The PASCAL Visual Object Classes Challenge 2012 (VOC2012)

Part I – Classification Challenge

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

PASCAL2
Pattern Analysis, Statistical Modeling and Computational Learning

Classification challenge

- Predict whether at least one object of a given class is present in an image

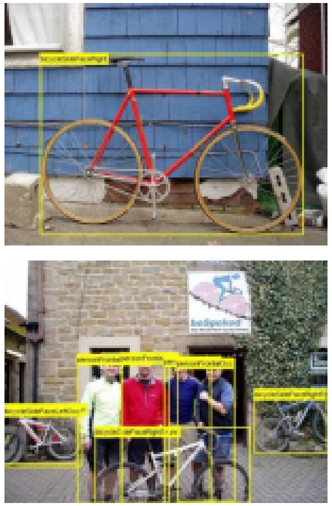
- Competition 1: Train on the supplied data
  - Which methods perform best given specified training data?

- Competition 2: Train on any (non-test) data
  - How well do state-of-the-art methods perform on these problems?

20 classes – examples

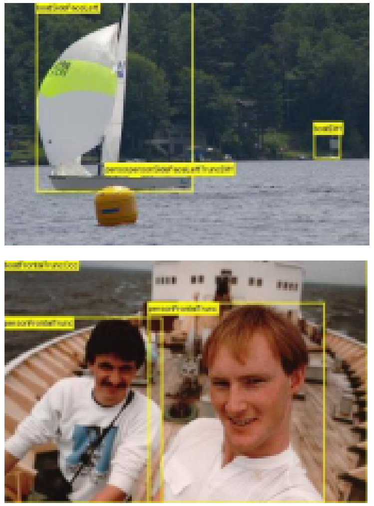
Aeroplane

Bicycle

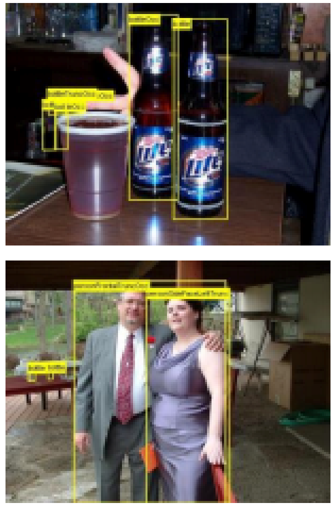


Bird

Boat

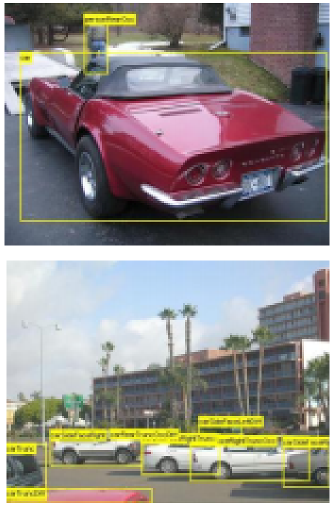


Bottle



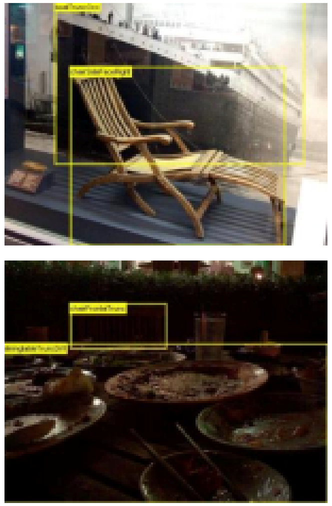
Bus

Car

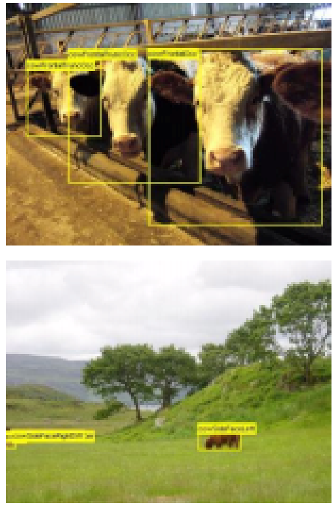


Cat

Chair



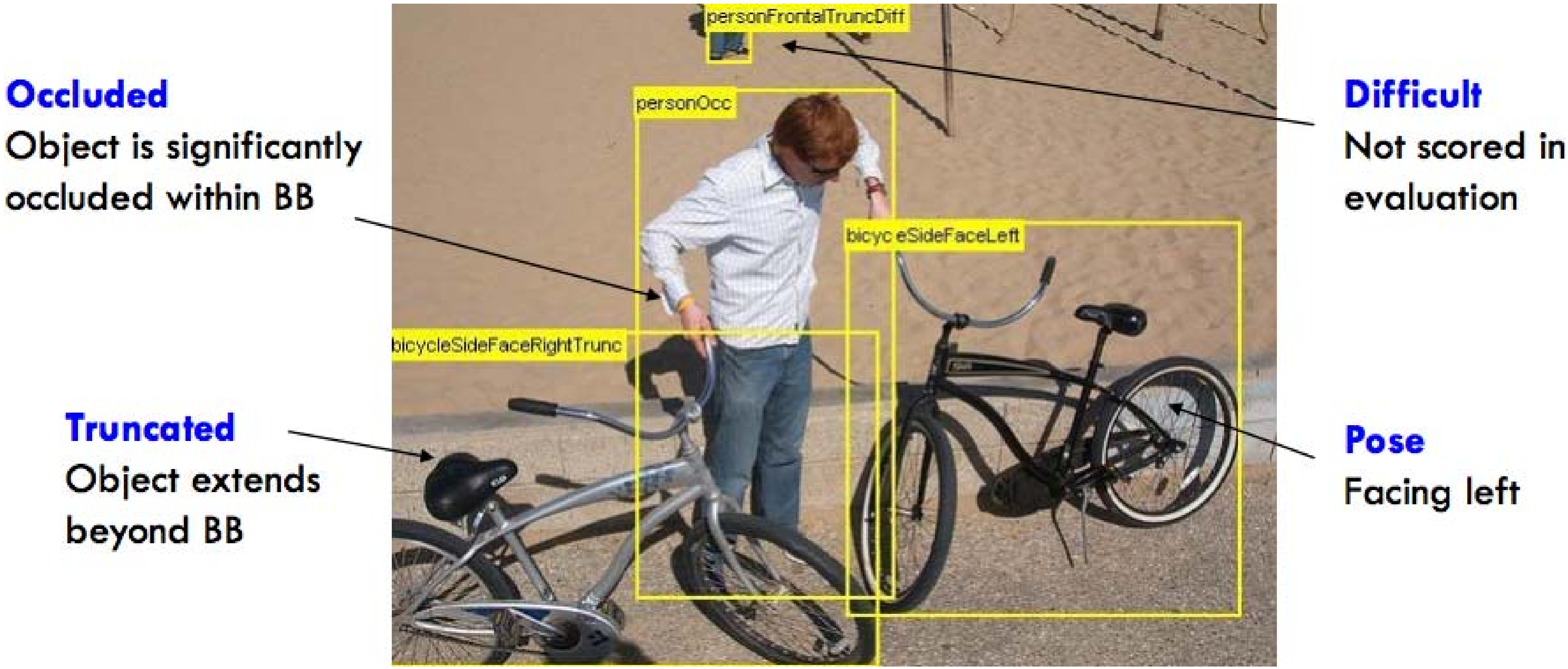
Cow



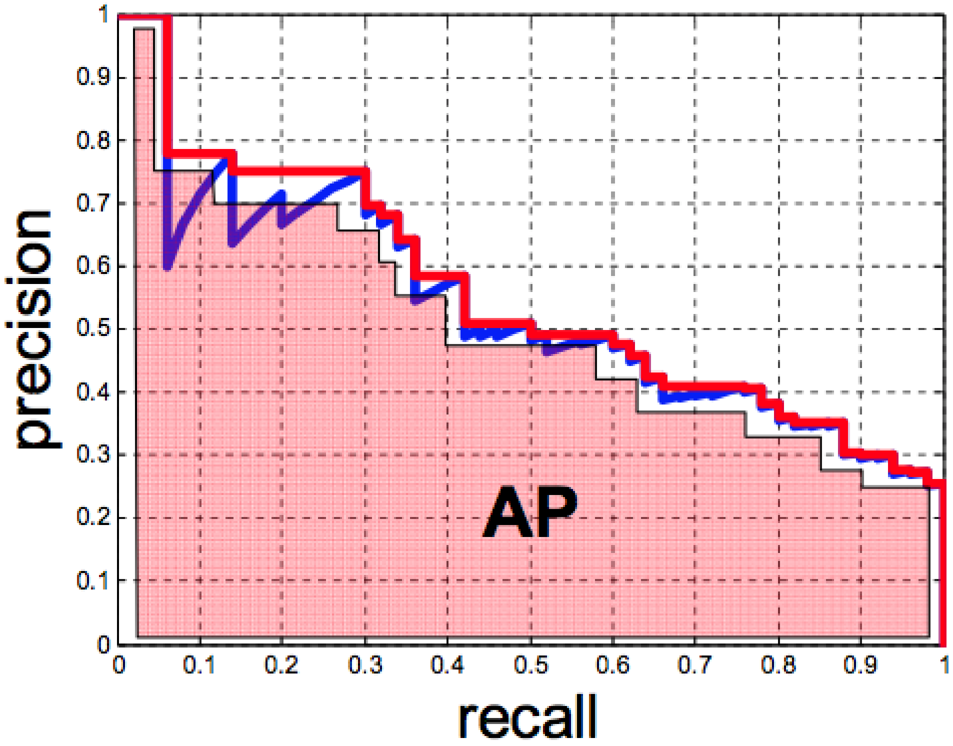
20 classes – examples

Annotation

- Complete annotation of objects from 20 categories



Average precision: VOC2010-2012



- Interpolate curve to create version for which the precision is monotonically non- increasing
- Measure area under interpolated curve

