Fraunhofer FIRST's Submission to ImageCLEF2009 Photo Annotation Task: Non-sparse Multiple Kernel Learning

Alexander Binder, Motoaki Kawanabe
Fraunhofer Institute FIRST, Kekuléstr. 7, 12489 Berlin, Germany
{alexander.binder, motoaki.kawanabe}@first.fraunhofer.de

23.08.2009

Abstract

In order to achieve good performance in image annotation tasks, it is necessary to combine information from various image features. In our submission, we applied the non-sparse multiple kernel learning for feature combination proposed by Kloft et al.(2009) to the ImageCLEF2009 photo annotation data. Since some of the concepts of the ImageCLEF task are rather abstract, we conjectured that color histograms are informative for some categories such as sky and snow. Therefore we tried pyramid histograms of pixel colors. Since the images are not aligned, we sorted histograms at different places, when computing similarity of two images. Short description of our methods will be presented and obtained results will be discussed in this manuscript.

Categories and Subject Descriptors

I.4 [**Image Processing and Computer Vision**]: I.4.8 Scene Analysis; I.4.9 Image Representation; I.5 [**Pattern Recognition**]: I.5.2 Design Methodology; I.5.4 Applications

General Terms

Measurement, Performance, Experimentation

Keywords

ImageCLEF2009, Photo Annotation, Support Vector Machine, Multiple Kernel Learning, Spatial Pyramid Representation, Bag of Words

1 Introduction

Recent research results show that combining information from various image features is inevitable to achieve good performance in image annotation tasks. With the support vector machine (SVM), this is implemented by mixing kernels (similarities between images) constructed from different image descriptors with appropriate weights. For instance, the average kernel with uniform weights or the optimal kernel trained by multiple kernel learning (MKL) have been used so far. Recently, Kloft et al.(2009) proposed the non-sparse MKL with L^p-regularizer, which bridges the abovementioned two extremes, i.e. the average kernel SVM and the standard MKL. The non-sparse MKL is successfully applied to object classification tasks; it could outperform the two baseline methods by optimizing the tuning parameter $p \geq 1$ through cross validation. In our submission, we applied the novel technique for feature combination to ImageCLEF2009 photo annotation data based on bag-of-words [2] over SIFT features which is similar to a subset of the Pascal VOC 2008 winners [10, 3] submission. Furthermore, we tested sorted pyramid kernel with color histograms which will be explained in Section 2.1.1.

2 Experiment Description

2.1 Features

We computed SIFT features [6, 7] on a dense grid of step size six over the color channels red, green, blue, normalized red, normalized green, normalized blue, opponent color 1, opponent color2, normalized opponent color 1,normalized opponent color 2, grey. We used radii 4,8,12,16 for the SIFT feature supports.

For the bag of words models we combined the color channels into the following five sets: grey, red-green-blue, normalized red-green-blue, opponent color 1 - opponent color2. The bag of words prototypes have been obtained using kmeans clustering with 800000 SIFT features randomly sampled from 3000 files. The number of prototypes was fixed to 4000 for each of the five channel combinations. Finally we computed the bag of words for image tilings 1x1, 2x2 and 3x1 in the sense of spatial pyrmaid levels [5, 1]. This results in 12 features of which we applied in each learning setting at most eleven due to 32 bit memory restrictions. Furthermore we employed pyramid histograms over color intensities features based on the color channels grey, opponent color 1,opponent color 2, green and blue. We chose the number of bins per pyramid cell to be ten.

2.1.1 Modifications of pyramid histograms over color intensities

In order to improve the information content of the spatial pyramids for pyramid histograms over color intensities (PHoCol), we sorted the cells within a pyramid level according to the slant of the histogram. This reflects the notion that an image patch having high intensities in a certain colour channel results in an intensity histogram slanted towards higher intensities. The slant was determined via the entropy of the accumulated histogram of a histogram:

$$\widetilde{a}(h)_i = \sum_{k \leq i} h_k, a(h)_i = \frac{\widetilde{a}(h)_i}{\sum_k \widetilde{a}(h)_k}, sl(h) = \sum_i -a(h)_i \ln(a(h)_i)$$

In addition to that we included the difference between sort costs in the kernel. Algebraically a sort cost is just a mapping of the sort permutation of a pyramid level into the real numbers. In our specific case we resorted to

$$sc(\pi) = C \sum_k \max(v(\pi(k)) - v(k), 0)$$

where $v(i)$ is the height of the cell with index i with C being chosen so that the sort cost is upperbounded by one and lies at a similar scale compared to the χ^2-distance. The intuition behind such a sort cost is to punish large changes of the vertical position of an image patch associated to a pyramid cell caused by sorting compared to the unsorted pyramid. This sort cost is obviously invariant to changes in horizontal position of pyramid cells.

The sort cost modified kernel is a product of the χ^2-Kernel and a gaussian kernel over the difference of the sortcosts between two features

$$k(x, y) = \exp(-\sigma(d_{\chi^2}(h_x, h_y) + (sc_x - sc_y)^2))$$

. While it would be straightforward to replace this fixed weight kernel modification by an linearized MKL-approach, we omitted it due to 32 bit memory restrictions. We also experimented with an alternative formulation

$$k(x, y) = \exp(-\sigma(d_{\chi^2}(h_x, h_y)(1 + (sc_x - sc_y)^2))$$

for which the kernel property is not proven which however turned out to be nonnegative definite.

