The PASCAL Visual Object Classes Challenge 2012 (VOC2012)

Part I – Detection Challenge

Mark Everingham, Luc Van Gool
Chris Williams, John Winn
Andrew Zisserman
Yusuf Aytar, Ali Eslami

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

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?

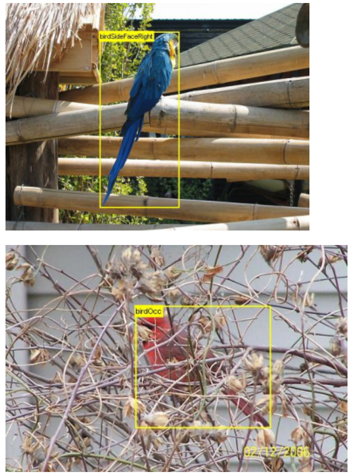
- Competition 4: Train on any (non-test) data
  - How well do state-of-the-art methods perform on these problems?

Examples

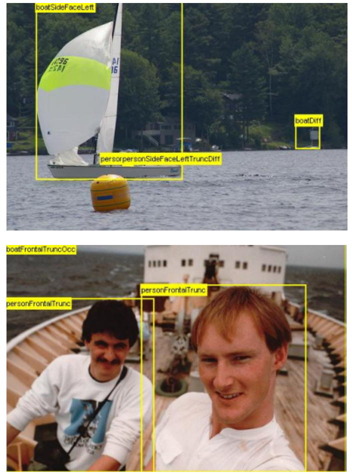
Aeroplane

Bicycle

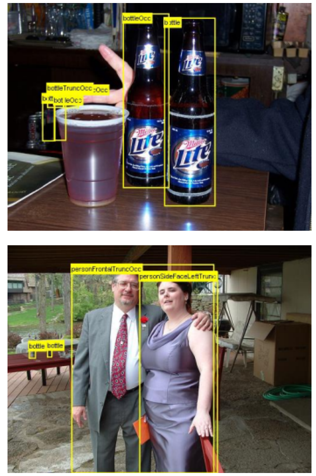
Bird



Boat

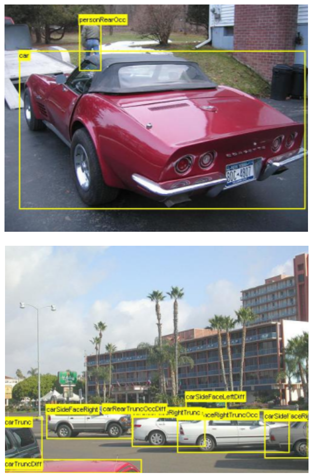


Bottle

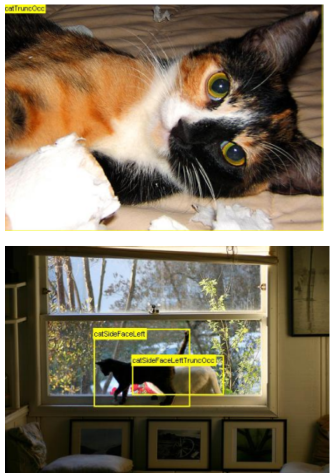


Bus

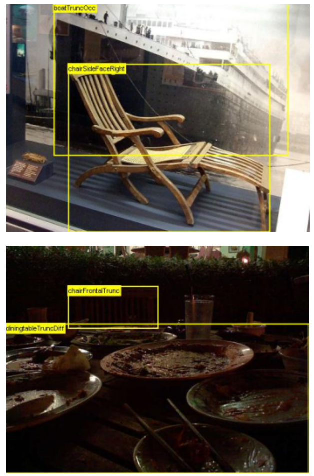
Car



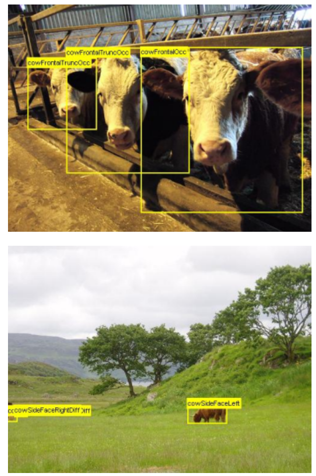
Cat



Chair

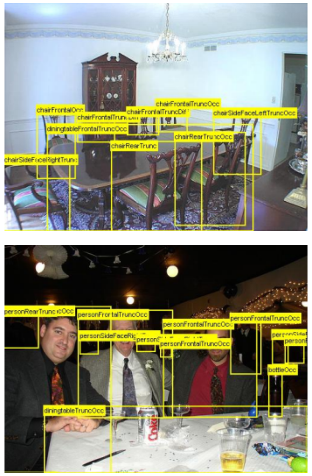


Cow



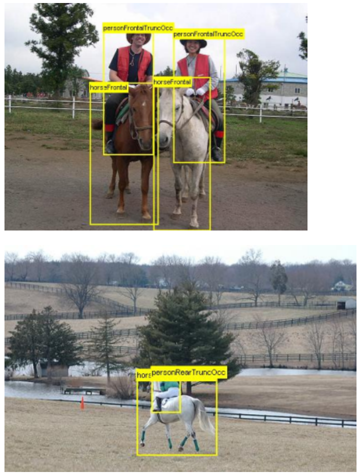
Examples