- Sawtooth shape is ignored
- Area is measured with maximum accuracy

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

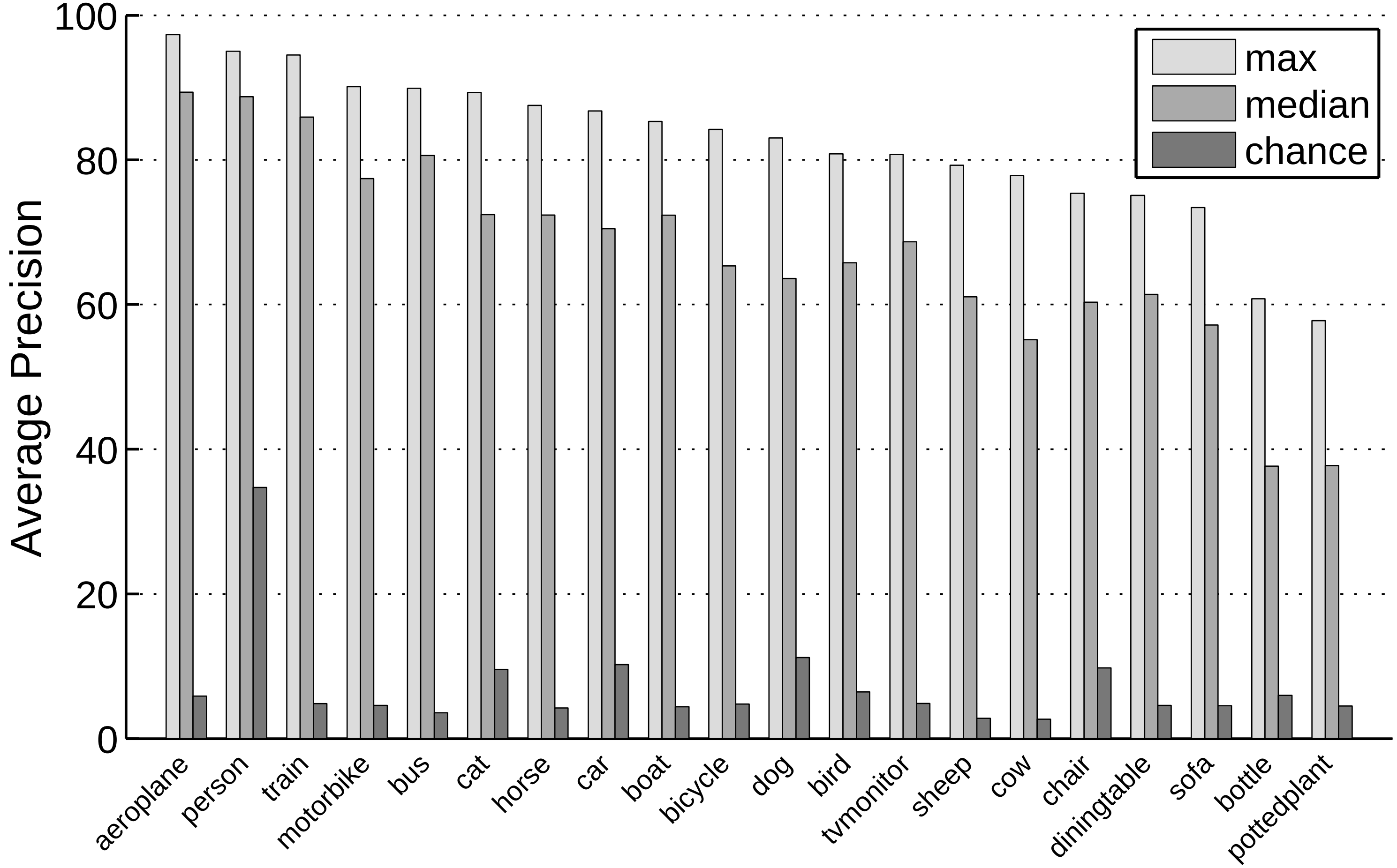
Submitted methods

- 7 methods, 5 groups
  - VOC 2011: 19 methods, 11 groups

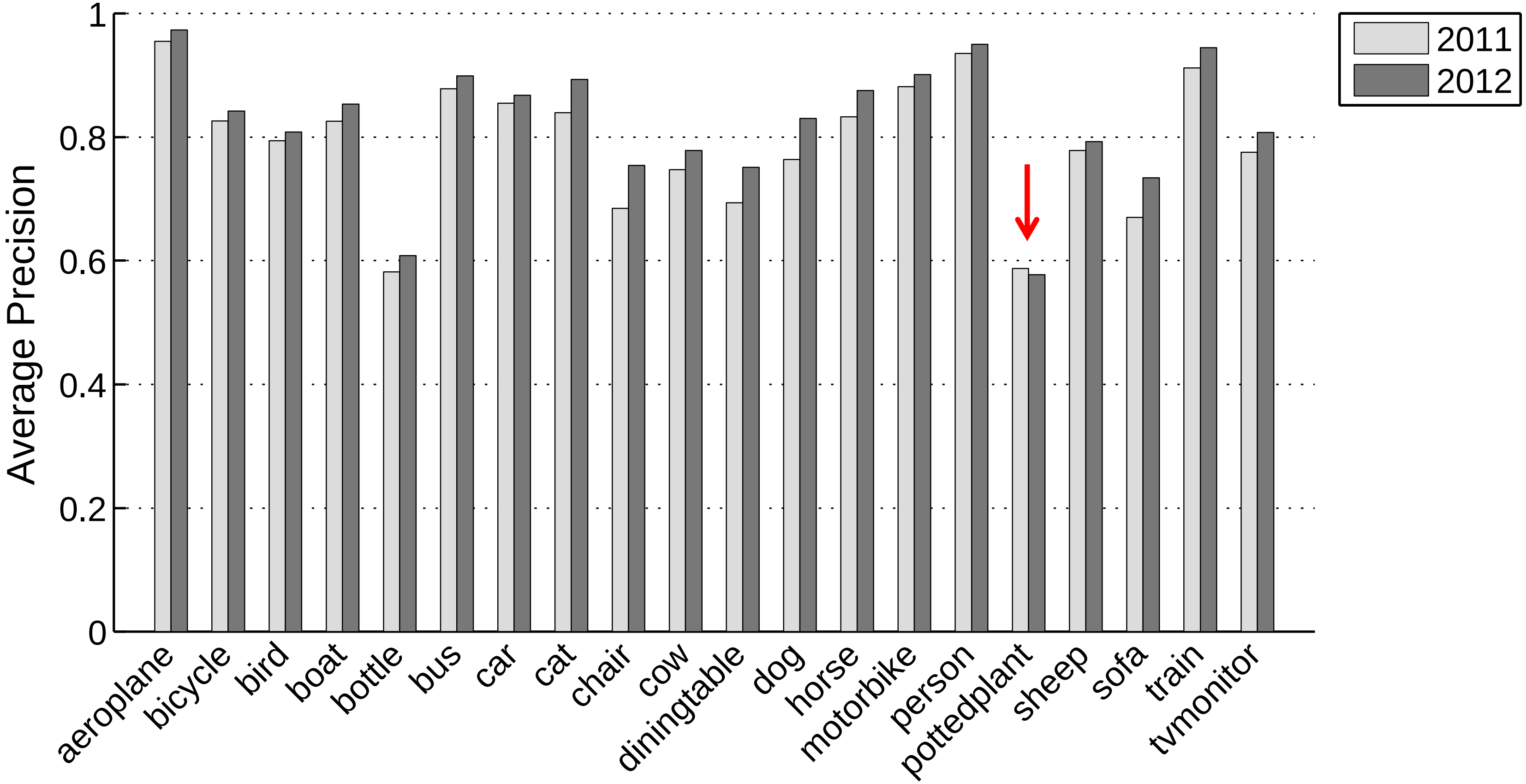
- Basic recipe:
  - Features: Dense SIFT, HOG, colour
  - Encodings: spatial pyramid, BOW, Fisher vector
  - Detectors: DPM
  - Classifier: SVM

- Additional ingredients:
  - Complex log-normal features
  - Sub-clusters for classes
  - Combinations of clusterings and projections

Average precision by class



Improvement over VOC2011



AP by class and method

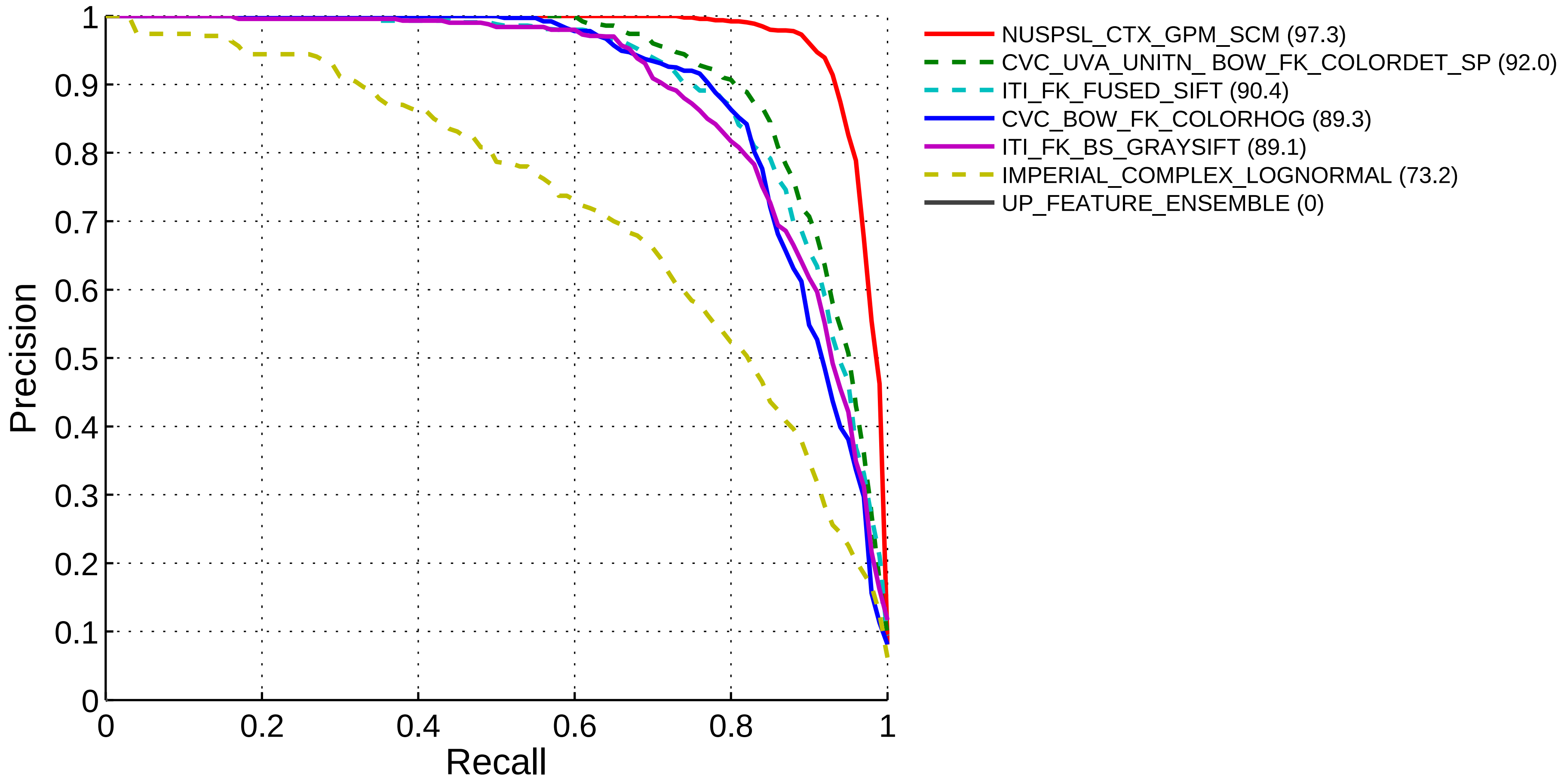
Trained on VOC 2012 data

	aeroplane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	diningtable	dog	horse	motorbike	person	pottedplant	sheep	sofa	train	tvmonitor
CVC_BOW_FK_COLORHOG	89.3	70.9	69.8	73.9	51.3	84.8	79.6	72.9	63.8	59.4	64.1	64.7	75.5	79.2	91.4	42.7	63.2	61.9	86.7	73.8
CVC_UVA_UNITN_ BOW_FK_COLORDET_SP	92.0	74.2	73.0	77.5	54.3	85.2	81.9	76.4	65.2	63.2	68.5	68.9	78.2	81.0	91.6	55.9	69.4	65.4	86.7	77.4
IMPERIAL_COMPLEX_LOGNORMAL	73.2	33.4	31.0	44.7	17.0	57.7	34.4	45.9	41.2	18.1	30.2	34.3	23.1	39.3	57.3	11.9	23.1	25.3	51.2	36.2
ITI_FK_BS_GRAYSIFT	89.1	62.3	60.0	68.1	33.4	79.8	66.9	70.3	57.4	51.0	55.0	59.3	68.6	74.5	83.1	25.6	57.2	53.8	83.4	64.9
ITI_FK_FUSED_SIFT	90.4	65.4	65.8	72.3	37.7	80.6	70.5	72.4	60.3	55.1	61.4	63.6	72.4	77.4	86.8	37.7	61.1	57.2	85.9	68.7
NUSPSL_CTX_GPM_SCM	97.3	84.2	80.8	85.3	60.8	89.9	86.8	89.3	75.4	77.8	75.1	83.0	87.5	90.1	95.0	57.8	79.2	73.4	94.5	80.7
UP_FEATURE_ENSEMBLE	-	-	-	-	-	-	-	-	-	-	-	-	-	-	88.7	-	-	-	-	-

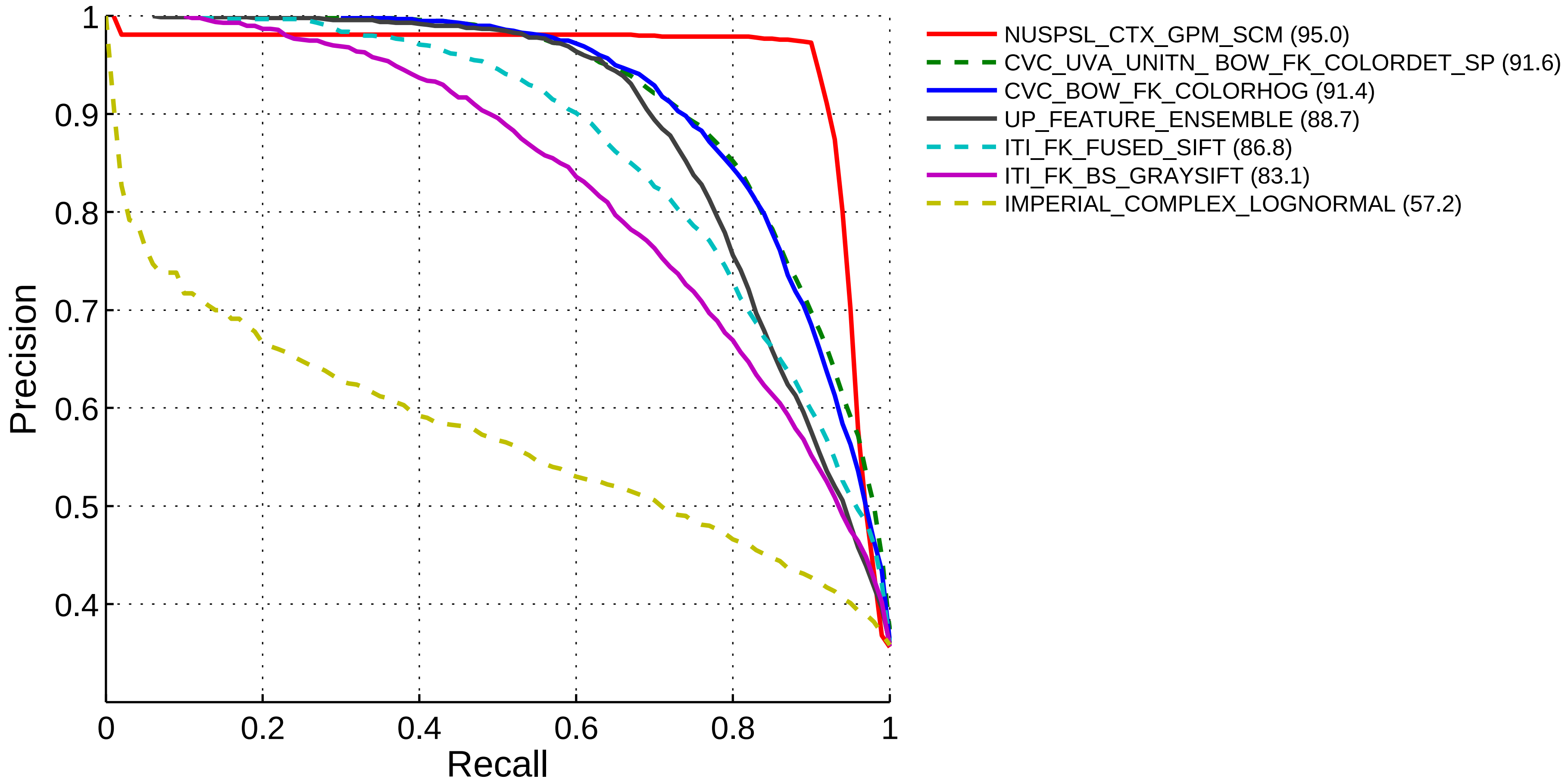
Trained on external data

	aeroplane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	diningtable	dog	horse	motorbike	person	pottedplant	sheep	sofa	train	tvmonitor
ITI_FK_FLICKR_GRAYSIFT_ENTROPY	88.1	63.0	61.9	68.6	34.9	79.6	67.4	70.5	57.5	52.0	55.3	60.1	68.7	74.3	83.2	26.4	57.6	53.4	83.0	64.0

