

The PASCAL Visual Object Classes Challenge 2007 (VOC2007)

Part 2 – Detection Task

Mark Everingham

Luc Van Gool

Chris Williams

John Winn

Andrew Zisserman

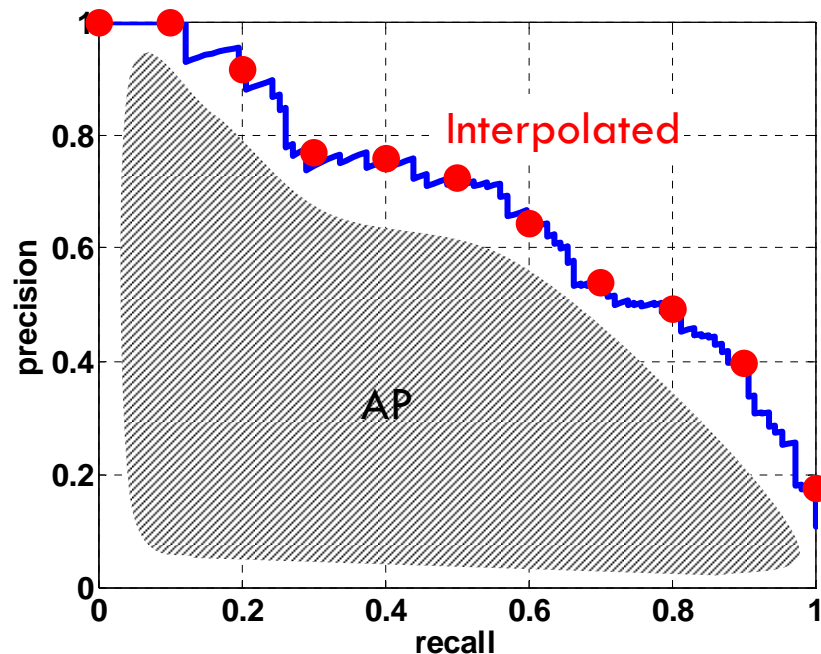


Detection Challenge

- Predict the bounding boxes of all objects of a given class in an image (if any)
- Competition 3: Train on the supplied data
 - Which methods perform best given specified training data?
- Competition 4: Train on any (non-test) data
 - How well do state-of-the-art methods perform on these problems?
 - **No results submitted**

Evaluation

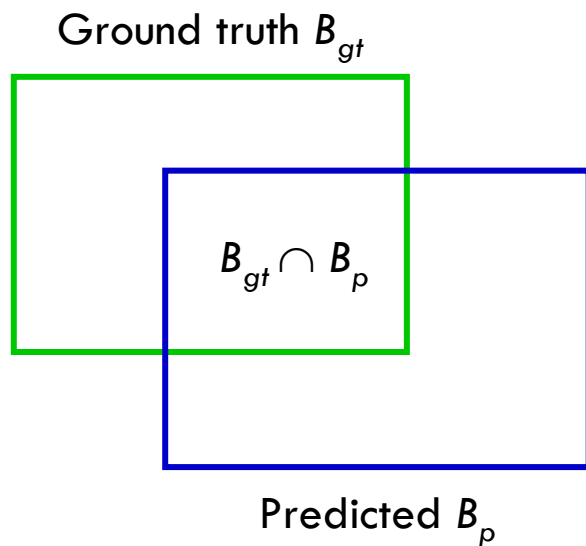
- Average Precision [TREC] averages precision over the entire range of recall
 - Curve interpolated to reduce influence of “outliers”



- A good score requires both high recall and high precision
- Application-independent
- Penalizes methods giving high precision but low recall

Evaluating Bounding Boxes

- Area of Overlap (AO) Measure



$$AO(B_{gt}, B_p) = \frac{|B_{gt} \cap B_p|}{|B_{gt} \cup B_p|}$$

- Need to define a threshold t such that $AO(B_{gt}, B_p)$ implies a correct detection: 50%

