg o_i)p(\neg o_i)}$$

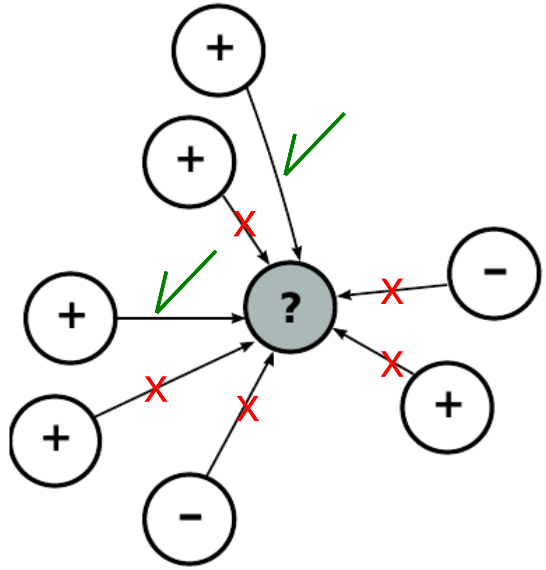
$$w_j = p(o_j|s_j) = \frac{p(s_j|o_j)p(o_j)}{p(s_j|o_j)p(o_j) + p(s_j|\neg o_j)p(\neg o_j)}$$



Cautious Inference

- McDowel et al., JMLR 2009
- Neville et al, SRL ws@AAAI 2000

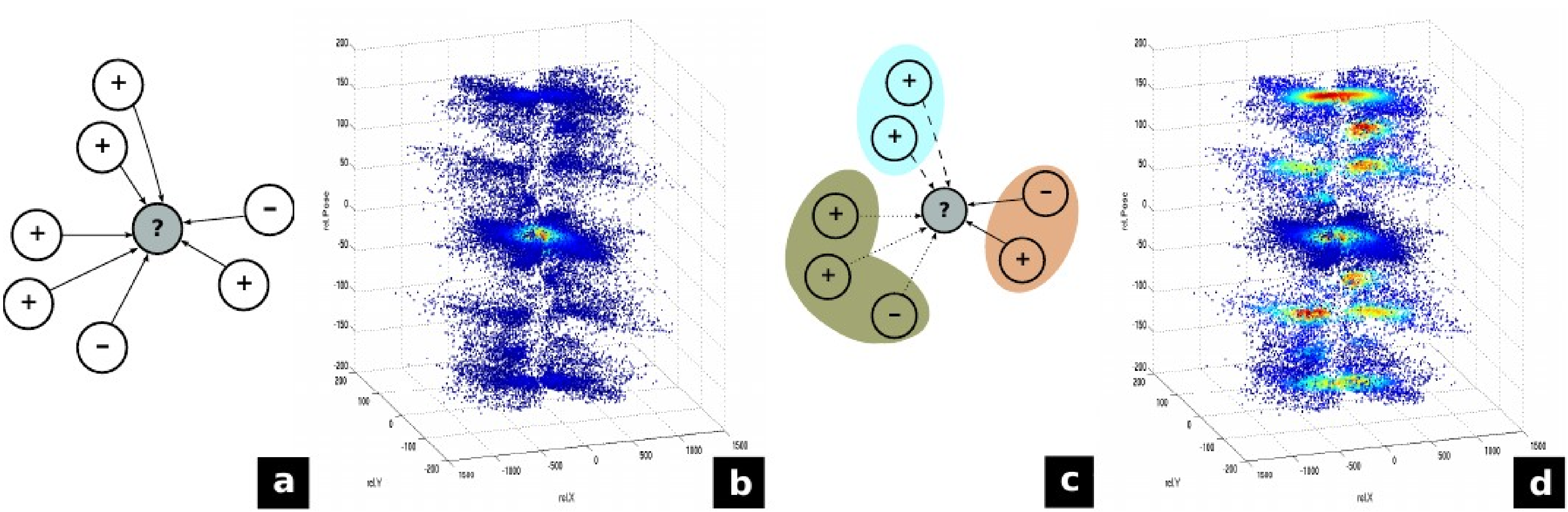
$$wvRN(o_i^u|N_i) = \frac{1}{z}\sum_{o_j^k \in (N_i \cap O^k)} p(o_i^u|o_j^k).w_{ij}$$



Proposed Method

How objects associate to each other ?