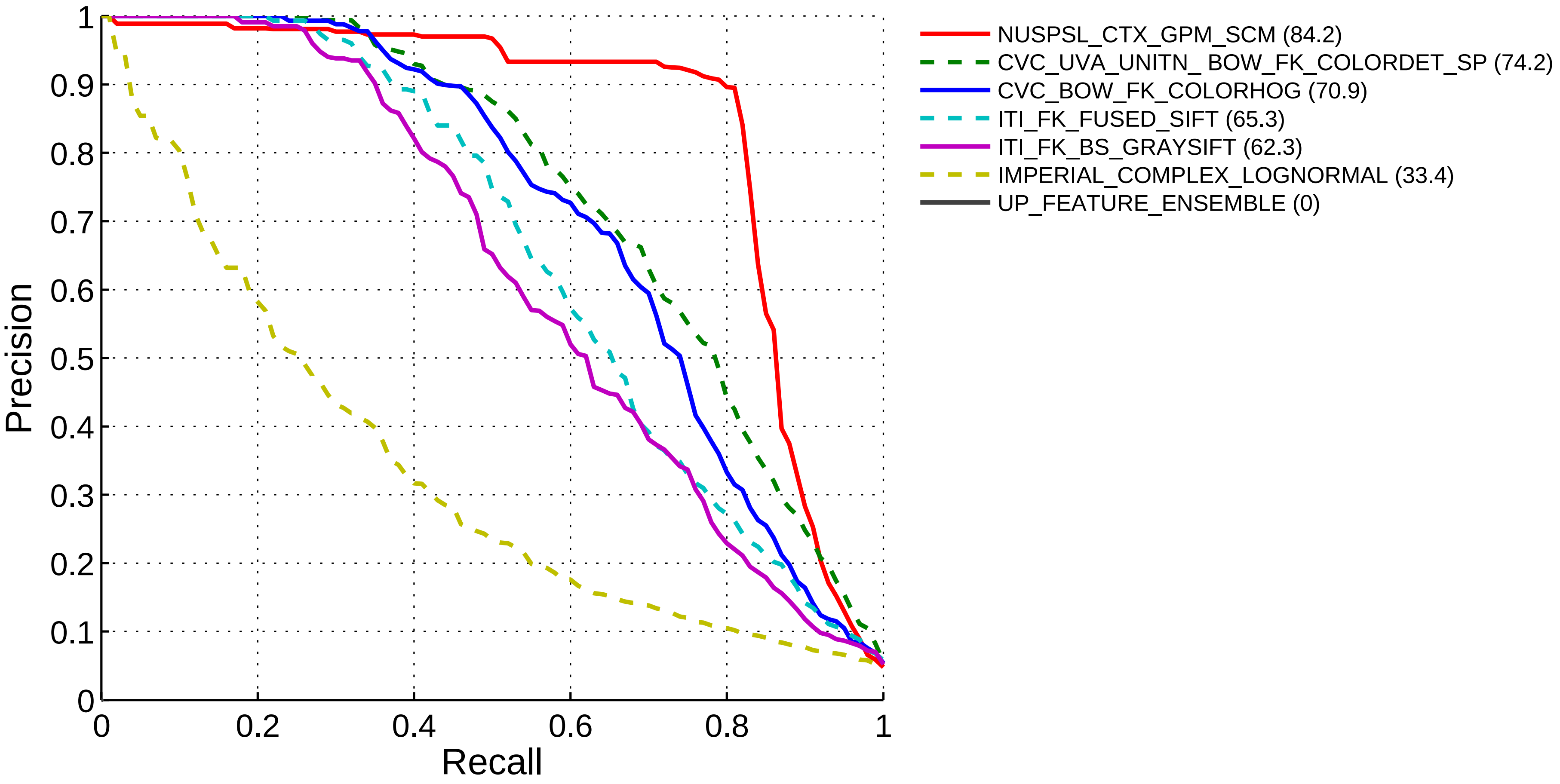
Precision/recall curves (aeroplane)



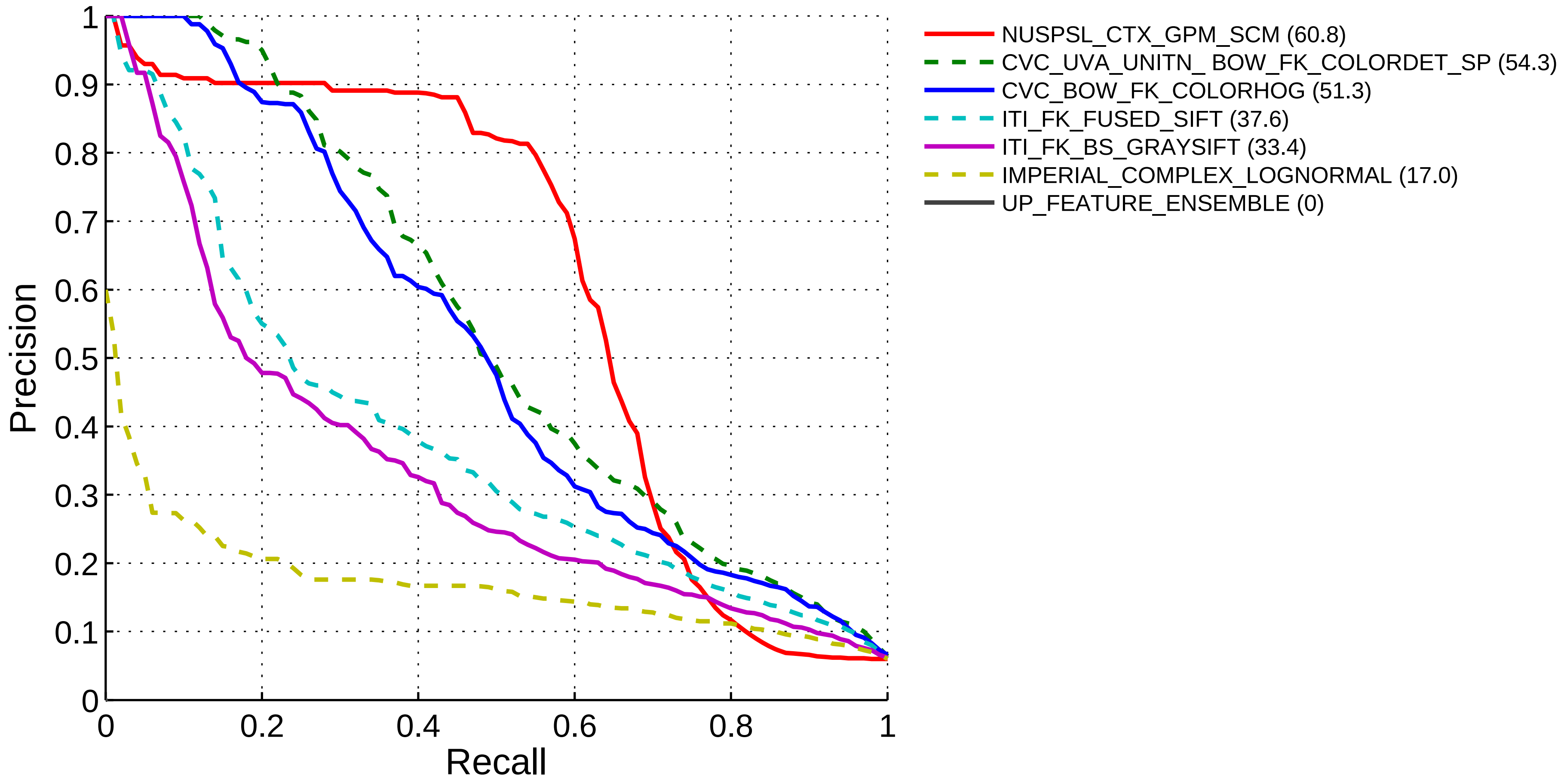
Precision/recall curves (person)



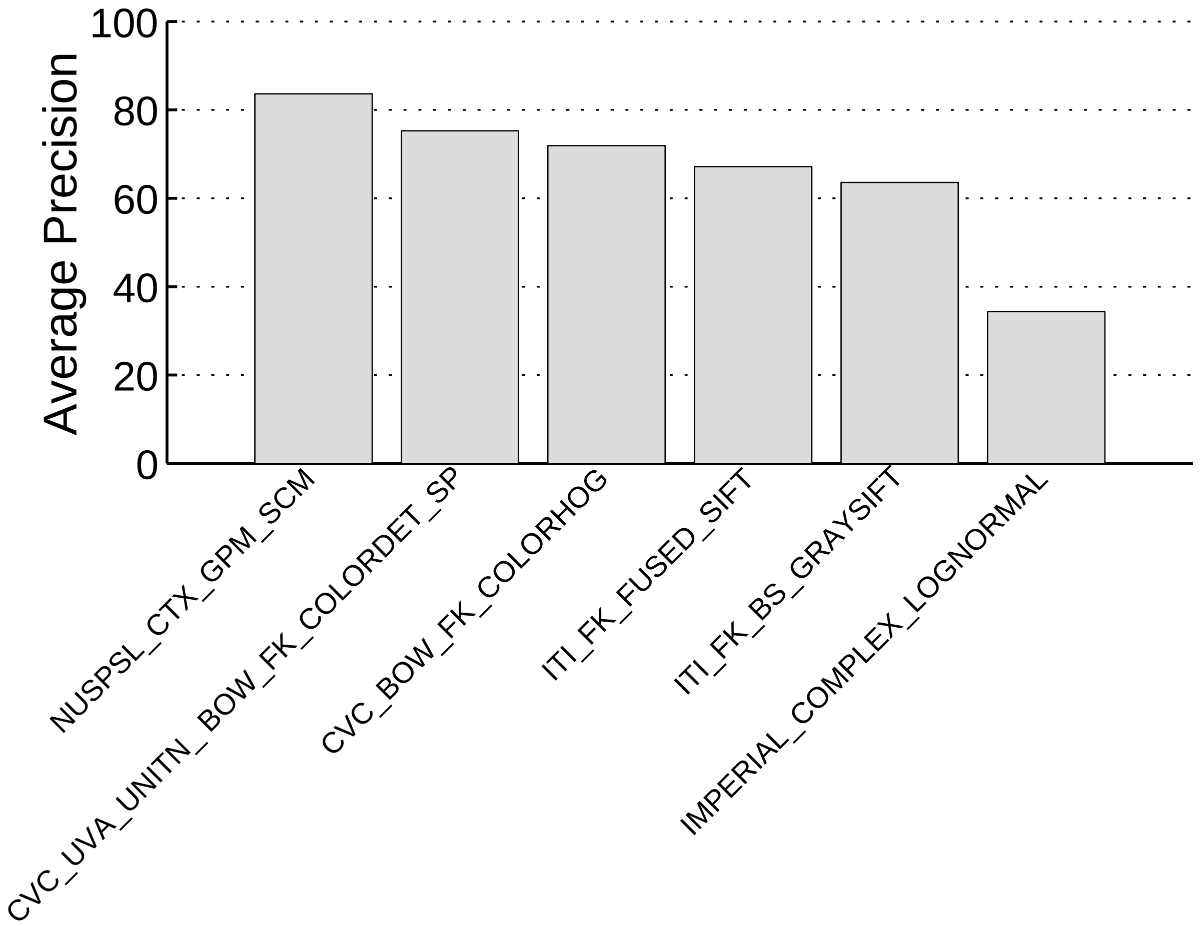
Precision/recall curves (bicycle)



Precision/recall curves (bottle)



Median average precision by method



Prizes

- Winner
  - **NUSPSL_CTX_GPM_SCM**

  Dong Jian, Chen Qiang, Song Zheng,
  Pan Yan, Xia Wei, Yan Shuicheng,
  Hua Yang, Huang Zhongyang, Shen Shengmei
  National University of Singapore
  Panasonic Singapore Laboratories
  Sun Yat-sen University