Dining Table

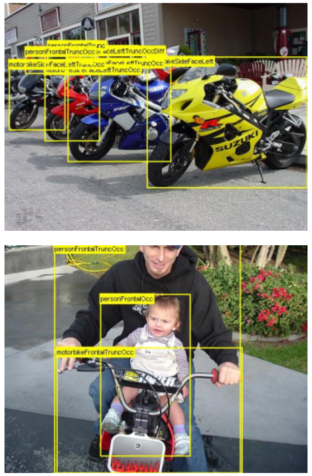


Dog

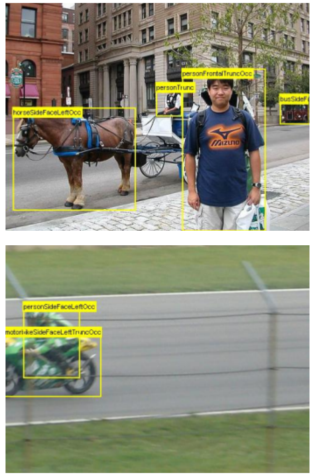
Horse



Motorbike

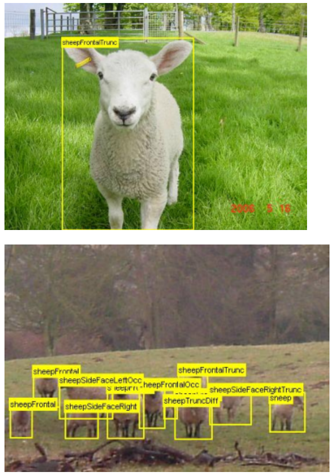


Person

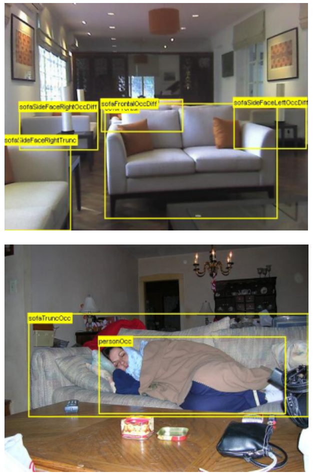


Potted Plant

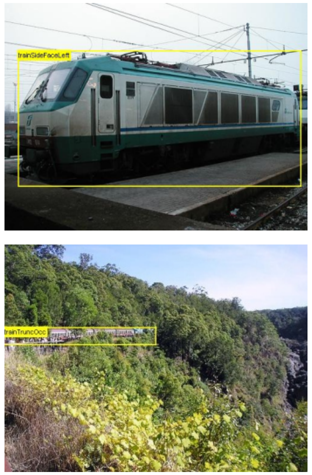
Sheep



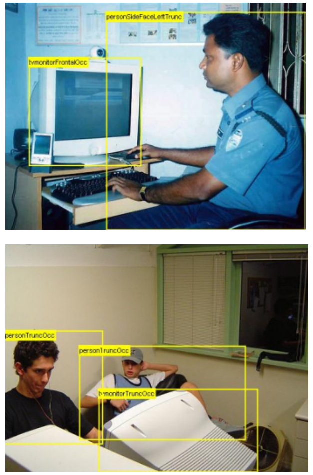
Sofa



Train



TV/Monitor

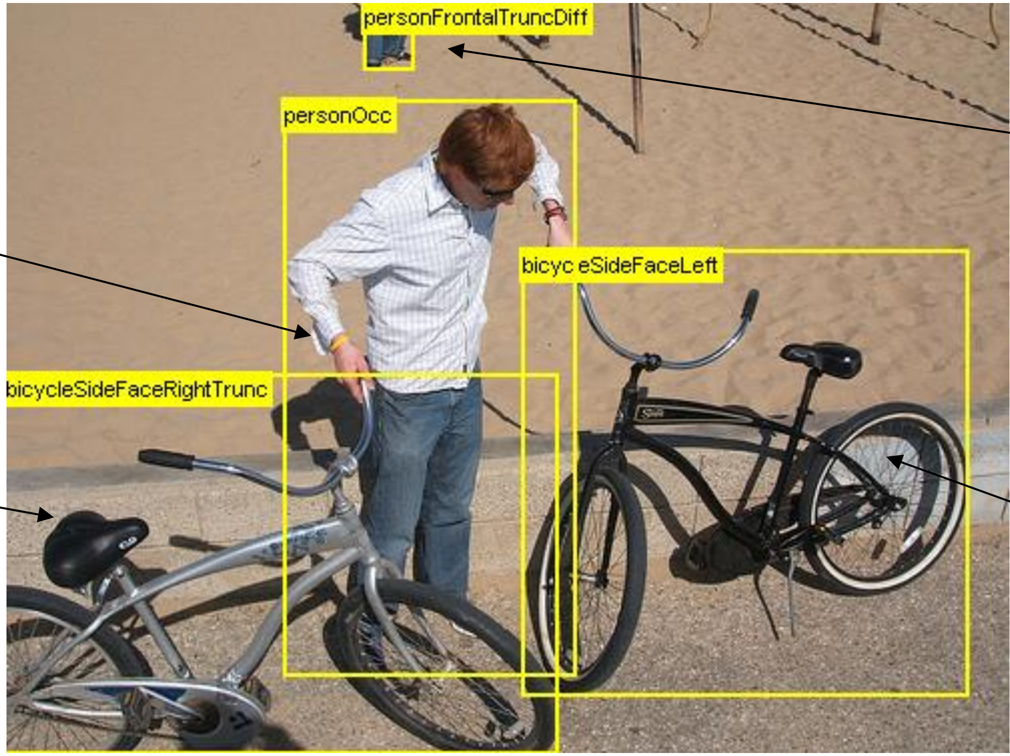


Annotation

- Complete annotation of objects from 20 categories

Occluded
Object is significantly occluded within BB

Truncated
Object extends beyond BB

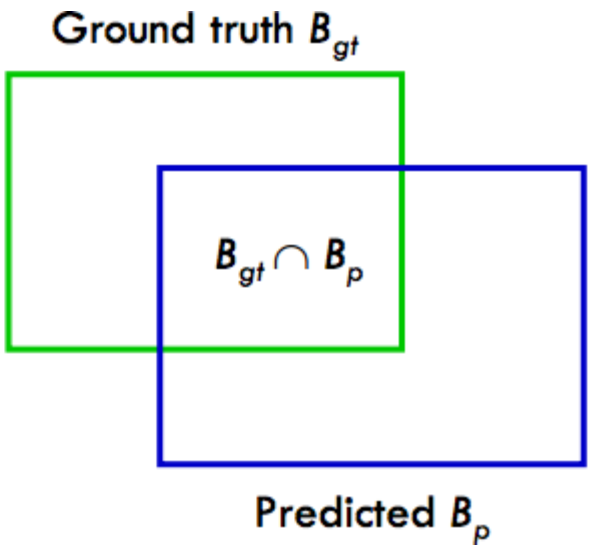


Difficult
Not scored in evaluation

Pose
Facing left

Evaluating bounding boxes

- Area of overlap (AO) measure



$$AO(B_{gt}, B_p) = \frac{|B_{gt} \bigcap B_p|}{|B_{gt} \bigcup B_p|}$$

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

Dataset statistics

- Same size as VOC2011.