Methods

- **“Sliding window classifier” predominant**
- Features
 - HoG
 - Pairs of adjacent edge segments (PAS)
 - Local histograms of 1st, 2nd derivative, color
 - Quantized SIFT at interest points (bag of words)
 - Coefficients of topics from LDA

Methods

- Classifiers
 - AdaBoost
 - Linear SVM
 - SVM (bag of words)
 - “Star model” of linear SVM’s (“part detectors”)
- Other aspects
 - Efficient search for window (non-sliding)
 - Windows seeded from interest points (non-sliding)
 - Combining confidence of detection and whole-image classification

Methods

- Other aspects...
 - Separate detectors for views and feature type
 - Separate detectors for “contexts” e.g. “person on motorbike”, “person on horse”
 - Sharing detection confidence between similar classes e.g. bicycle/motorbike
- Other approaches
 - Prediction of object in middle of image (?)
 - Classification of segmented regions using bag of words and shape features of regions

AP by Method and Class

	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	pers	plant	sheep	sofa	train	tv
Darmstadt	-	-	-	-	-	-	30.1	-	-	-	-	-	-	-	-	-	-	-	-	-
INRIA_Normal	9.2	24.6	1.2	0.2	6.8	19.7	26.5	1.8	9.7	3.9	1.7	1.6	22.5	15.3	12.1	9.3	0.2	10.2	15.7	24.2
INRIA_PlusClass	13.6	28.7	4.1	2.5	7.7	27.9	29.4	13.2	10.6	12.7	6.7	7.1	33.5	24.9	9.2	7.2	1.1	9.2	24.2	27.5
IRISA	-	28.1	-	-	-	-	31.8	2.6	9.7	11.9	-	-	28.9	22.7	22.1	-	17.5	-	-	25.3
MPI_Center	6.0	11.0	2.8	3.1	0.0	16.4	17.2	20.8	0.2	4.4	4.9	14.1	19.8	17.0	9.1	0.4	9.1	3.4	23.7	5.1
MPI_ESSOL	15.2	15.7	9.8	1.6	0.1	18.6	12.0	24.0	0.7	6.1	9.8	16.2	3.4	20.8	11.7	0.2	4.6	14.7	11.0	5.4
Oxford	26.2	40.9	-	-	-	39.3	43.2	-	-	-	-	-	-	37.5	-	-	-	-	33.4	-
TKK	18.6	7.8	4.3	7.2	0.2	11.6	18.4	5.0	2.8	10.0	8.6	12.6	18.6	13.5	6.1	1.9	3.6	5.8	6.7	9.0
UoCTTI	20.6	36.9	9.3	9.4	21.4	23.2	34.6	9.8	12.8	14.0	0.2	2.3	18.2	27.6	21.3	12.0	14.3	12.7	13.4	28.9