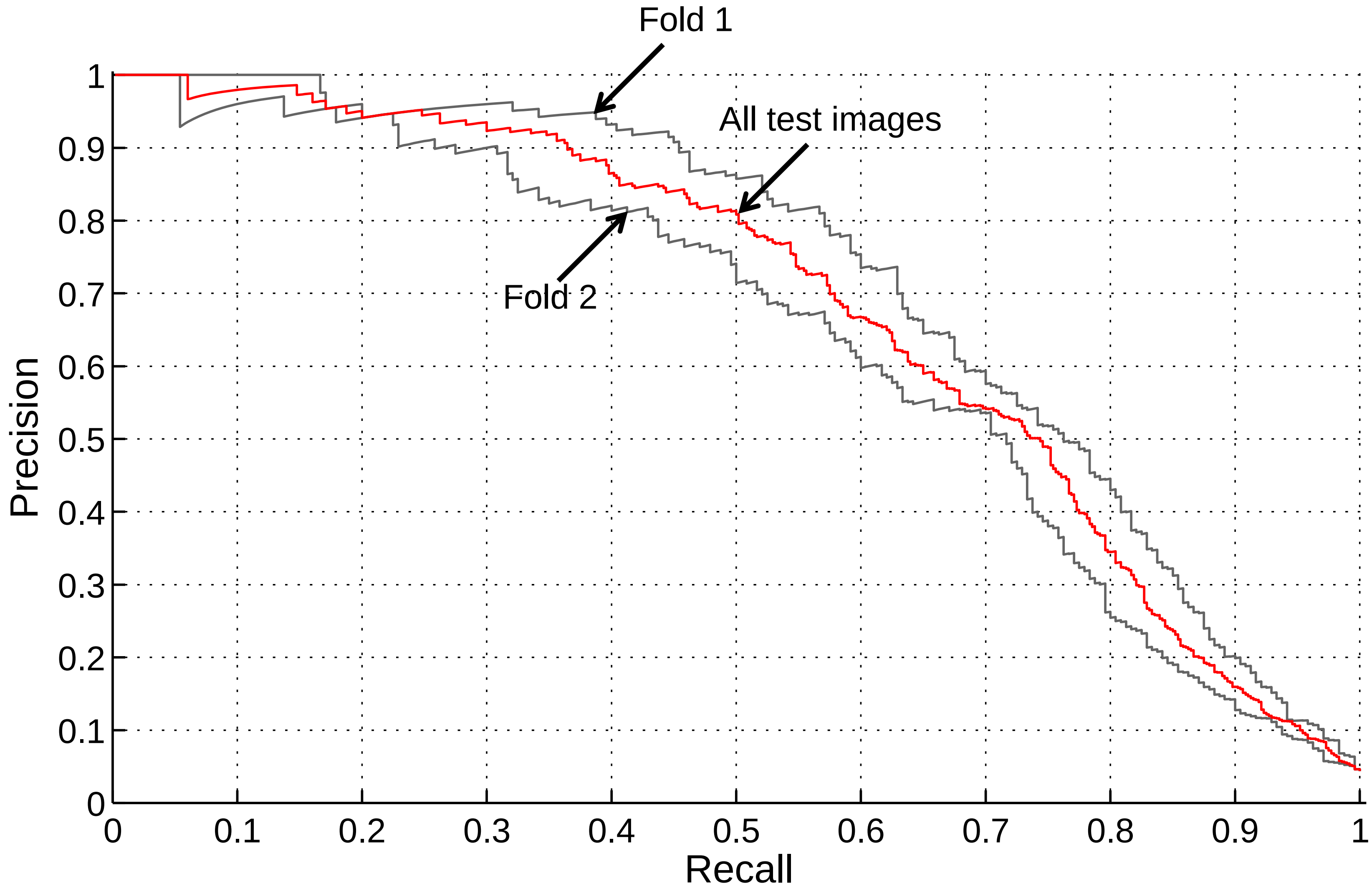
- Honourable mention
  - **CVC_UVA_UNITN_ BOW_FK_COLORDET_SP**

  Fahad Khan, Jan van Gemert, Camp Davesa, Jasper Uijlings , Albert Gordo, Sezer Karaoglu, Koen van de Sande, Pep Gonfaus, Rao Muhammad Anwer, Joost van de Weijer, Cees Snoek, Ramon Baldrich, Nicu Sebe, Theo Gevers
  Computer Vision Barcelona
  University of Amsterdam
  University of Trento

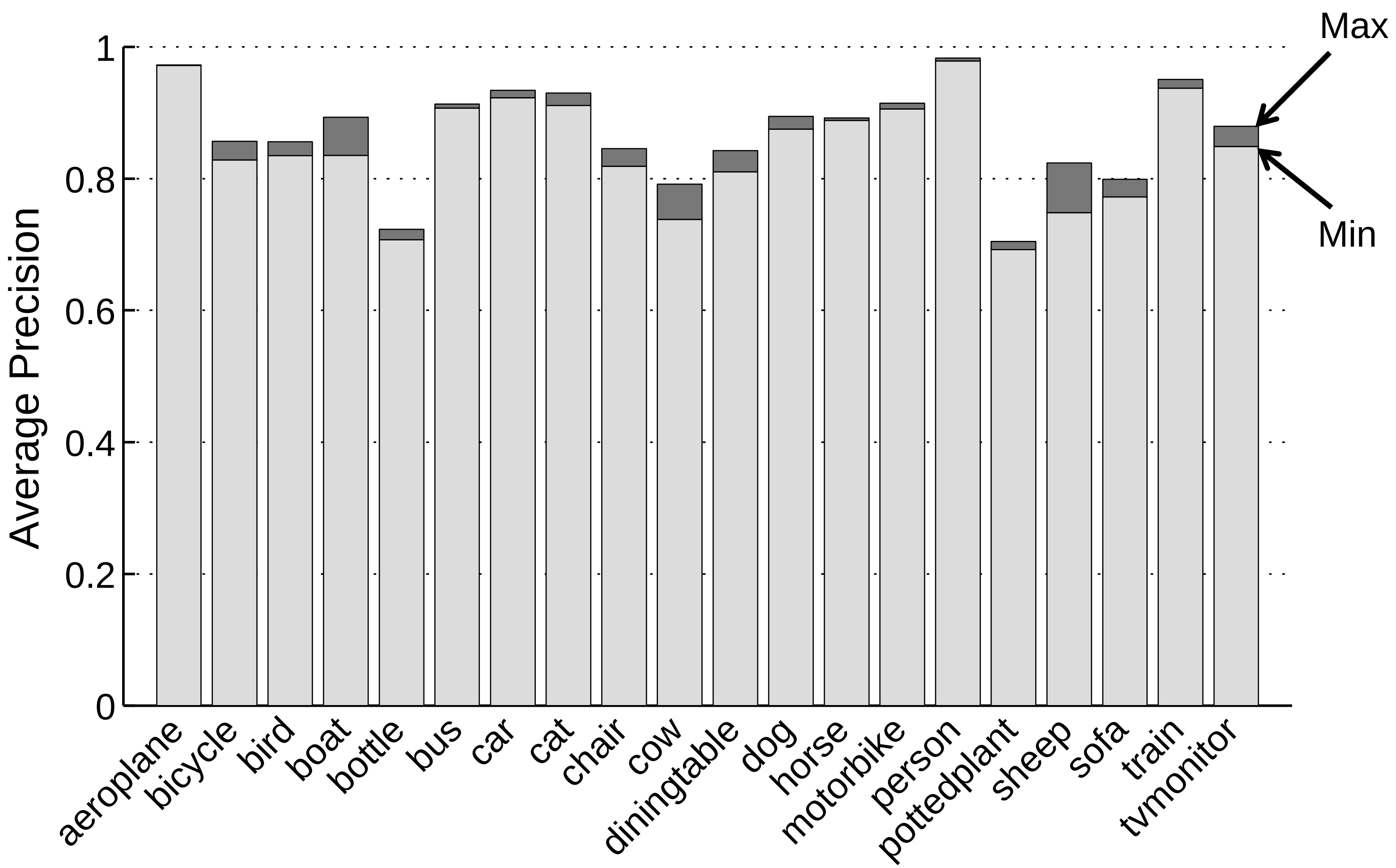
Super-classifiers

- Split test data into two sets, **A** and **B**
- Use the score of each method as a feature for each image of set **A** (feature vector length is equal to number of methods)
- Learn a linear L2 binary SVM classifier (with C=1) to predict classification of each image for each class (learn 20 classifiers)
- Compute PR curves and AP for each classifier
- Repeat by training on **B** and testing on **A**

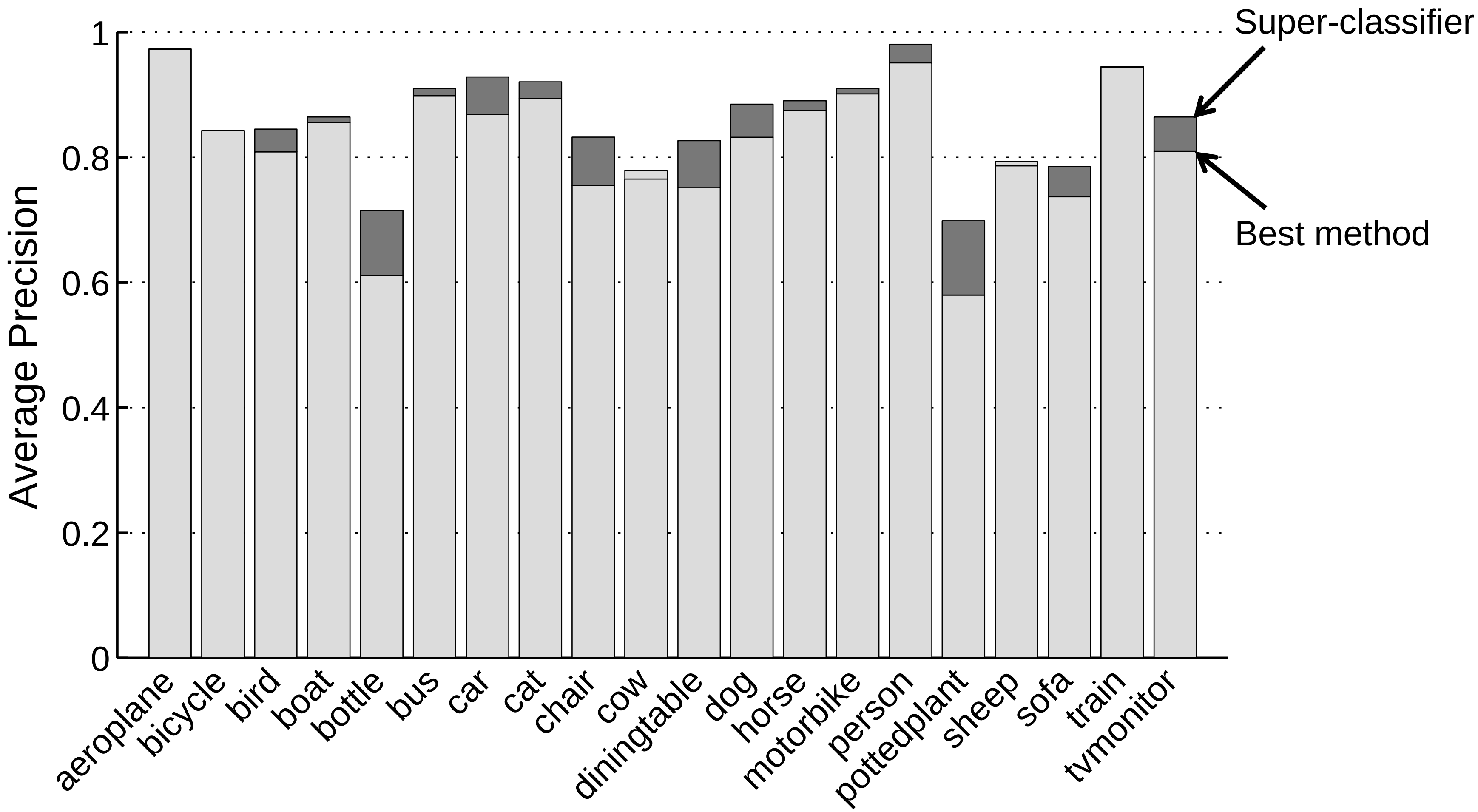
Sensitivity to test split (boat)