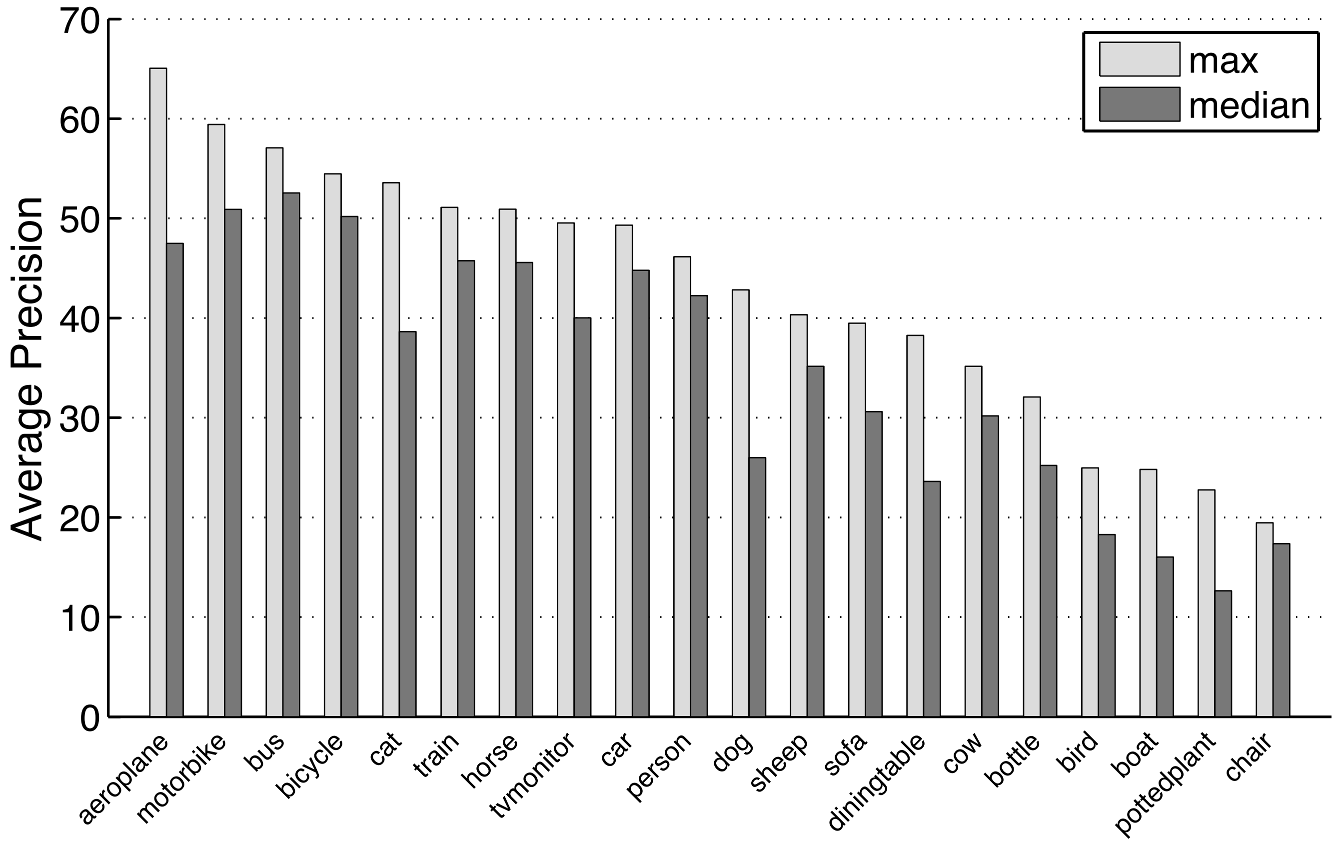
	Training	Testing
Images	11,540	10,994
Objects	27,450	27,078

- Minimum ~600 training objects per category
- ~2,000 cars, 1,500 dogs, 8,500 people
- Approximately equal distribution across training and test datasets

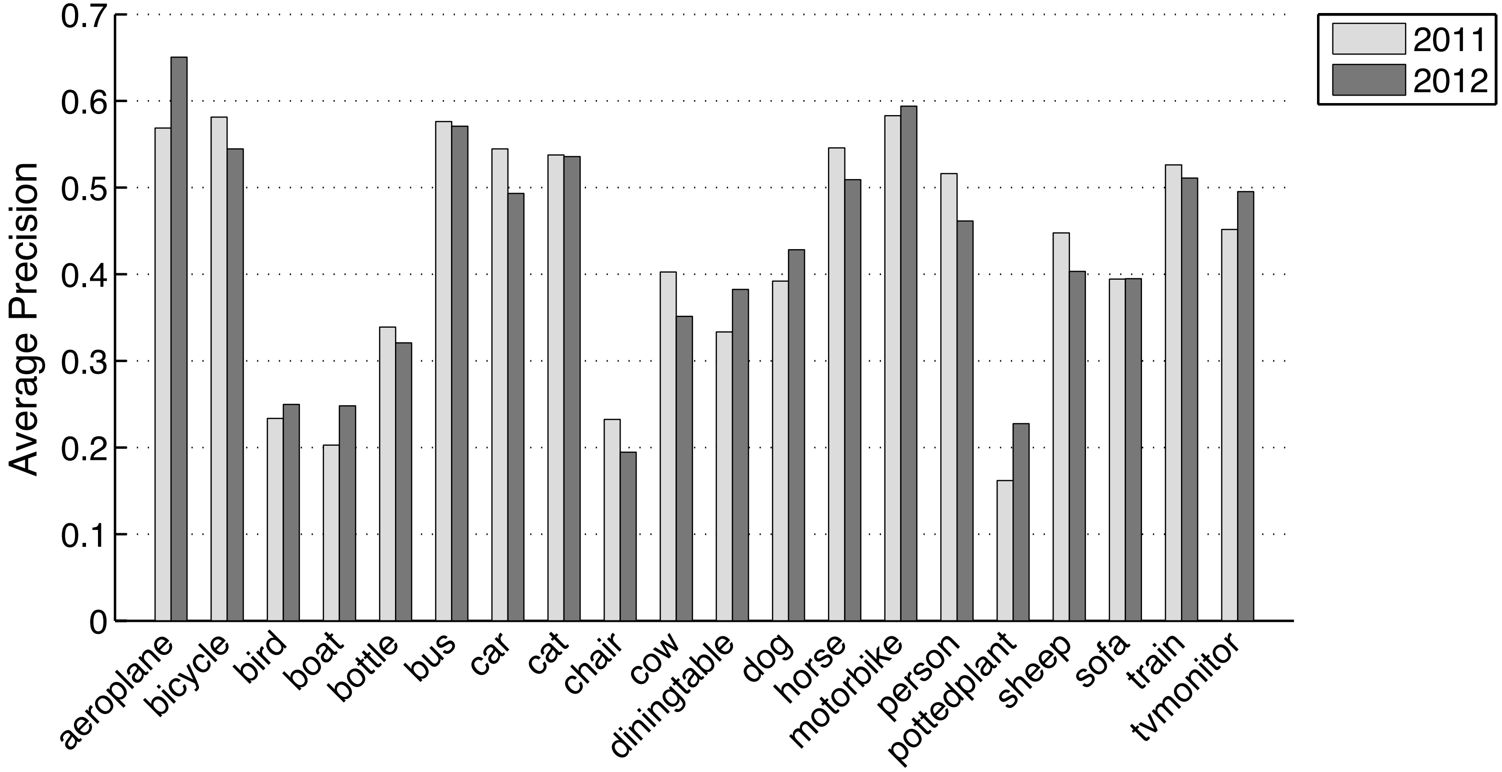
Submitted methods

- 8 methods, 7 groups
  - VOC2011: 13 methods, 15 groups
- Common approach:
  - Deformable Part Model (Girshick, Felzenszwalb, McAllester) with variations, e.g.
    - HOG-LBP features
    - Colour features
    - Multiple kernel learning
- New approaches:
  - Selective search (UVA, NEC_STANFORD)
  - Dynamic AND-OR tree