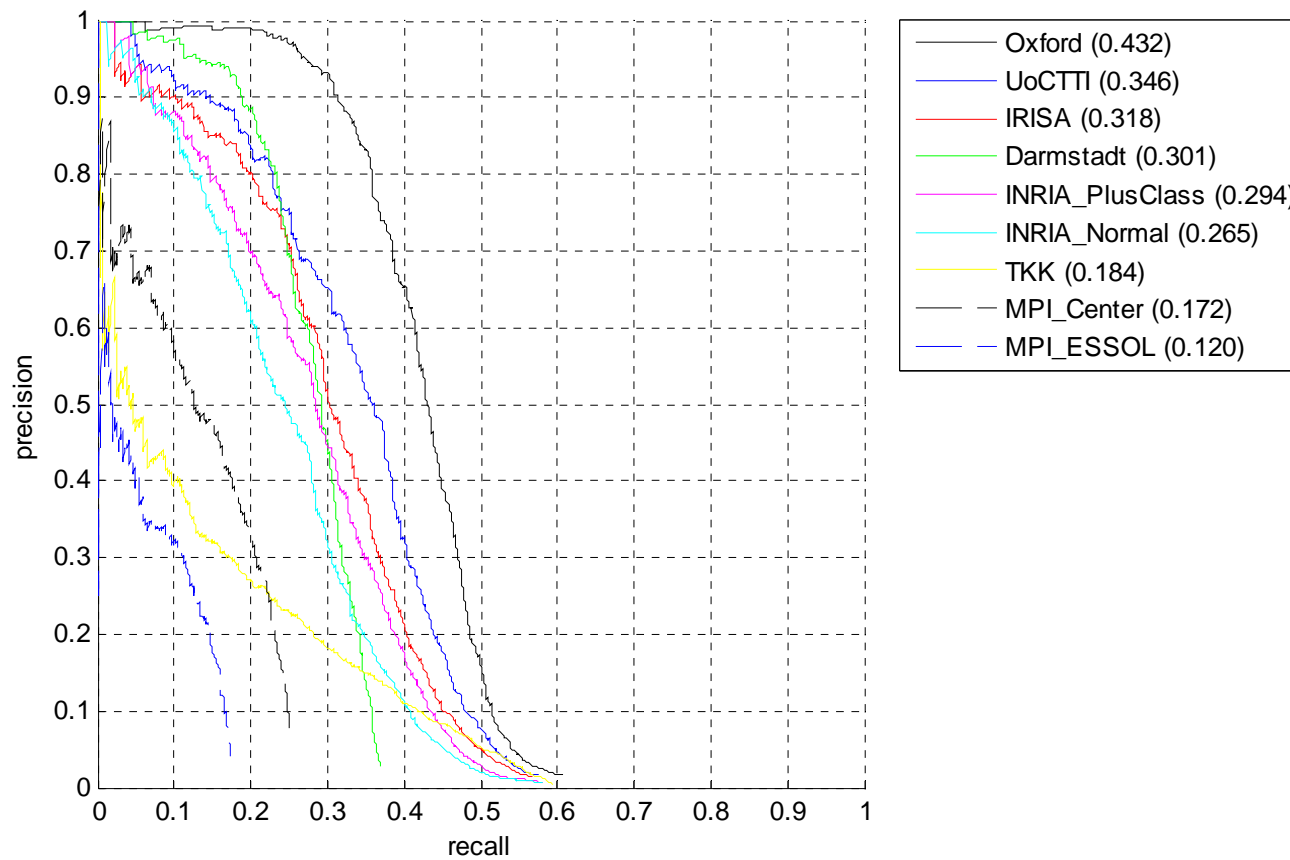
- Oxford method is convincingly best for 6 classes entered
- Best method for other classes varies greatly

Ranks by Method

	1st	2nd	3rd
Oxford	6	-	-
UoCTTI	6	8	1
MPI_ESSOL	5	0	0
IRISA	2	1	4
INRIA_PlusClass	1	6	5
INRIA_Normal	0	1	5
MPI_Center	0	2	3
TKK	0	2	3
Darmstadt	0	0	0