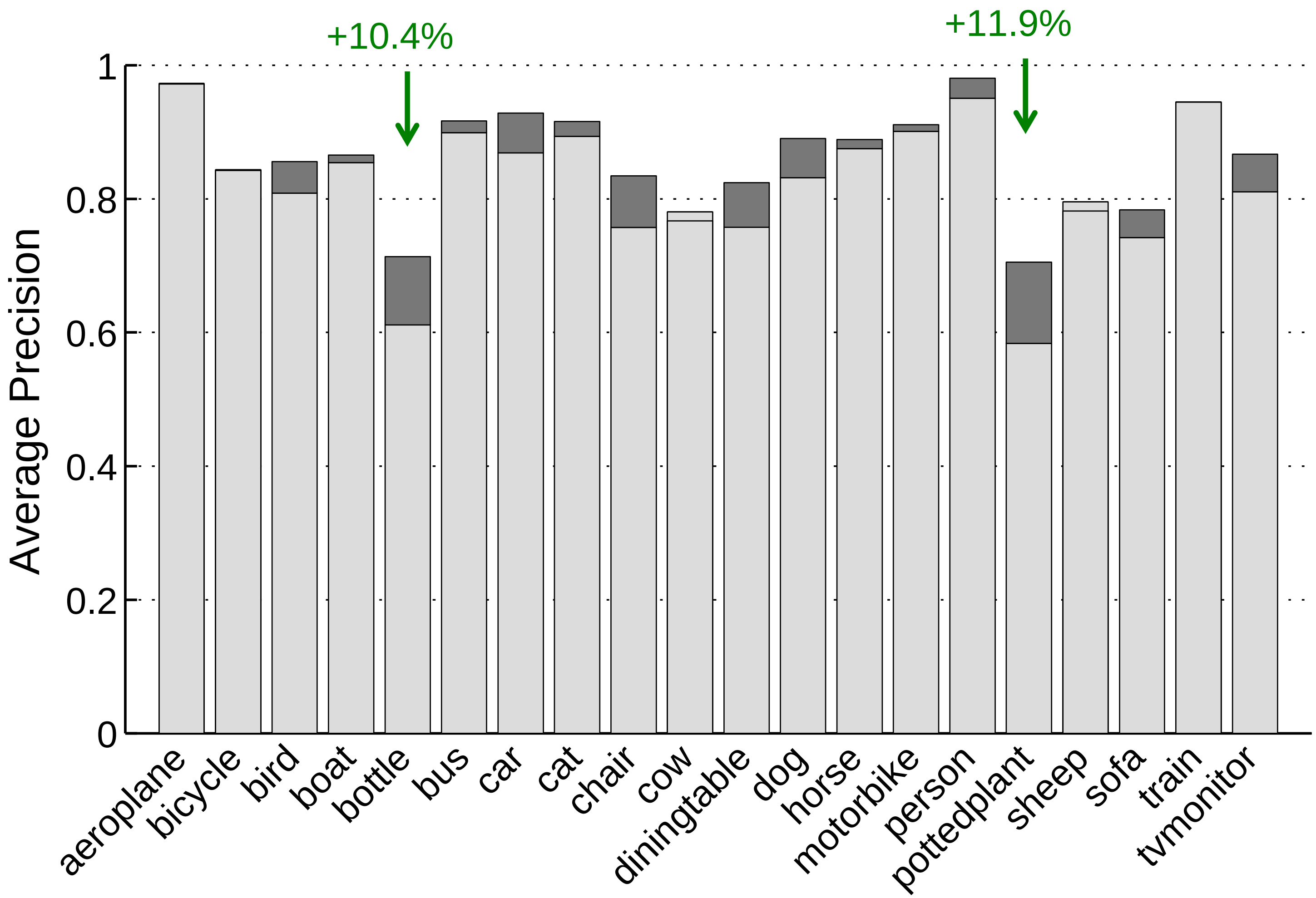
Super-classifiers (sensitivity to test split)



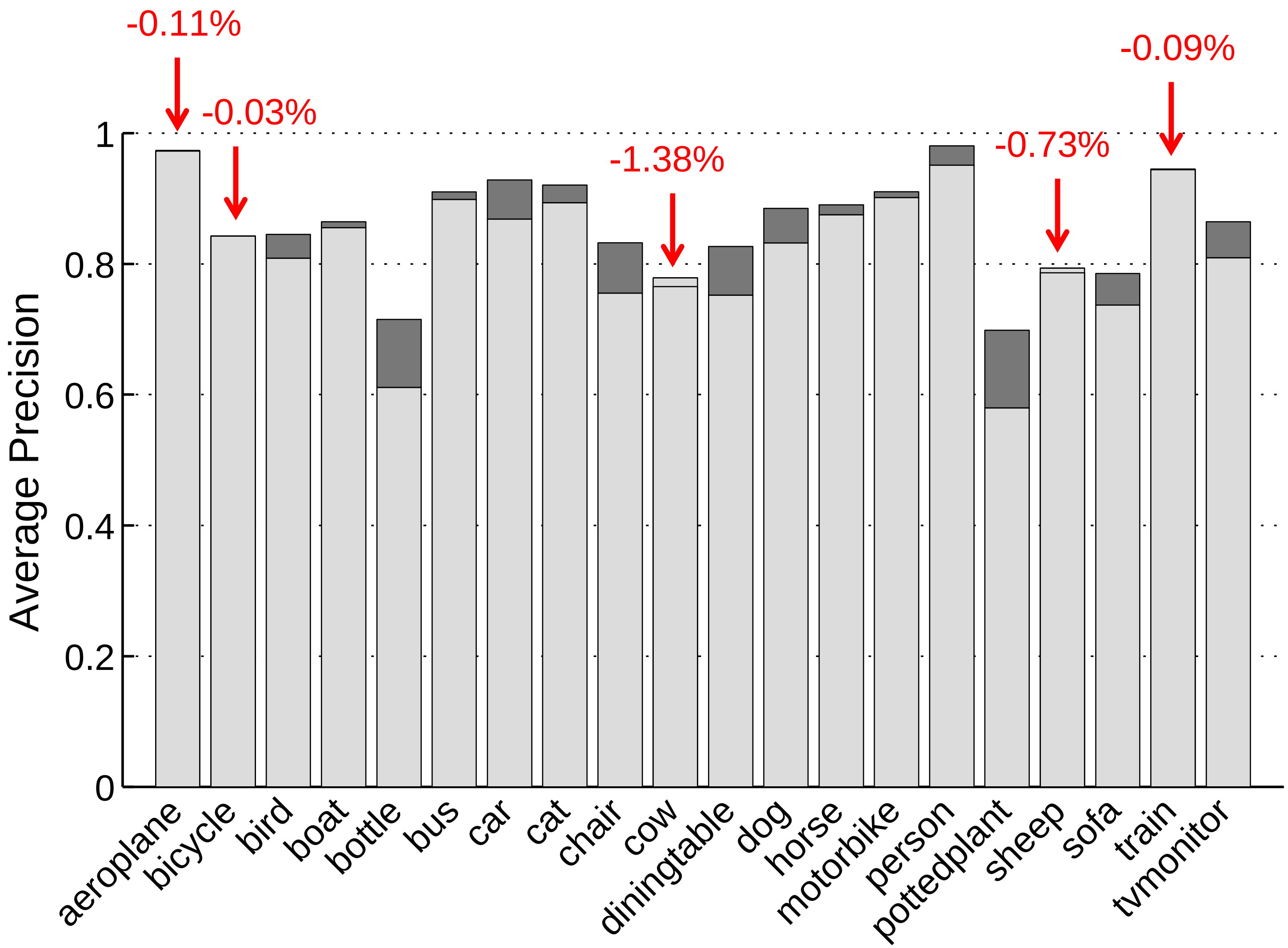
Super-classifiers



Super-classifiers

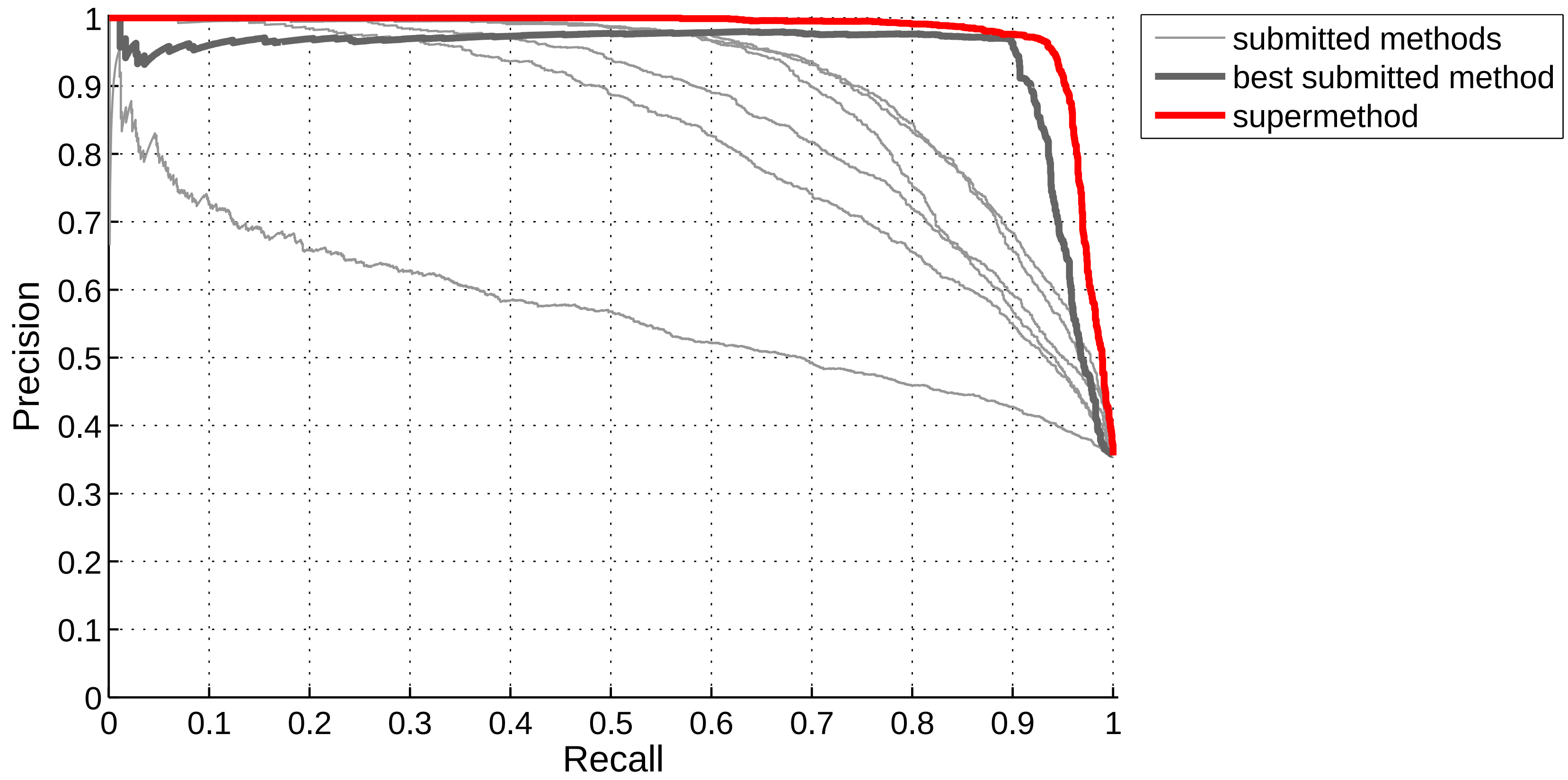


Super-classifiers



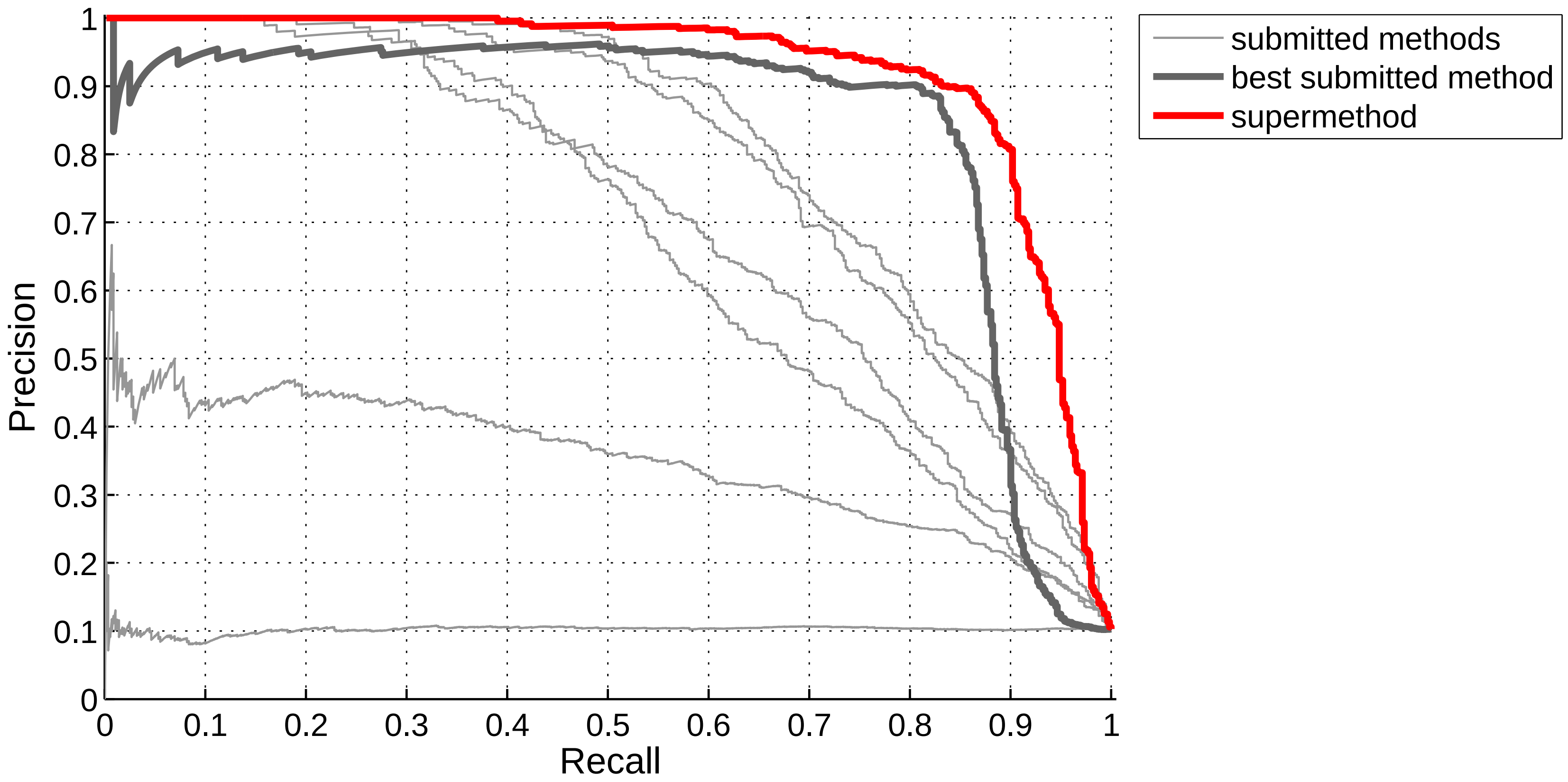
Super-classifiers (person)