Average precision by class



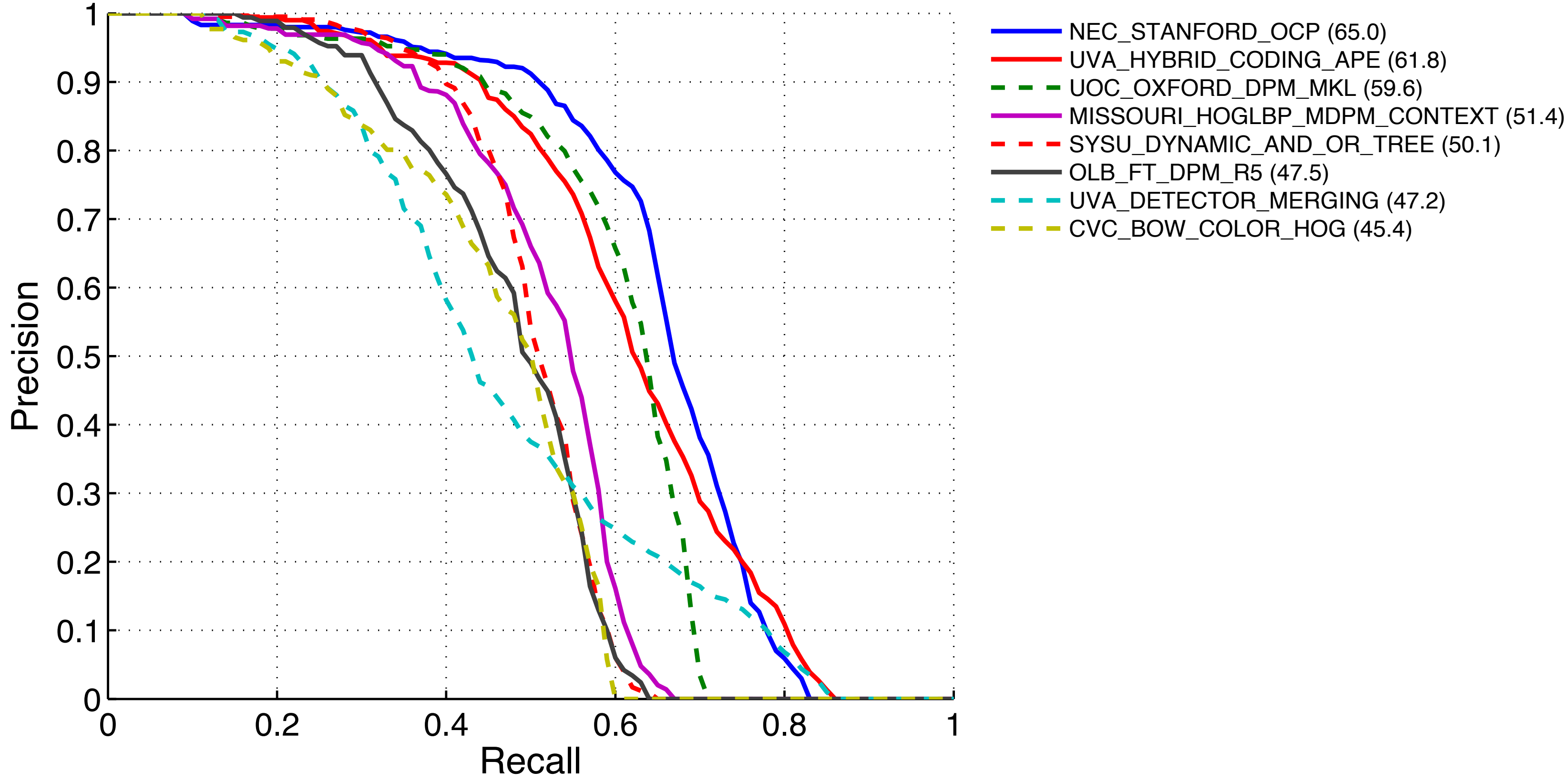
Improvement over VOC2011



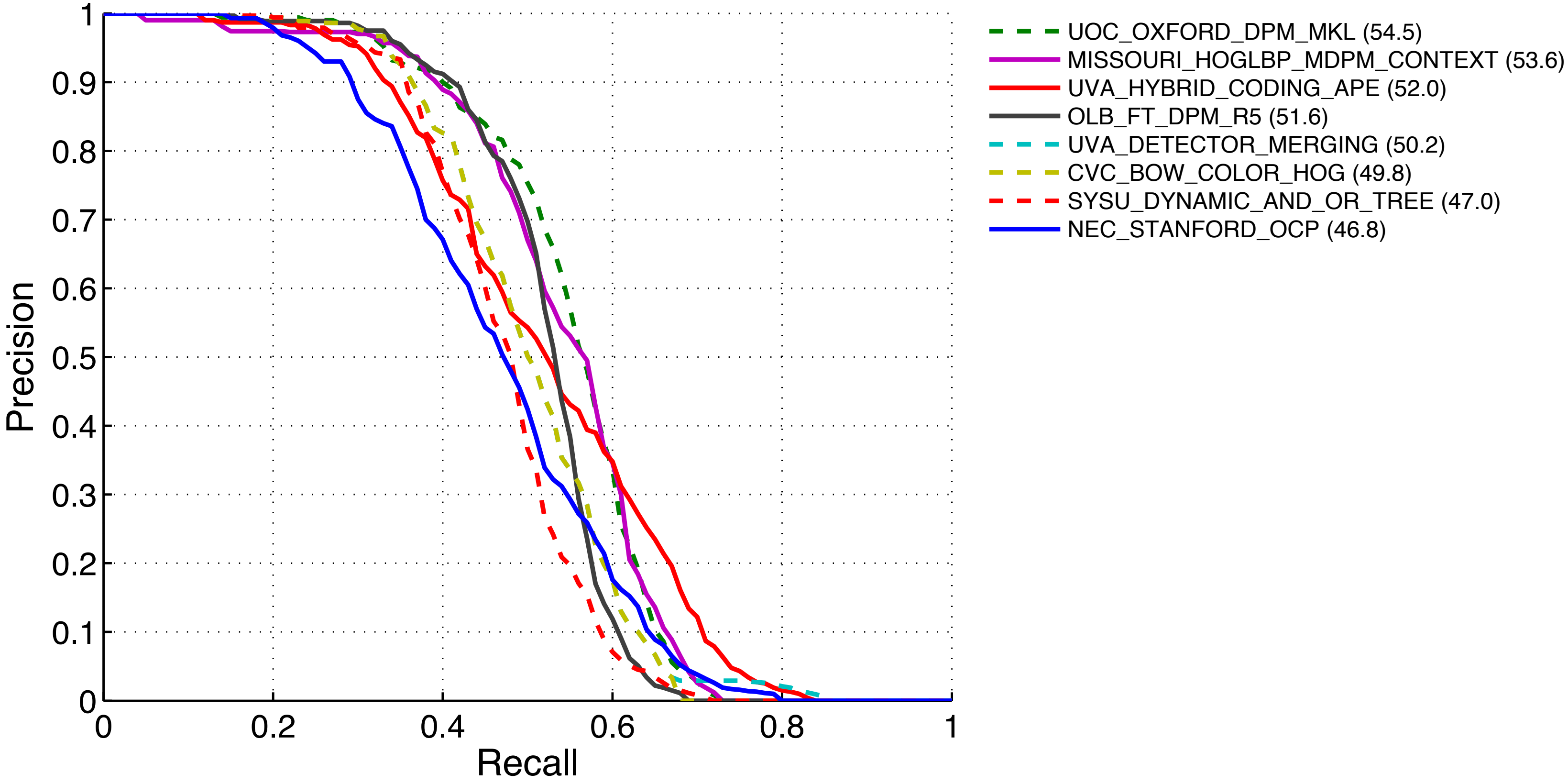
AP by class and method

	aeroplane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	diningtable	dog	horse	motorbike	person	pottedplant	sheep	sofa	train	tvmonitor
CVC_BOW_COLOR_HOG	45.4	49.8	15.7	16.0	26.3	54.6	44.8	35.1	16.8	31.3	23.6	26.0	45.6	49.6	42.2	14.5	30.5	28.5	45.7	40.0
MISSOURI_HOGLBP_MDPM_CONTEXT	51.4	53.7	18.3	15.6	31.6	56.5	47.1	38.6	19.5	32.0	22.1	25.0	50.3	51.9	44.9	11.9	37.7	30.6	50.8	39.3
NEC_STANFORD_OCP	65.1	46.8	25.0	24.6	16.0	51.0	44.9	51.5	13.0	26.6	31.0	40.2	39.7	51.5	32.8	12.6	35.7	33.5	48.0	44.8