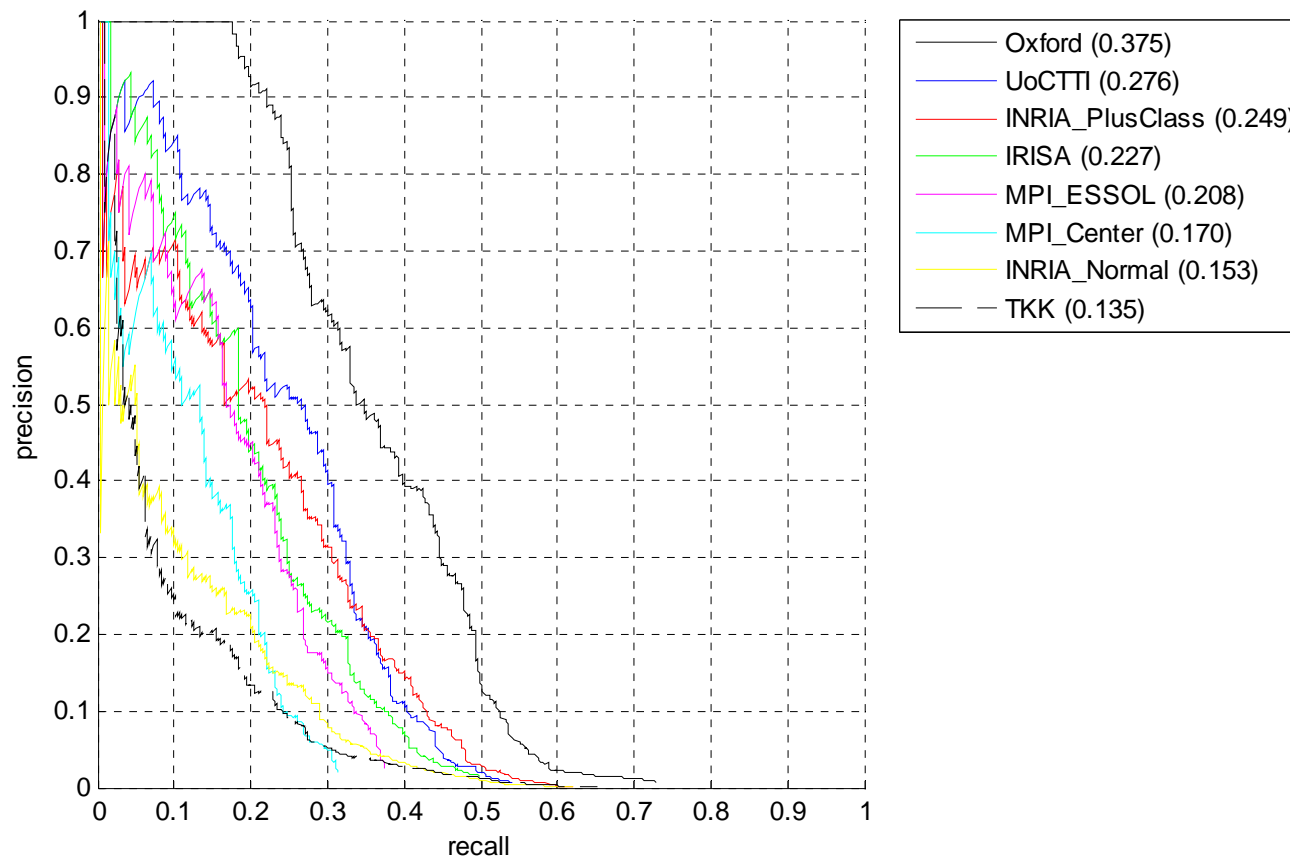
Example Precision/Recall: Car

- Car



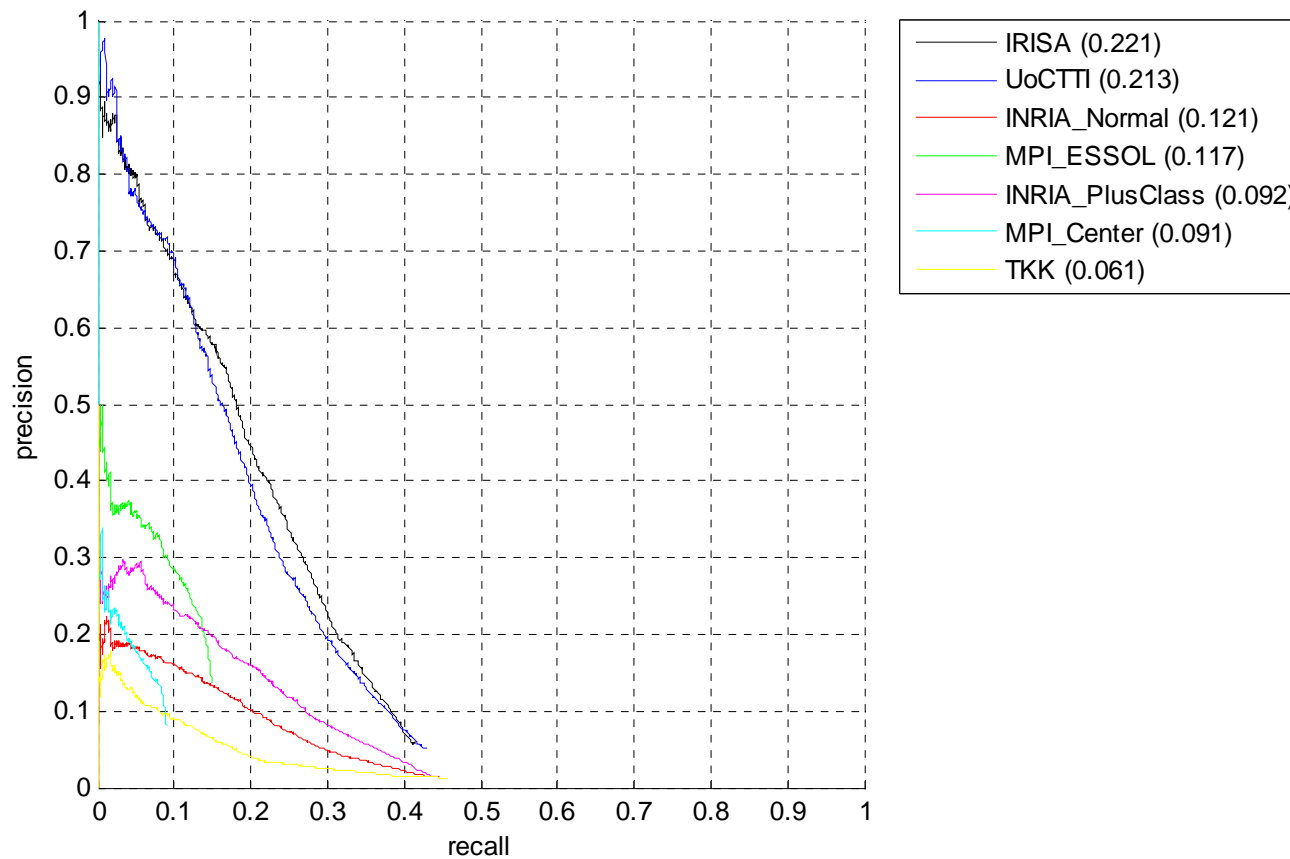
Example Precision/Recall: Motorbike

- Motorbike



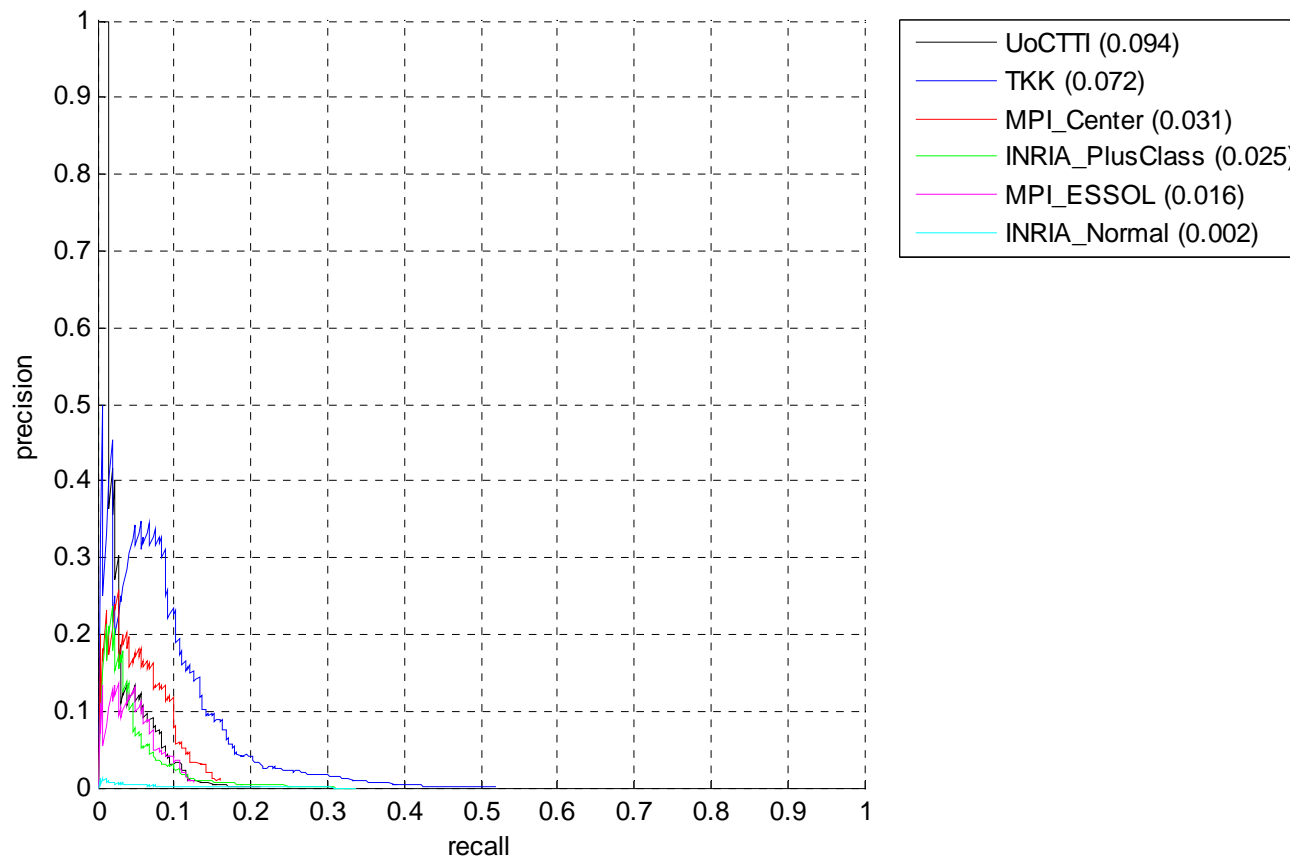
Example Precision/Recall: Person

- Person

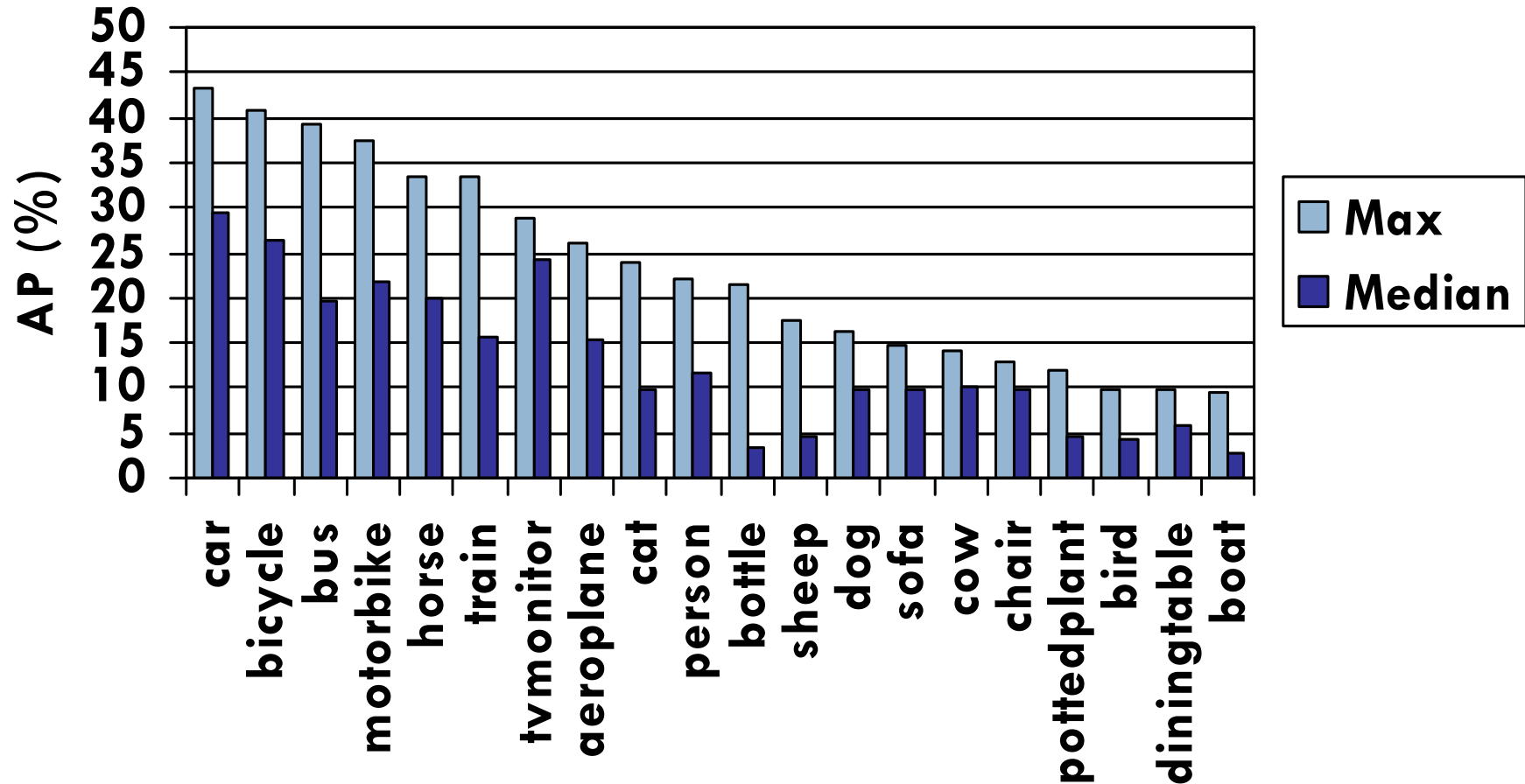


Example Precision/Recall: Boat