AP +3.0



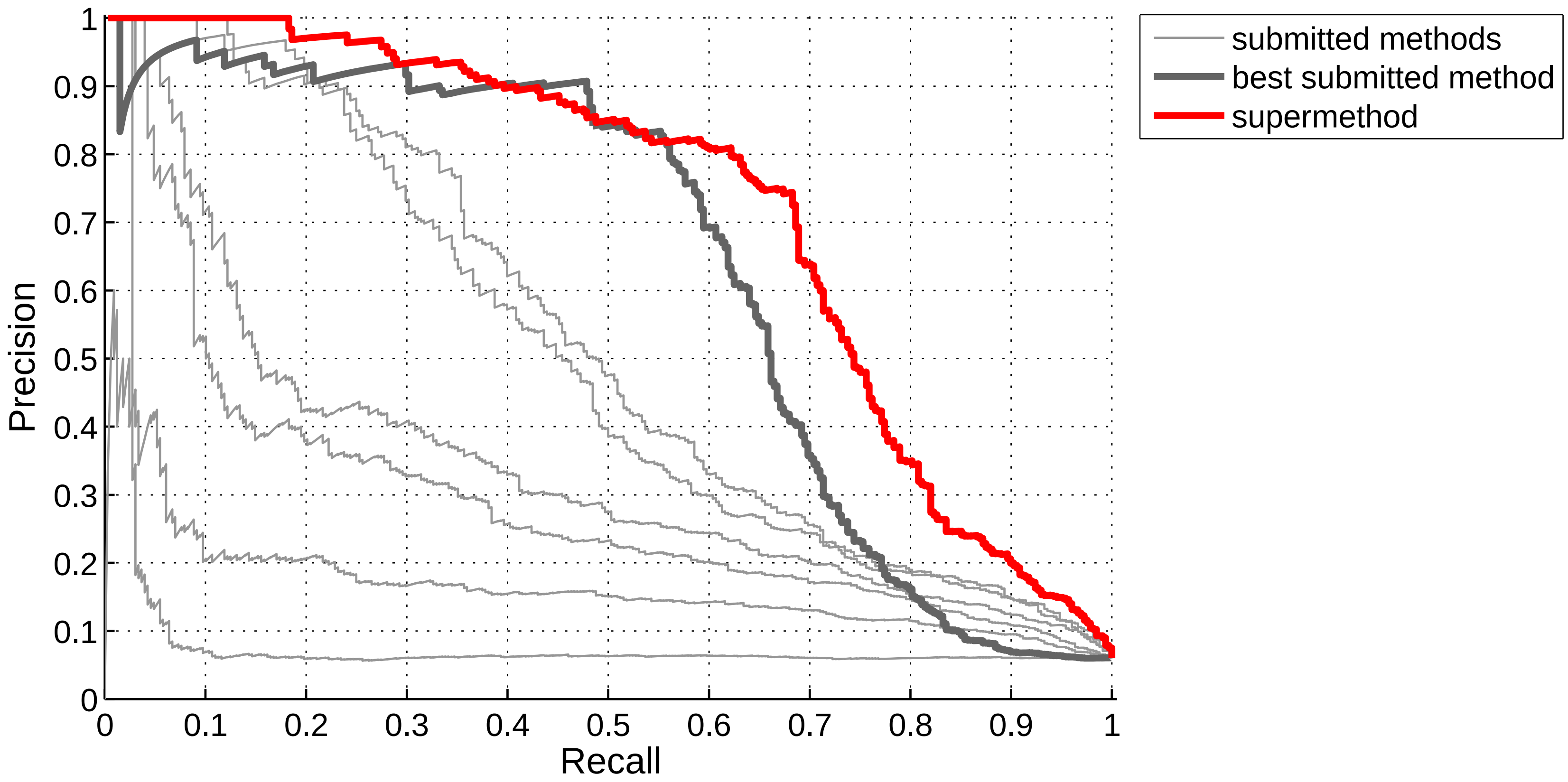
Super-classifiers (car)

AP +6.0



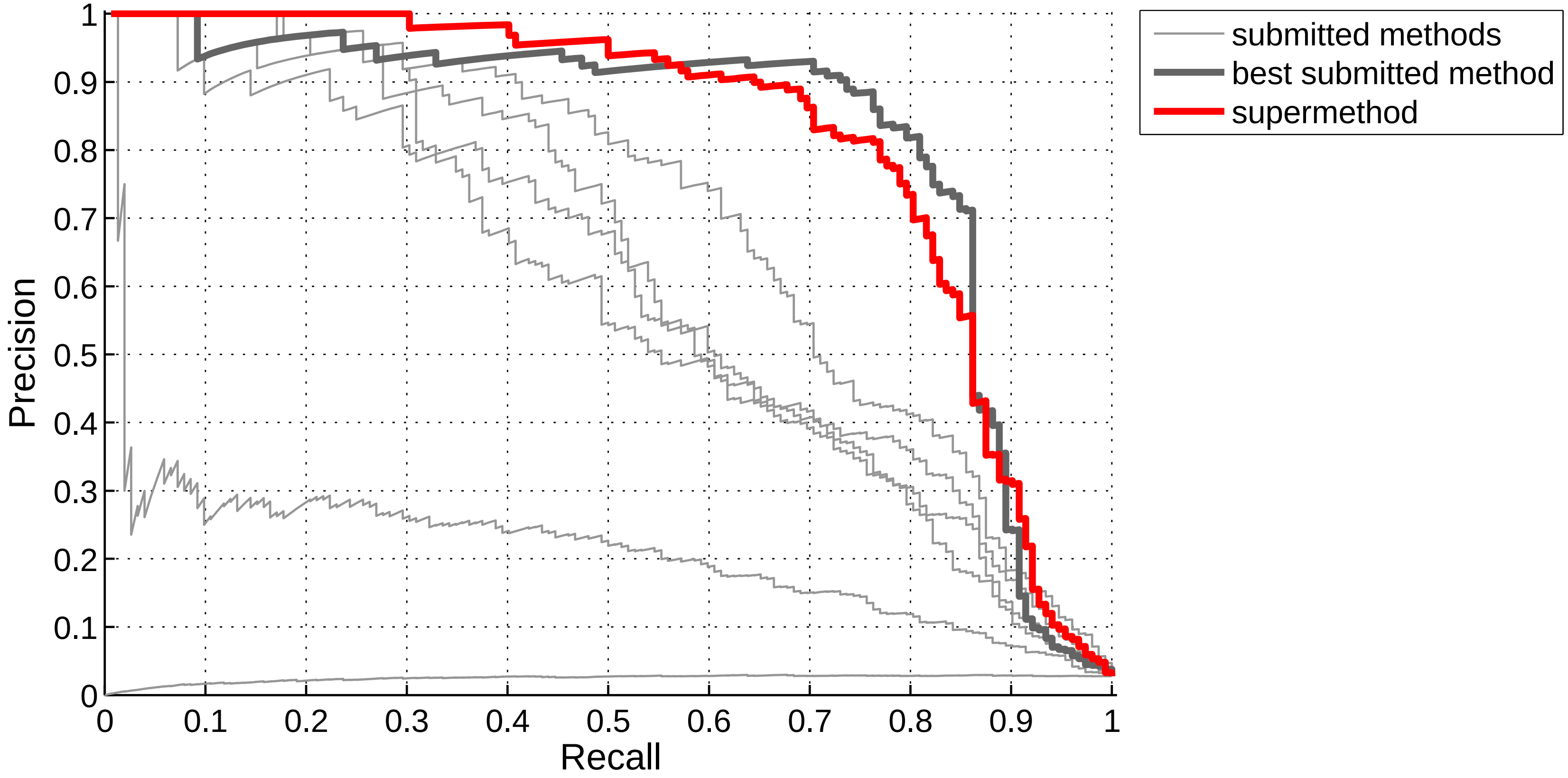
Super-classifiers (bottle)

AP +10.4



Super-classifiers (sheep)

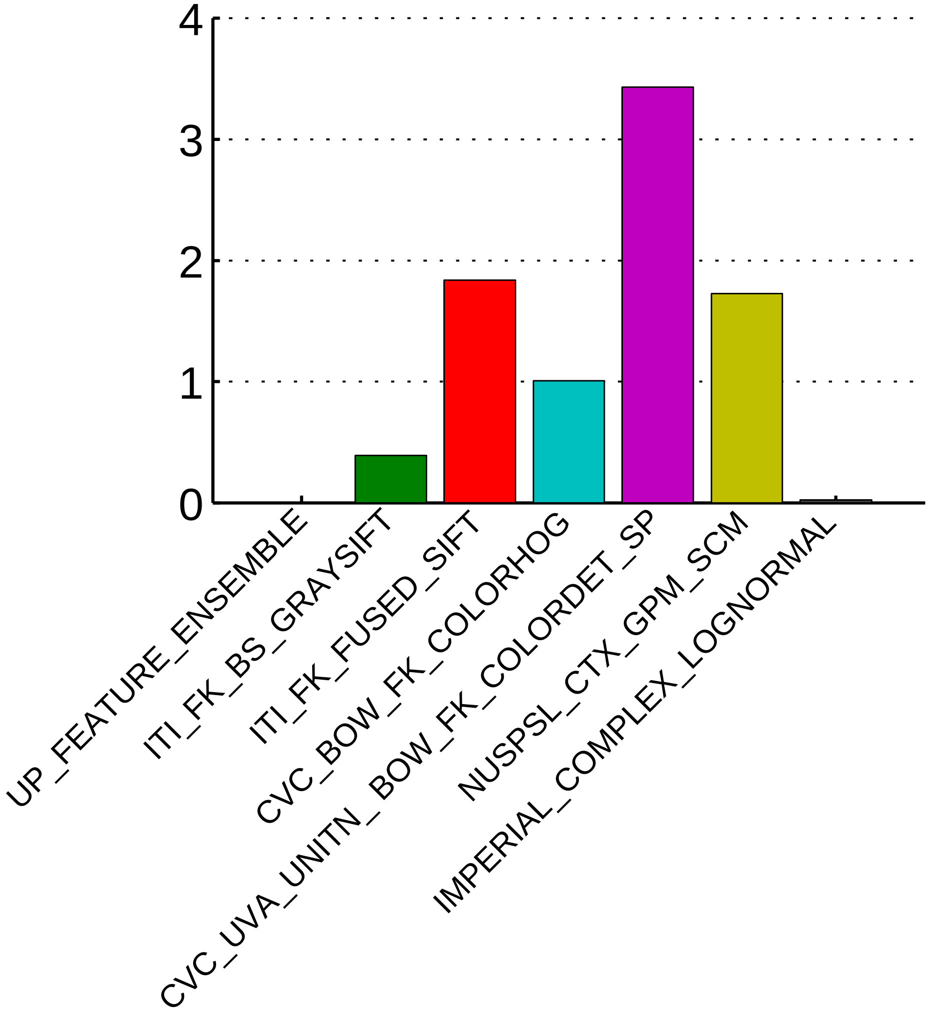
AP -0.73

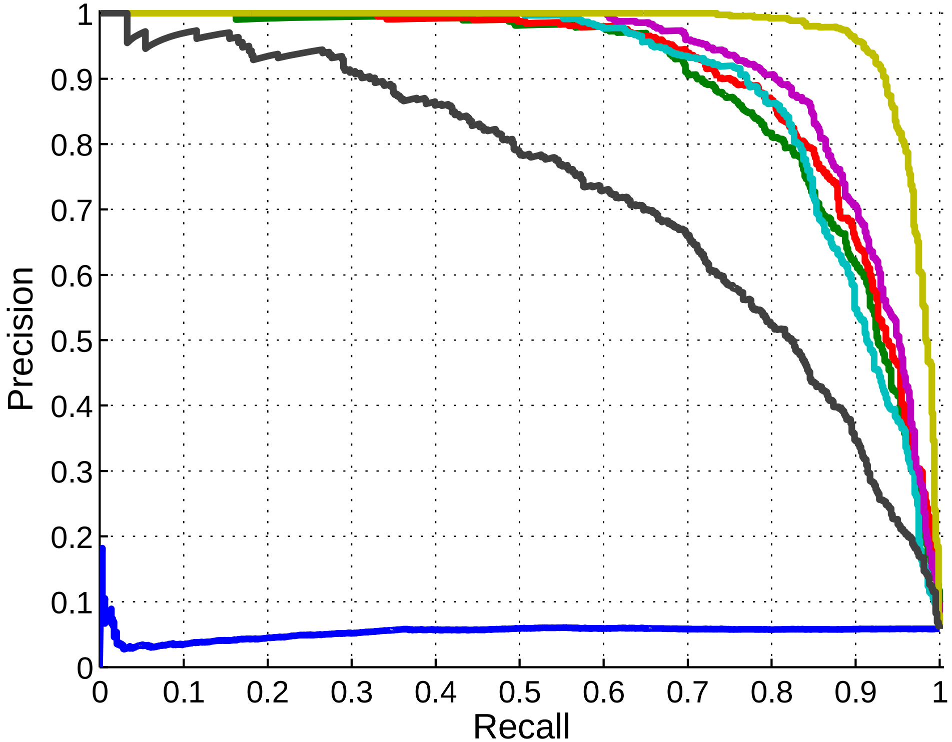


Super-classifiers (weights)