OLB_FT_DPM_R5	47.5	51.7	14.2	12.6	27.3	51.8	44.2	25.3	17.8	30.2	18.1	16.9	46.9	50.9	43.0	9.5	31.2	23.6	44.3	22.1
SYSU_DYNAMIC_AND_OR_TREE	50.1	47.0	7.9	3.8	24.8	47.2	42.8	31.2	17.5	24.2	10.0	21.3	43.5	46.4	37.5	7.9	26.4	21.5	43.1	36.7
UOC_OXFORD_DPM_MKL	59.6	54.5	21.9	21.6	32.1	52.5	49.3	40.8	19.1	35.2	28.9	37.2	50.9	49.9	46.1	15.6	39.3	35.6	48.9	42.8
UVA_DETECTOR_MERGING	47.2	50.2	18.3	21.4	25.2	53.3	46.3	46.3	17.5	27.8	30.3	35	41.6	52.1	43.2	18	35.2	31.1	45.4	44.4
UVA_HYBRID_CODING_APE	61.8	52	24.6	24.8	20.2	57.1	44.5	53.6	17.4	33	38.3	42.8	48.8	59.4	35.7	22.8	40.3	39.5	51.1	49.5

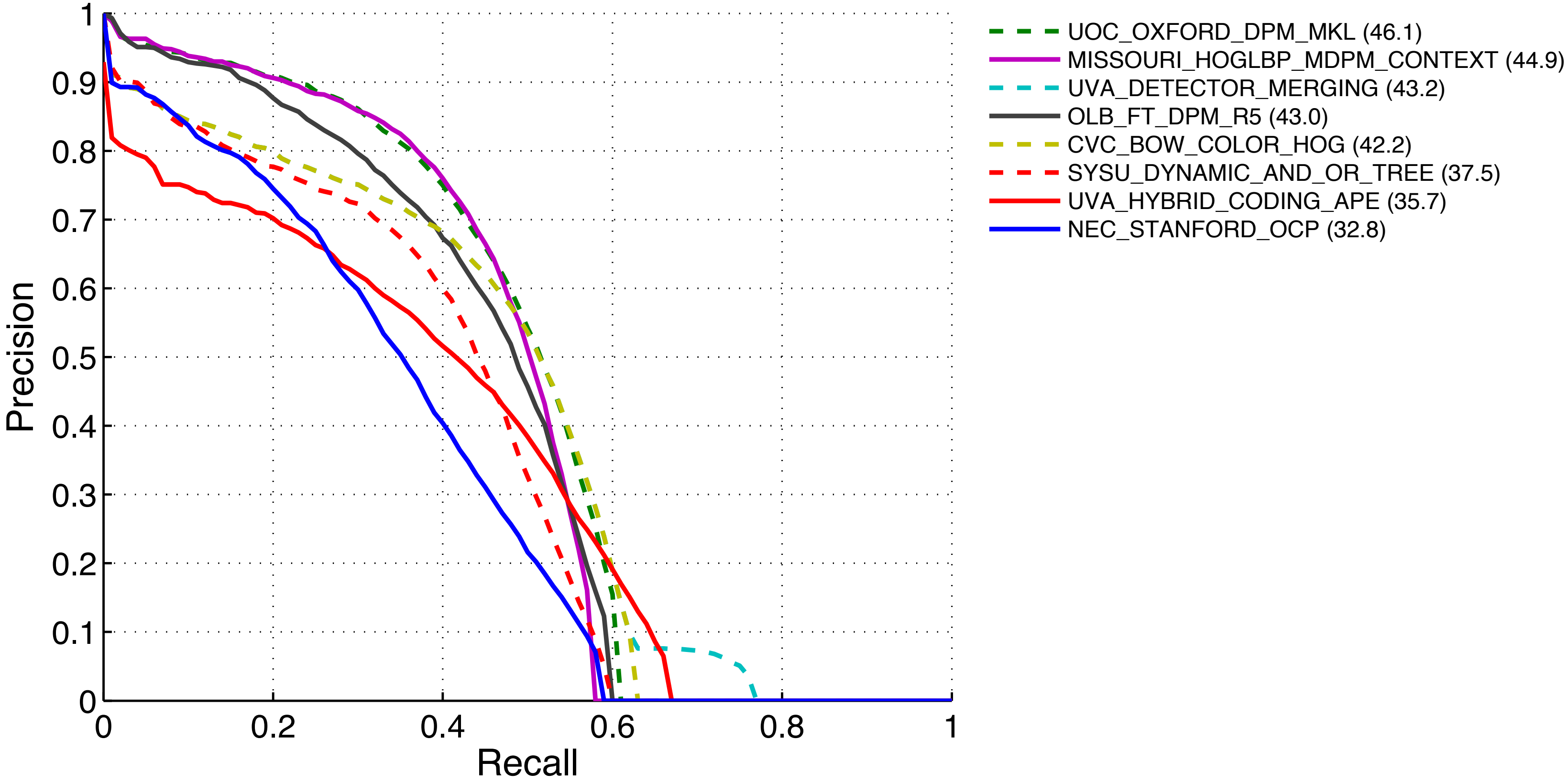
Precision/recall curves (aeroplane)