- Boat



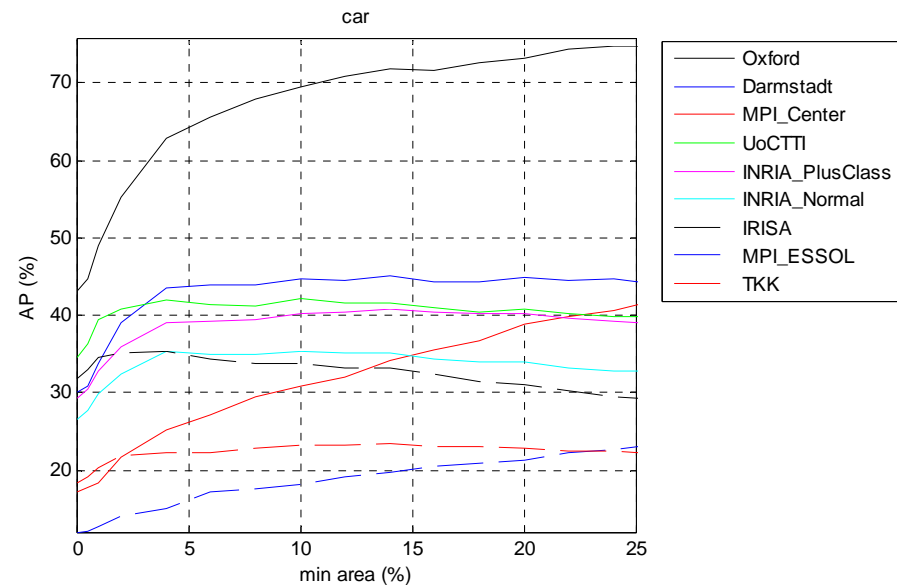
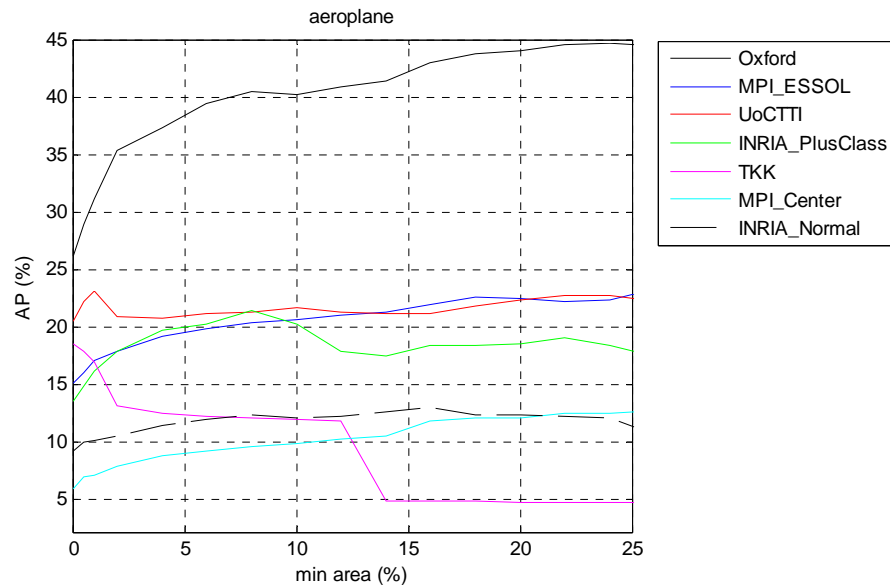
AP by Class



- Some counter-intuitive results e.g. good performance on “train” due to “whole image” methods?

AP vs. Object Area

- Do these methods have a bias toward larger objects?