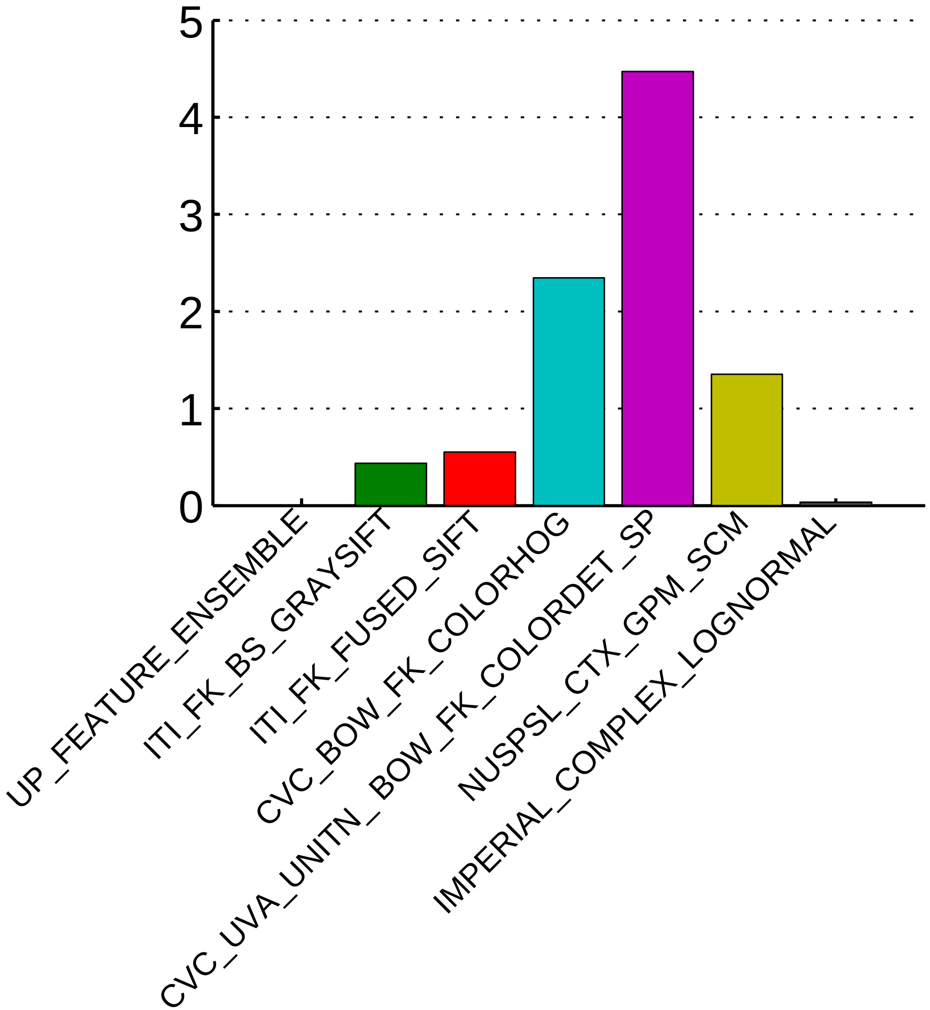
- Train super-classifier on *all* of test data.
- Inspect weight magnitudes for different features.

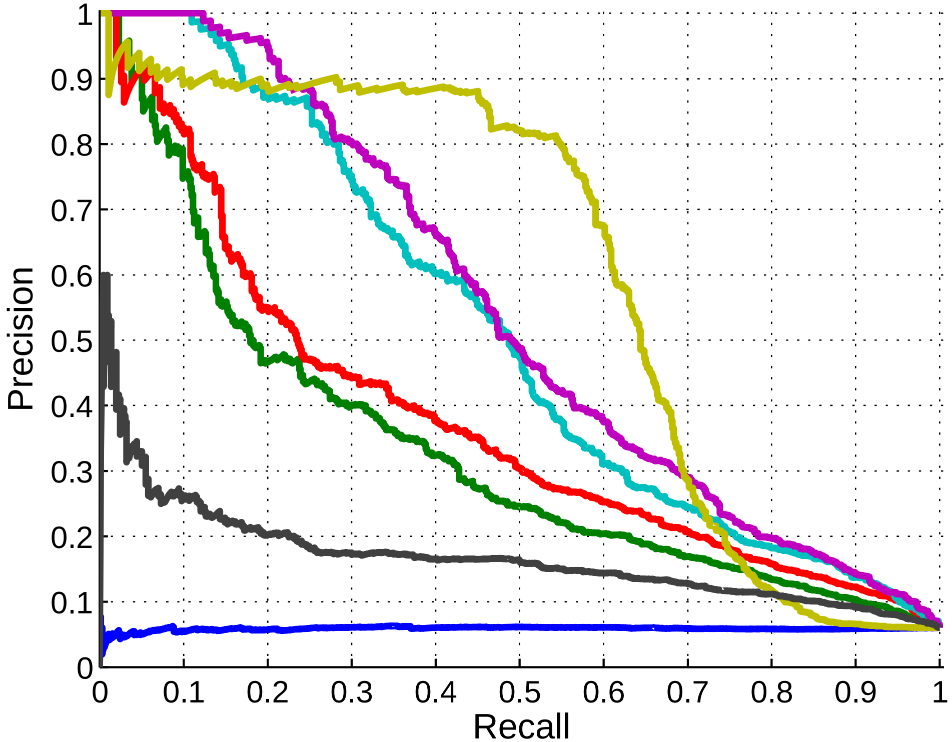
Super-classifiers (weights - aeroplane)



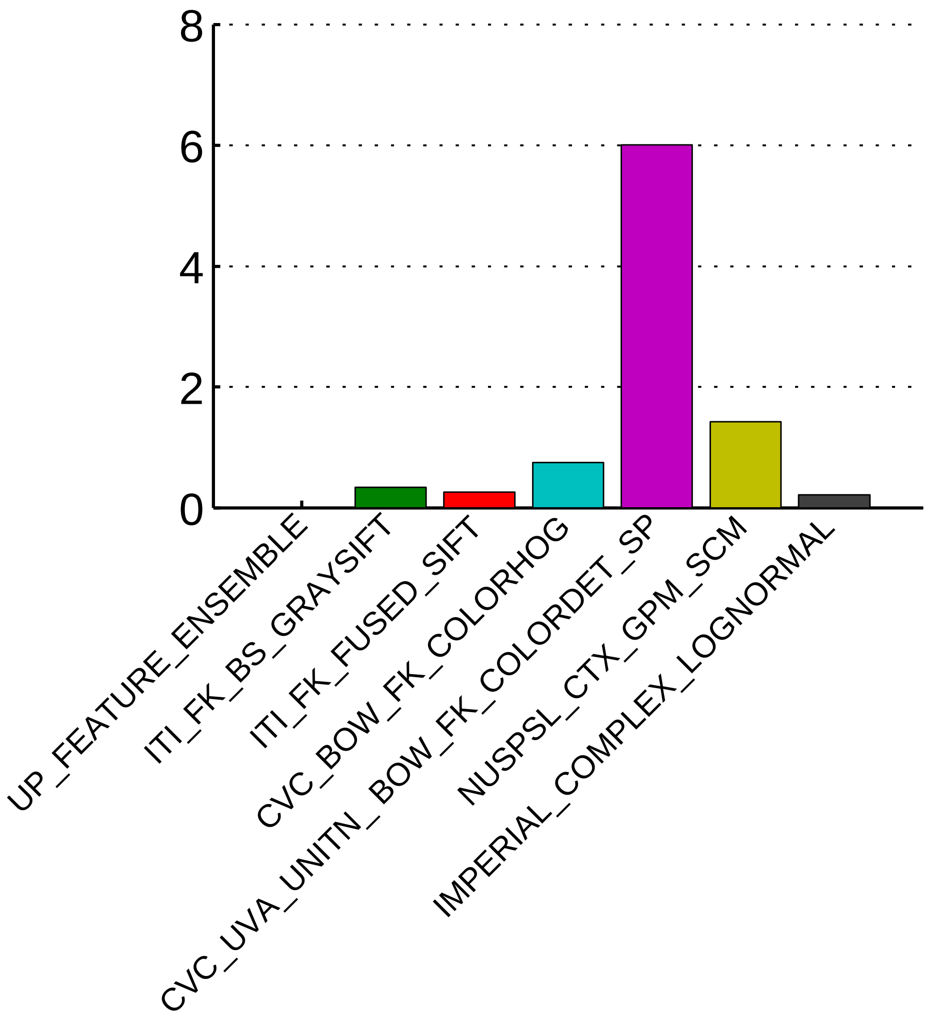


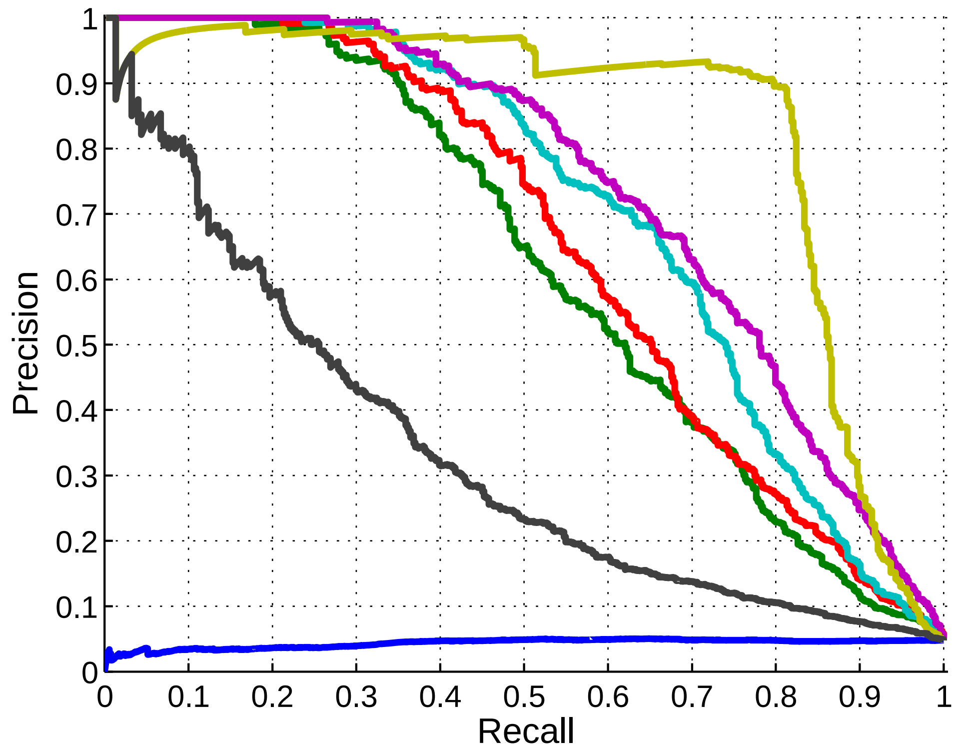
Super-classifiers (weights - bottle)