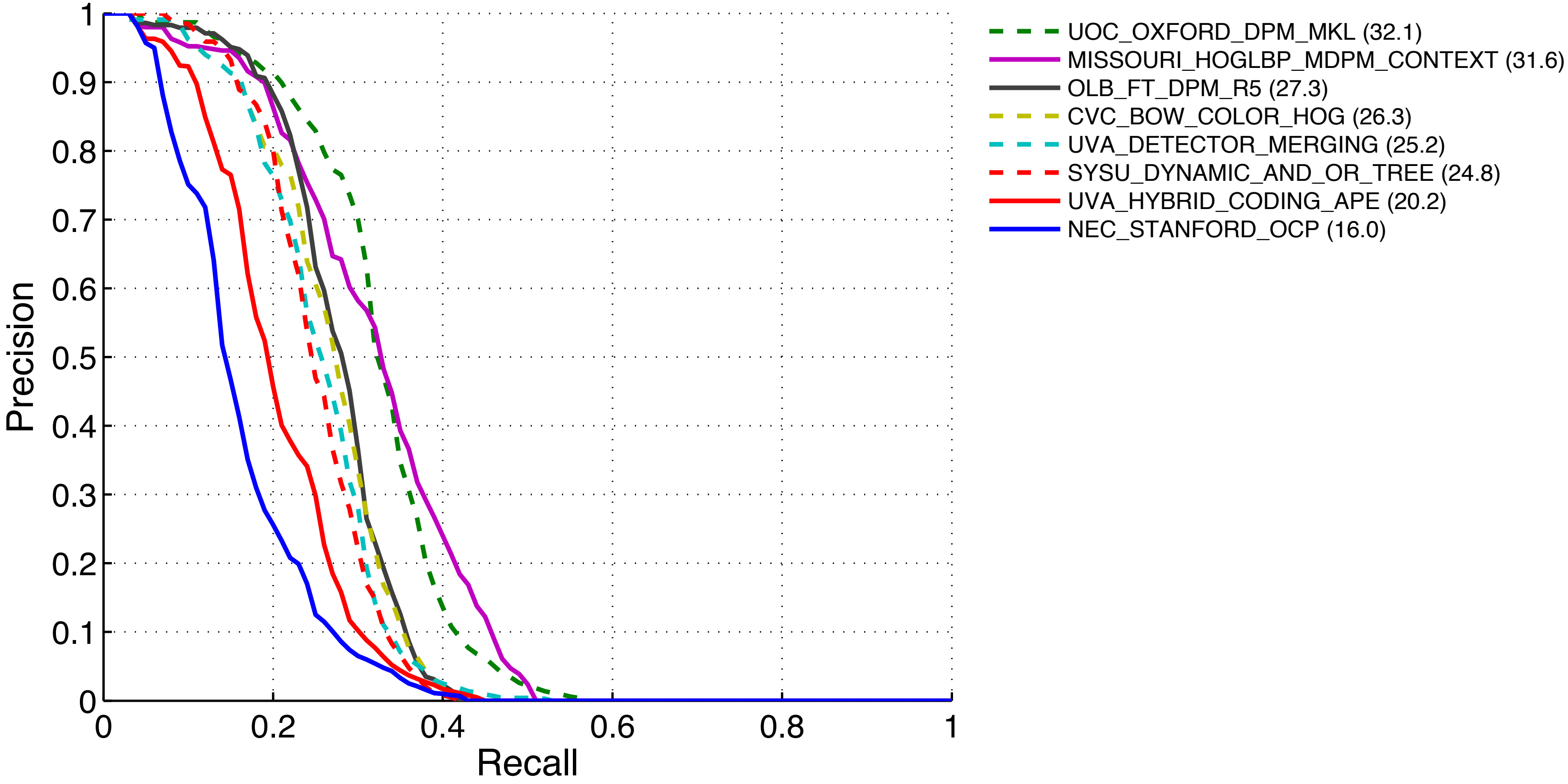
Precision/recall curves (bicycle)