Class-based Homophily | **Relationship-based Homophily**

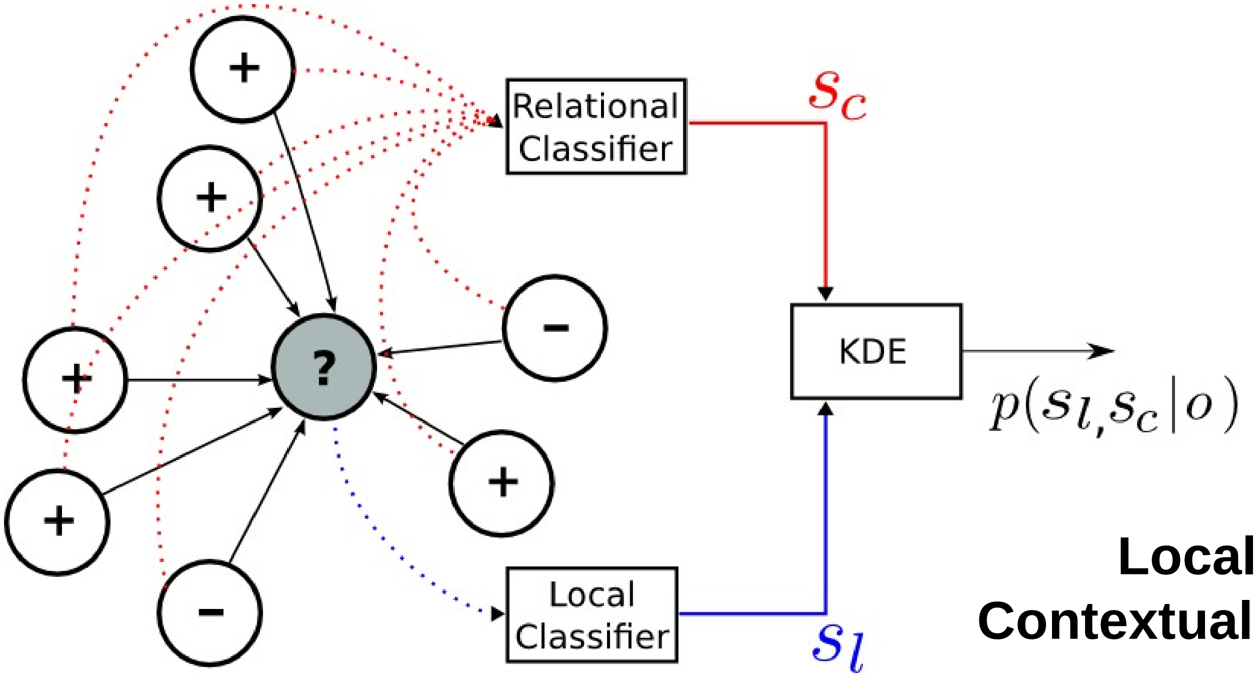


$$wvRN(o_i|N_i) = \frac{1}{z} \sum_{o_j \in N_i} p(o_i|o_j).w_{ij}$$

Proposed Method

Combining local and contextual sources of Information

Inspired by the score combination method of Perko et al., CVIU 2010.



Local Support : $s_l = p(\hat{o}_i)$

Contextual Support : $s_c = wvRN(o_i | N_i)$

- Score Combination

$$p(o|s_l, s_c) = \frac{p(s_l, s_c|o)p(o)}{p(s_l, s_c|o)p(o) + p(s_l, s_c|\neg o)p(\neg o)}$$