- Oxford method greatly affected by object size, “whole image methods” somewhat, “sliding window” methods very little

VOC2006 vs. VOC2007 Test Data

		bike	bus	car	cat	cow	dog	horse	mbike	pers	sheep
Test on 2007	IRISA	28.1	-	31.8	2.6	11.9	-	28.9	22.7	22.1	17.5
	Oxford	40.9	39.3	43.2	-	-	-	-	37.5	-	-
	UoCTTI	36.9	23.2	34.6	9.8	14.0	2.3	18.2	27.6	21.3	14.3
Test on 2006	IRISA	35.2	-	48.2	9.4	20.9	-	18.3	33.3	21.1	26.2
	Oxford	56.8	36.0	53.5	-	-	-	-	53.9	-	-
	UoCTTI	56.2	23.6	55.5	10.3	21.2	9.9	17.3	43.9	26.2	22.1
VOC2006	Best	44.0	16.9	44.4	16.0	25.2	11.8	14.0	39.0	16.4	25.1

- High correlation between results on 2007 and 2006 test data
- For 7/10 classes, 2007 methods did better than the best 2006 method
 - Caveat: 2007 training data helped or hindered?

Prizes



- Winner: University of Oxford
- Honorable Mention:
 - University of Chicago/TTI Chicago (UoCTTI)
 - Max-Planck Institute Tuebingen (MPI_ESSOL)
 - IRISA/INRIA Rennes (IRISA)
 - INRIA Rhones-Alpes (INRIA_PlusClass)