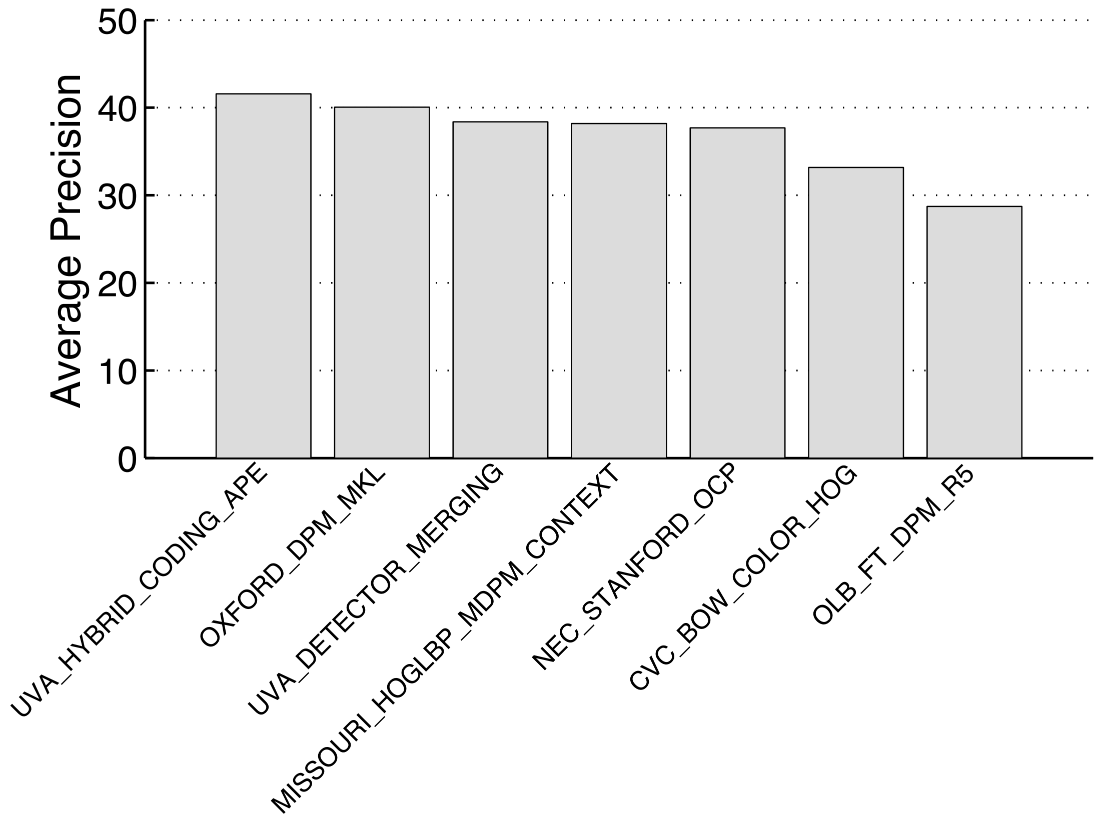
Precision/recall curves (person)



Precision/recall curves (bottle)



Median average precision by method



Prizes

- Winner
  - **UVA_HYBRID_CODING_APE**
    Koen E. A. van de Sande,
    Jasper R. R. Uijlings,
    Cees G. M. Snoek,
    Arnold W. M. Smeulders
    University of Amsterdam

- Honourable mention
  - **OXFORD_DPM_MKL**
    Ross Girshick, Andrea Vedaldi,
    Karen Simonyan
    University of Oxford

The PASCAL Visual Object Classes Challenge 2012 (VOC2012)

Part I – Detection Ranking Uncertainty

Mark Everingham, Luc Van Gool
Chris Williams, John Winn
Andrew Zisserman
Yusuf Aytar, Ali Eslami

PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning

Ranking uncertainty

- Only one AP curve per class and method
- However, we can use bootstrap resampling to obtain multiple AP curves (see e.g. blog post by Brendan O'Connor, 2010)
- Compare AP or rank of two methods **A** and **B**
- Can obtain a confidence interval for AP
- If `rank(`**A**`) < rank(`**B**`)` with high probability then **A** is significantly different from **B**

Ranking uncertainty

for each replication

1. sample a subset of the test images
2. compute AP of each submission on sample
3. compute rank of each submission based on APs

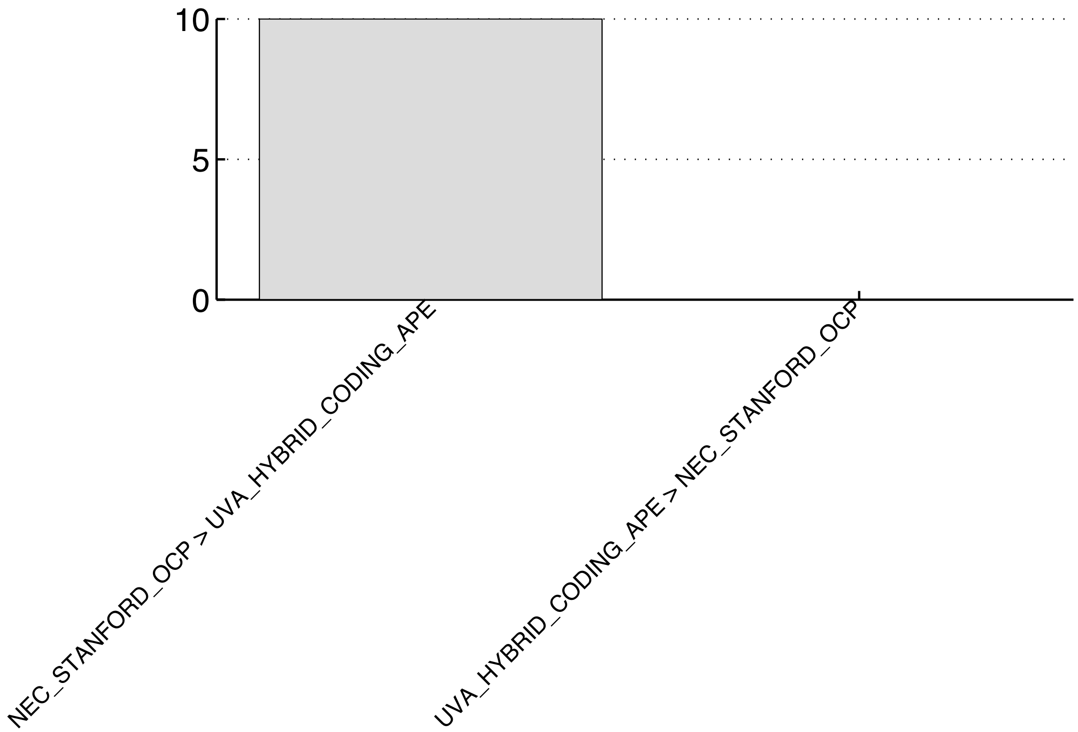
for each pair m^1 and m^2

1. m^1 and m^2 equivalent if rank of one method is not higher than the rank of the other in at least in 95% of replications