Evaluation

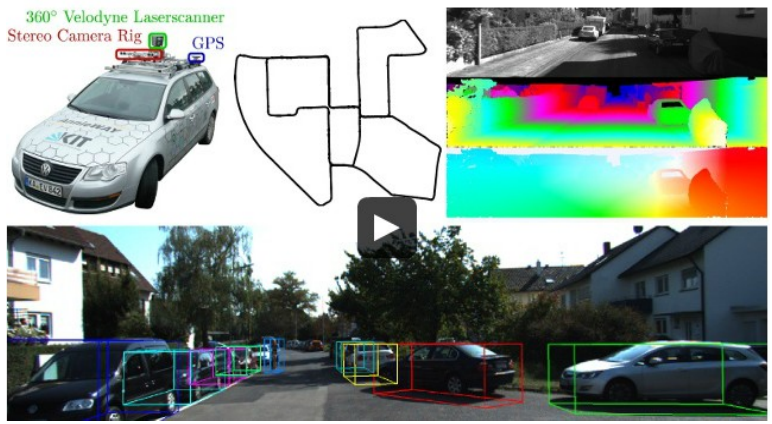
Experimental details

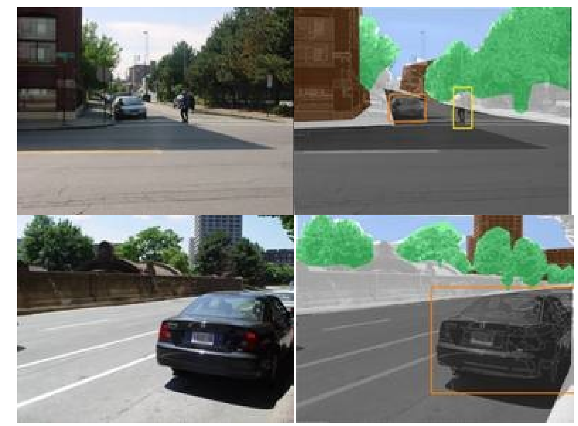
Datasets:

KITTI benchmark
(Geiger et al, CVPR 2012)

MIT StreetScenes
(Bileschi et al)

- Class of Interest : **car**
- Focus on Images with more than 2 objects



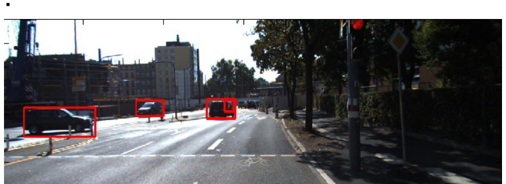


Object Localization:

Object Detection (2D)

[1] López et al. ,ws@ICCV 2011.

.



Parameters:

- **Inference Type:** *Aggressive* | *Cautious .*
- **Homophily Type:** *class-based* | *relation-based .*

Evaluation

Only using contextual information

Dataset **KITTI benchmark**	**Relations Representation : RF1**				**Relations Representation : RF2**			
	Class-based Hom.		Relation-based Hom.		Class-based Hom.		Relation-based Hom.	
	Global		Global		Global		Global	
Set	aggre.	caut.	aggre.	caut.	aggre.	caut.	aggre.	caut.
all	0.29	***0.38***	0.28	***0.37***	0.32	0.40	0.41	***0.50***

Dataset **MIT StreetScenes**	**Relations Representation : RF3**				**Relations Representation : RF2**			
	Class-based Hom.		Relation-based Hom.		Class-based Hom.		Relation-based Hom.	
	Global		Global		Global		Global	
Set	aggre.	caut.	aggre.	caut.	aggre.	caut.	aggre.	caut.
all	0.54	***0.63***	0.51	0.59	0.51	***0.56***	0.49	0.55

Combination of Local and Contextual Information

Dataset **KITTI benchm**[1]		**RF1**		**RF2**	
		Class-based Homophily		Relation-based Homophily	
		Global		Global	
Set	Detector [14]	aggre.	caut.	aggre.	caut.
all	0.61 ± 0.011	0.61 ± 0.009	0.63 ± 0.007	0.65 ± 0.011	$\boldsymbol{0.68\pm0.003}$

Dataset **MIT StreetScei**[1]		**RF3**		**RF2**	
		Class-based Homophily		Class-based Homophily	
		Global		Global	
Set	Detector [14]	aggre.	caut.	aggre.	caut.
all	0.69 ± 0.006	0.77 ± 0.001	$\boldsymbol{0.80\pm0.028}$	0.73 ± 0.011	0.76 ± 0.014