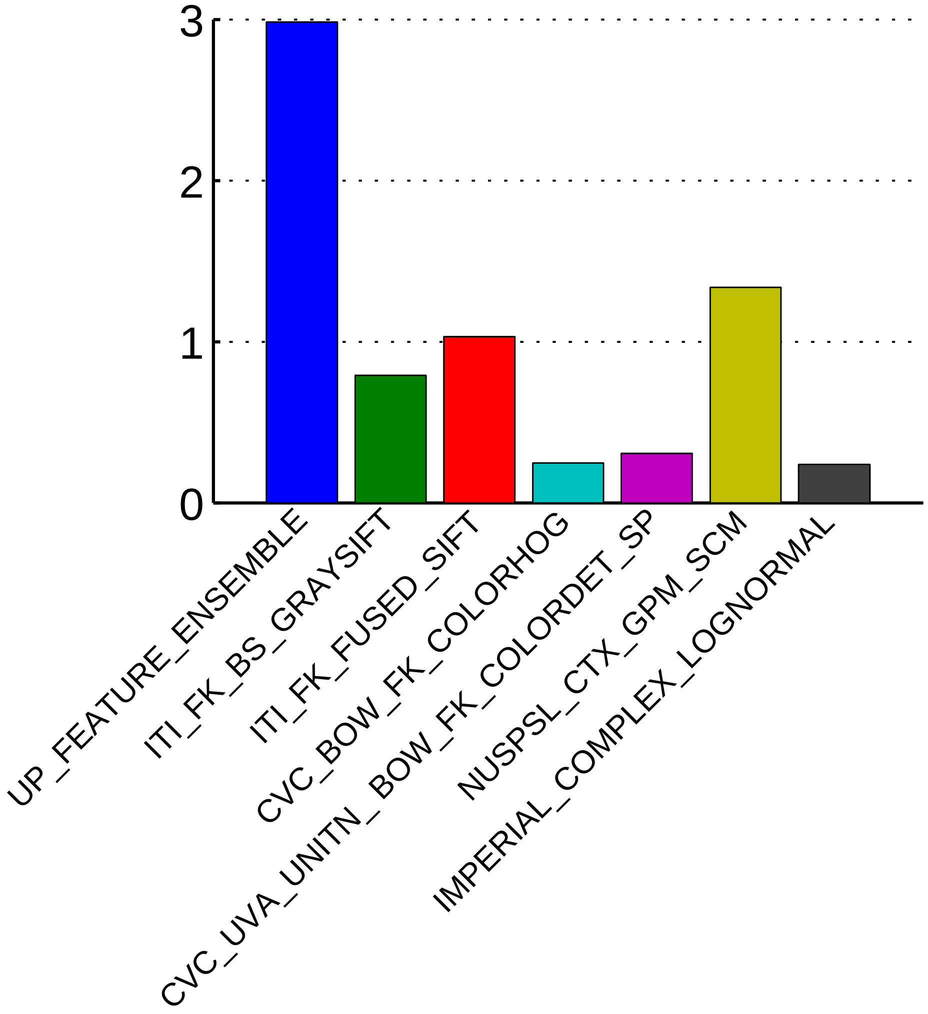


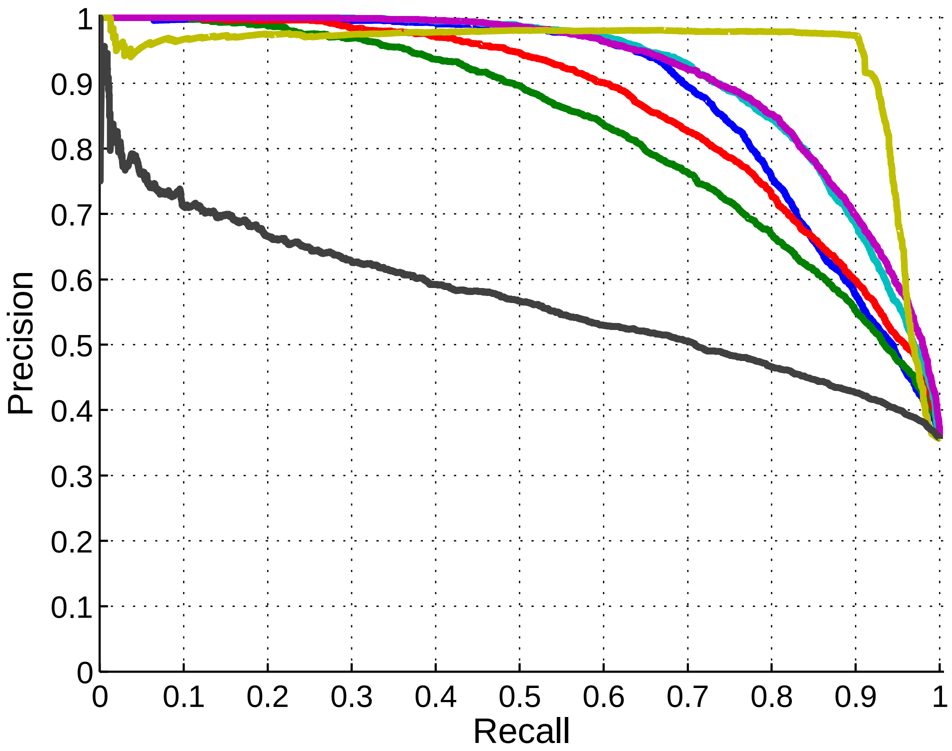
Super-classifiers (weights - bicycle)



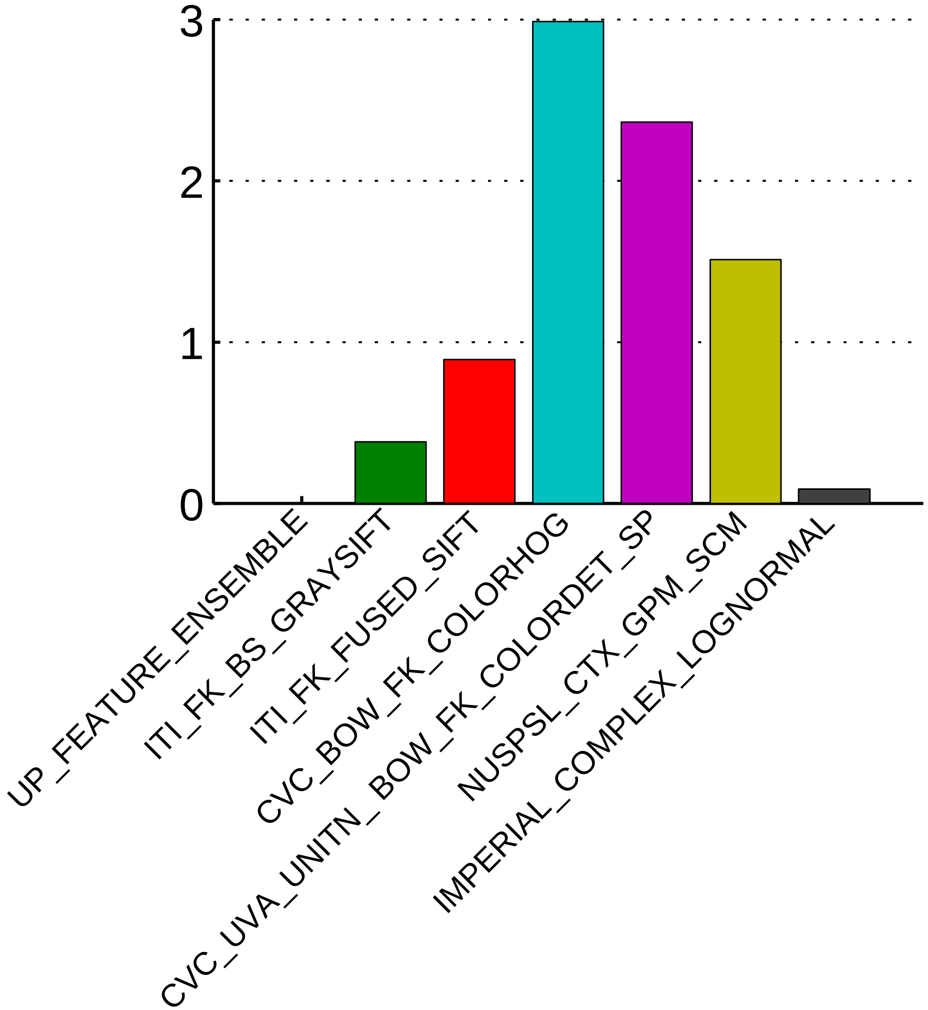


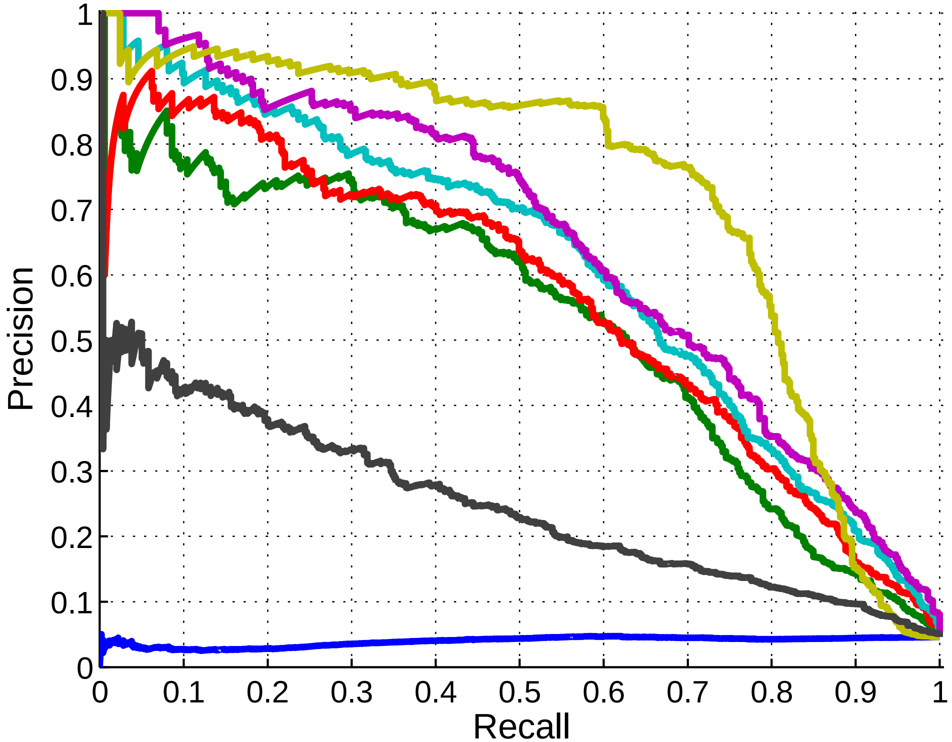
Super-classifiers (weights - person)





Super-classifiers (weights - sofa)





Collaboration recommendations

- Consider all combinations of pairs of submissions
- Train a super-classifier using only predictions of each pair as features
- Choose
  - The pair of submissions that leads to maximum possible (pairwise) AP
  - The pair of submissions that leads to maximum possible (pairwise) relative *increase* in AP

Collaboration recommendations

Class	Top AP	Collaborator 1 Name	AP	Collaborator 2 Name	AP	C^1+C^2 AP	All combined AP
aeroplane	**97.34**	UP_FEATURE_ENSEMBLE	6.00	NUSPSL_CTX_GPM_SCM	**97.34**	97.34	97.23
bicycle	**84.28**	ITI_FK_FUSED_SIFT	65.35	NUSPSL_CTX_GPM_SCM	**84.28**	85.30	84.25
bird	**80.89**	CVC_BOW_FK_COLORHOG	69.78	NUSPSL_CTX_GPM_SCM	**80.89**	85.09	84.54
boat	**85.52**	CVC_UVA_UNITN_ BO...	77.53	NUSPSL_CTX_GPM_SCM	**85.52**	86.38	86.44
bottle	**61.12**	CVC_UVA_UNITN_ BO...	54.37	NUSPSL_CTX_GPM_SCM	**61.12**	69.94	71.51
bus	**89.86**	CVC_BOW_FK_COLORHOG	84.79	NUSPSL_CTX_GPM_SCM	**89.86**	91.12	91.00
car	**86.87**	CVC_UVA_UNITN_ BO...	81.90	NUSPSL_CTX_GPM_SCM	**86.87**	93.07	92.84
cat	**89.37**	CVC_UVA_UNITN_ BO...	76.54	NUSPSL_CTX_GPM_SCM	**89.37**	92.19	92.06
chair	**75.56**	CVC_UVA_UNITN_ BO...	65.21	NUSPSL_CTX_GPM_SCM	**75.56**	83.49	83.22
cow	**77.88**	UP_FEATURE_ENSEMBLE	3.93	NUSPSL_CTX_GPM_SCM	**77.88**	77.96	76.50
diningtable	**75.24**	CVC_UVA_UNITN_ BO...	68.59	NUSPSL_CTX_GPM_SCM	**75.24**	82.26	82.64
dog	**83.19**	CVC_UVA_UNITN_ BO...	68.94	NUSPSL_CTX_GPM_SCM	**83.19**	89.25	88.48
horse	**87.53**	ITI_FK_FUSED_SIFT	72.39	NUSPSL_CTX_GPM_SCM	**87.53**	88.84	89.02
motorbike	**90.14**	CVC_BOW_FK_COLORHOG	79.21	NUSPSL_CTX_GPM_SCM	**90.14**	90.78	91.03
person	**95.11**	CVC_UVA_UNITN_ BO...	91.62	NUSPSL_CTX_GPM_SCM	**95.11**	97.87	98.08
pottedplant	**57.99**	CVC_UVA_UNITN_ BO...	56.24	NUSPSL_CTX_GPM_SCM	**57.99**	70.11	69.86
sheep	**79.34**	NUSPSL_CTX_GPM_SCM	**79.34**	IMPERIAL_COMPLEX_...	23.86	79.48	78.61
sofa	**73.69**	ITI_FK_FUSED_SIFT	57.42	NUSPSL_CTX_GPM_SCM	**73.69**	78.56	78.55
train	**94.49**	UP_FEATURE_ENSEMBLE	5.10	NUSPSL_CTX_GPM_SCM	**94.49**	94.49	94.41
tvmonitor	**80.95**	CVC_UVA_UNITN_ BO...	77.37	NUSPSL_CTX_GPM_SCM	**80.95**	85.75	86.41

Collaborations to maximise combined AP: AP^{1+2}