Equivalencies by class and method

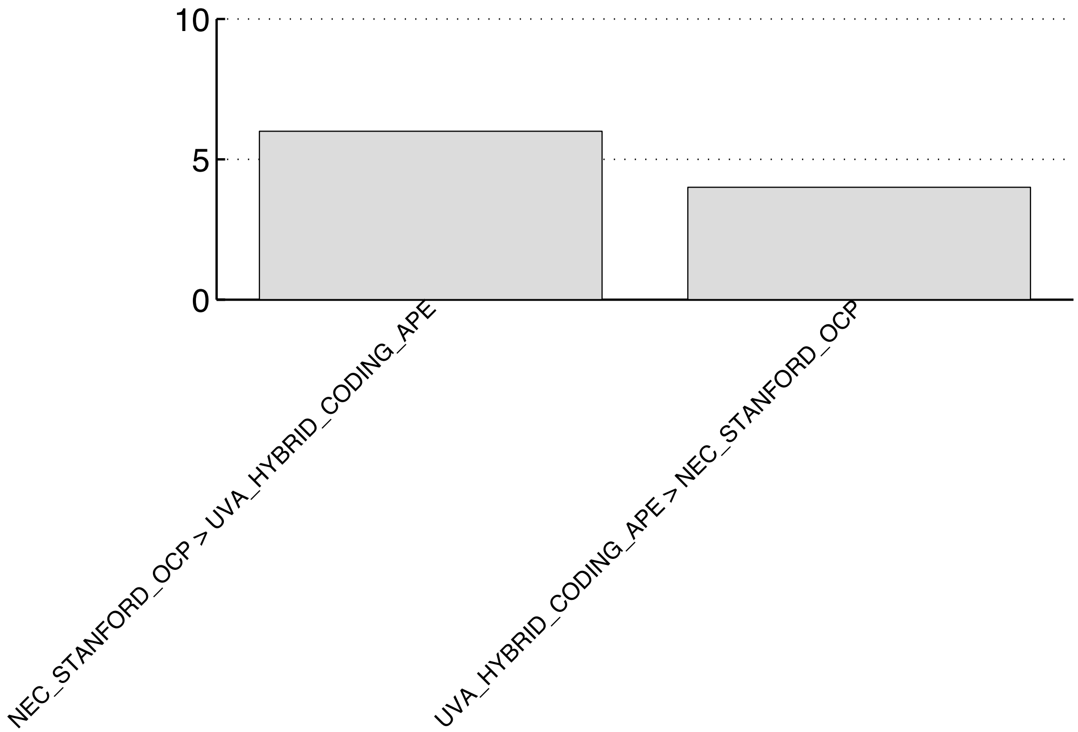
	aeroplane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	diningtable	dog	horse	motorbike	person	pottedplant	sheep	sofa	train	tvmonitor
CVC_BOW_COLOR_HOG	45.4	49.8	15.7	16.0	26.3	54.6	44.8	35.1	16.8	31.3	23.6	26.0	45.6	49.6	42.2	14.5	30.5	28.5	45.7	40.0
MISSOURI_HOGLBP_MDPM_CONTEXT	51.4	53.7	18.3	15.6	31.6	56.5	47.1	38.6	19.5	32.0	22.1	25.0	50.3	51.9	44.9	11.9	37.7	30.6	50.8	39.3
NEC_STANFORD_OCP	65.1	46.8	25.0	24.6	16.0	51.0	44.9	51.5	13.0	26.6	31.0	40.2	39.7	51.5	32.8	12.6	35.7	33.5	48.0	44.8
OLB_FT_DPM_R5	47.5	51.7	14.2	12.6	27.3	51.8	44.2	25.3	17.8	30.2	18.1	16.9	46.9	50.9	43.0	9.5	31.2	23.6	44.3	22.1
SYSU_DYNAMIC_AND_OR_TREE	50.1	47.0	7.9	3.8	24.8	47.2	42.8	31.2	17.5	24.2	10.0	21.3	43.5	46.4	37.5	7.9	26.4	21.5	43.1	36.7
UOC_OXFORD_DPM_MKL	59.6	54.5	21.9	21.6	32.1	52.5	49.3	40.8	19.1	35.2	28.9	37.2	50.9	49.9	46.1	15.6	39.3	35.6	48.9	42.8
UVA_DETECTOR_MERGING	47.2	50.2	18.3	21.4	25.2	53.3	46.3	46.3	17.5	27.8	30.3	35	41.6	52.1	43.2	18	35.2	31.1	45.4	44.4
UVA_HYBRID_CODING_APE	61.8	52	24.6	24.8	20.2	57.1	44.5	53.6	17.4	33	38.3	42.8	48.8	59.4	35.7	22.8	40.3	39.5	51.1	49.5

Equivalencies by class and method



Difference is statistically significant

Equivalencies by class and method



Difference is not statistically significant

Equivalencies by class and method

	aeroplane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	diningtable	dog	horse	motorbike	person	pottedplant	sheep	sofa	train	tvmonitor
CVC_BOW_COLOR_HOG	45.4	49.8	15.7	16.0	26.3	54.6	44.8	35.1	16.8	31.3	23.6	26.0	45.6	49.6	42.2	14.5	30.5	28.5	45.7	40.0
MISSOURI_HOGLBP_MDPM_CONTEXT	51.4	53.7	18.3	15.6	**31.6**	**56.5**	47.1	38.6	**19.5**	32.0	22.1	25.0	**50.3**	51.9	**44.9**	11.9	37.7	30.6	**50.8**	39.3
NEC_STANFORD_OCP	**65.1**	46.8	**25.0**	**24.6**	16.0	51.0	44.9	51.5	13.0	26.6	31.0	40.2	39.7	51.5	32.8	12.6	35.7	33.5	**48.0**	44.8
OLB_FT_DPM_R5	47.5	51.7	14.2	12.6	27.3	51.8	44.2	25.3	17.8	30.2	18.1	16.9	46.9	50.9	43.0	9.5	31.2	23.6	44.3	22.1
SYSU_DYNAMIC_AND_OR_TREE	50.1	47.0	7.9	3.8	24.8	47.2	42.8	31.2	17.5	24.2	10.0	21.3	43.5	46.4	37.5	7.9	26.4	21.5	43.1	36.7
UOC_OXFORD_DPM_MKL	59.6	**54.5**	21.9	21.6	**32.1**	52.5	**49.3**	40.8	**19.1**	**35.2**	28.9	37.2	**50.9**	49.9	**46.1**	15.6	**39.3**	35.6	**48.9**	42.8
UVA_DETECTOR_MERGING	47.2	50.2	18.3	21.4	25.2	53.3	46.3	46.3	17.5	27.8	30.3	35	41.6	52.1	43.2	18	35.2	31.1	45.4	44.4
UVA_HYBRID_CODING_APE	61.8	52	**24.6**	**24.8**	20.2	**57.1**	44.5	**53.6**	17.4	33	**38.3**	**42.8**	48.8	**59.4**	35.7	**22.8**	**40.3**	**39.5**	**51.1**	**49.5**