Collecting Hypotheses using DPM

Dataset **KITTI benchmark**		**RF3**		**RF2**	
		Class-based Homophily		Relation-based Homophily	
		Global		Global	
Set	Detector [2]	aggre.	caut.	aggre.	caut.
all	0.65 ± 0.003	0.68 ± 0.007	0.71 ± 0.007	0.72 ± 0.009	$\mathbf{0.75\pm0.003}$

Dataset **MIT StreetScenes**		**RF3**		**RF2**	
		Class-based Homophily		Class-based Homophily	
		Global		Global	
Set	Detector [2]	aggre.	caut.	aggre.	caut.
all	0.62 ± 0.004	0.66 ± 0.011	$\mathbf{0.71\pm0.012}$	0.65 ± 0.026	0.69 ± 0.014

[1] López et al. ,ws@ICCV 2011.
[2] Felzenszwalb et al. ,TPAMI 2010.

Evaluation

Only using contextual information

Dataset **KITTI benchmark**	**Relations Representation : RF1**				**Relations Representation : RF2**			
	Class-based Hom.		Relation-based Hom.		Class-based Hom.		Relation-based Hom.	
	Global		Global		Global		Global	
Set	aggre.	caut.	aggre.	caut.	aggre.	caut.	aggre.	caut.
all	0.29	***0.38***	0.28	***0.37***	0.32	0.40	0.41	***0.50***

Dataset **MIT StreetScenes**	**Relations Representation : RF3**				**Relations Representation : RF2**			
	Class-based Hom.		Relation-based Hom.		Class-based Hom.		Relation-based Hom.	
	Global		Global		Global		Global	
Set	aggre.	caut.	aggre.	caut.	aggre.	caut.	aggre.	caut.
all	0.54	***0.63***	0.51	0.59	0.51	***0.56***	0.49	0.55

Combination of Local and Contextual Information

Dataset **KITTI benchm[1]**		**RF1**		**RF2**	
		Class-based Homophily		Relation-based Homophily	
		Global		Global	
Set	Detector [14]	aggre.	caut.	aggre.	caut.
all	0.61 ± 0.011	0.61 ± 0.009	0.63 ± 0.007	0.65 ± 0.011	$\boldsymbol{0.68\pm0.003}$

Dataset **MIT StreetSce[1]**		**RF3**		**RF2**	
		Class-based Homophily		Class-based Homophily	
		Global		Global	
Set	Detector [14]	aggre.	caut.	aggre.	caut.
all	0.69 ± 0.006	0.77 ± 0.001	$\boldsymbol{0.80\pm0.028}$	0.73 ± 0.011	0.76 ± 0.014

Collecting Hypotheses using DPM

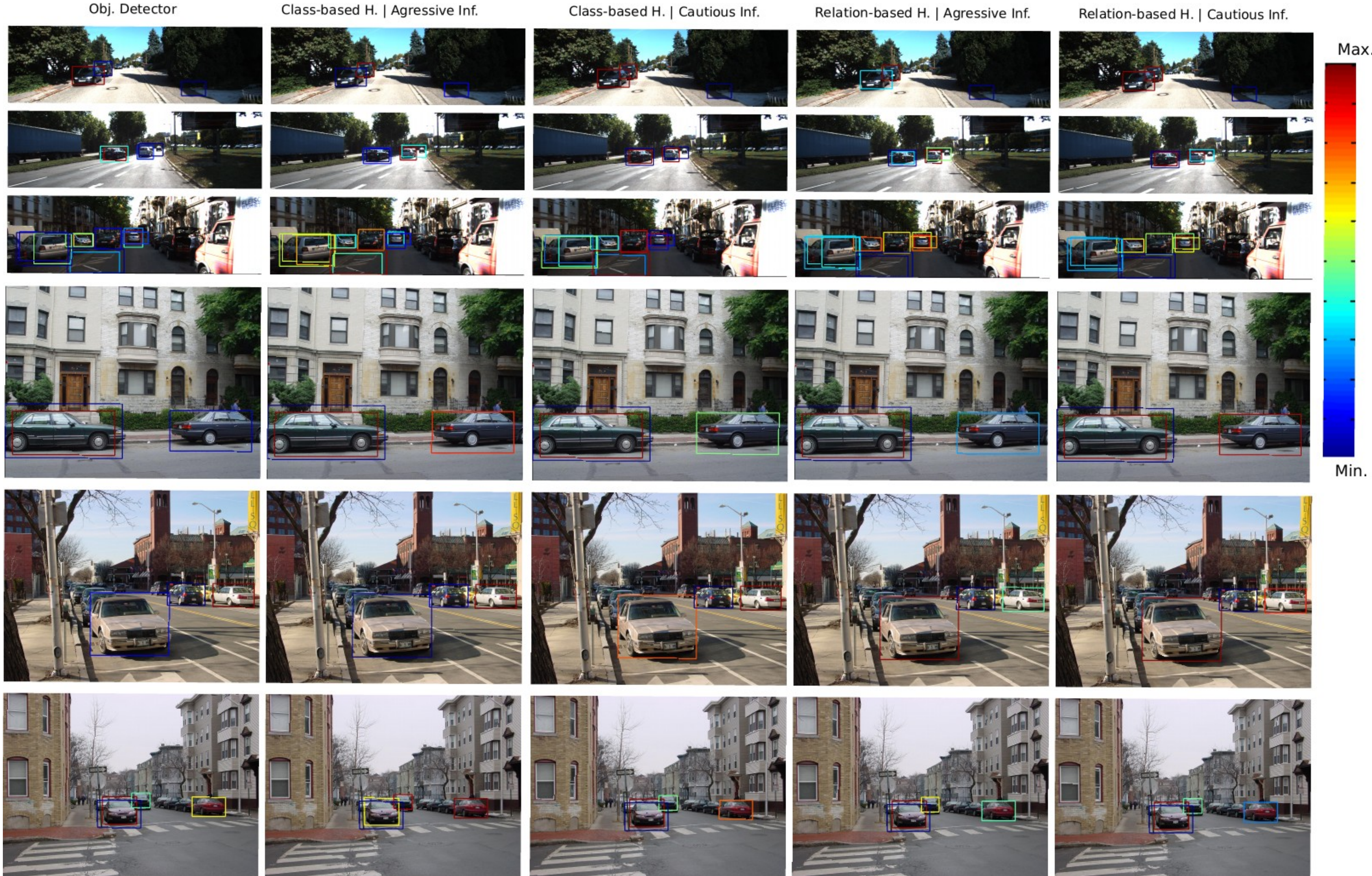
Dataset **KITTI benchmark**		**RF3**		**RF2**	
		Class-based Homophily		Relation-based Homophily	
		Global		Global	
Set	Detector [2]	aggre.	caut.	aggre.	caut.
all	0.65 ± 0.003	0.68 ± 0.007	0.71 ± 0.007	0.72 ± 0.009	$\mathbf{0.75\pm0.003}$

Dataset **MIT StreetScenes**		**RF3**		**RF2**	
		Class-based Homophily		Class-based Homophily	
		Global		Global	
Set	Detector [2]	aggre.	caut.	aggre.	caut.
all	0.62 ± 0.004	0.66 ± 0.011	$\mathbf{0.71\pm0.012}$	0.65 ± 0.026	0.69 ± 0.014

[1] López et al. ,ws@ICCV 2011.
[2] Felzenszwalb et al. ,TPAMI 2010.

Evaluation

Qualitative Results



Conclusions

- **Cautious Inference about object relations outperforms traditional aggressive counterparts.**

- **Relation-based Homophily is good for scenarios where there is no local information about the unknown object.**
  (e.g. in an inpainting scenario)

- **Future Work**
  - Better representations for reasoning in 3D space.
  - Investigate methods to recover the underlying structures in the relational space.
  - Investigate the generality of the method in the context of other object categories or alternative applications.

Questions?

KU LEUVEN

PSI VISICS

Towards Cautious Collective Inference for Object Verification.

José Oramas M., Luc De Raedt, Tinne Tuytelaars
KU Leuven
March 24th 2014