In the end we computed these features using tilings 4x4, no sorting, 3x1, sorted with sort cost, 16x16, sorted without sort cost over color channels noted above. While we do not regard such features as good standalone features, we realized their speed and potential value for contributing to some classifiers of the 53 classes in a proper MKL-procedure. We had low structured concepts with high variance in mind like sky, water, portrait, daytime.

2.2 Kernels

We used the χ^2-Kernel with the exception of pyramid histograms over color intensities. All kernels have been normalized.

2.3 Experiments

We used support vector machines with average kernels, sparse L^1-MKL, and the recently developed general non-sparse L^p-MKL [4]. The regularization parameter was fixed to one.

We submitted one single kernel result, namely phows over red-green-blue with 1x1 tiling as an early fallback, furthermore we submitted for each binary classifier the single best support vector learning algorithm measured by AP/AUC and FRR rate score over ten-fold crossvalidation (L^p-joint) using which generates three additional submissions

A) eleven bag of words kernels, namely red-green-blue, normalized red-green-blue and grey, all tilings, and opponent color 1 - opponent color2, 1x1 and 3x1 tiling.

B) nine bag of words kernels, namely red-green-blue, normalized red-green-blue and grey, all tilings,

C) 10 kernels, namely bag of words red-green-blue, all tilings, grey 1x1 tiling, PHoCol over green and blue channels 4x4,3x1,16x16 tilings (various sort strategies)

D) eleven kernels, namely bag of words red-green-blue and grey, all tilings, opponent color 1 - opponent color2 and normalized opponent color 1 - opponent color2, both with 1x1 tiling, PHoCol over grey channels 4x4,3x1,16x16 tilings (various sort strategies)

E) 10 kernels, namely bag of words red-green-blue, all tilings, grey 1x1 tiling, PHoCol over opponent color 1 and 2 channels 4x4,3x1,16x16 tilings (various sort strategies)

We used for the standard SVM and the L^p-MKL code from the shogun toolbox [9, 8].

The experiments were conducted with C++-code on an Opteron-Cluster running on a 32 bit Linux which implied a hard limit for memory usage of 3GB. On average we used 20 CPUs.

2.4 Conclusion

On average combinations of only Bag of words features perform best. There exists classes for which adding the low level color descriptors improves performance. Since each binary classifier can be selected separately based on crossvalidation error adding the color features improves the final submission. By comparing the results from A and B versus C, D, and E one can clearly observe overfitting. In our point of view it comes from noisy feature extraction in connection with small sample sizes for many of the smaller concept classes. We doubt that the applied features permit a Bayes Error close to zero. In that sense there is still a need for better feature extraction despite the success story of Bag of Words representations.

Below are AP and AUC scores over 5-fold crossvalidation, the result of the best method is marked bold. For multiclass subproblems we computed for each method the average score over the subproblem and marked the subproblem in total bold.

AP / methods	A+avg	A+L1	A+L1.5	A+L2	B+L1.0625	B+avg
average	43.904	42.751	43.823	43.839	42.474	43.353
Partylife	**26.32**	21.73	26.05	26.15	22.47	26.01
Familiy_Friends	**40.58**	38.35	39.06	39.13	38.63	39.98
Beach_Holidays	**19.49**	14.99	18.80	19.39	16.15	17.83
Building_Sights	35.92	34.12	**36.33**	36.27	34.62	35.47
Snow	**16.89**	13.19	13.61	14.57	10.50	15.25
Citylife	**48.26**	45.17	48.12	48.17	46.18	48.00
Landscape_Nature	**63.51**	62.50	63.43	63.46	62.48	62.84
Sports	5.06	4.10	5.01	4.94	3.32	4.74
Desert	17.33	12.96	17.15	17.16	14.62	12.90
Spring	**18.00**	19.03	18.02	18.11	17.21	17.69
Summer	**45.62**	44.74	44.97	45.02	43.32	44.31
Autumn	**15.33**	16.40	15.97	15.97	16.20	13.95
Winter	**26.67**	23.88	26.67	26.60	21.76	25.54
No_Visual_Season	**94.02**	93.51	93.89	93.92	93.65	93.85
Indoor	68.27	67.27	**68.57**	68.55	68.67	68.24
Outdoor	89.70	89.65	**90.15**	90.17	89.25	89.53
No_Visual_Place	39.55	38.91	**39.61**	39.42	39.09	39.51
Plants	62.96	62.88	**63.79**	63.73	61.57	62.77
Flowers	56.94	56.02	57.46	**57.56**	54.54	56.85
Trees	52.21	52.15	52.07	51.81	51.86	52.53
Sky	**72.74**	72.68	72.51	72.39	71.50	72.03
Clouds	73.84	73.44	73.88	**73.97**	71.74	73.16
Water	55.13	53.76	**55.34**	55.27	52.33	54.47
Lake	28.01	25.83	27.59	27.75	30.05	27.47
River	26.11	20.63	**26.37**	25.77	20.01	24.70
Sea	**45.44**	44.21	45.35	45.37	44.90	45.06
Mountains	33.77	29.76	33.76	**34.05**	30.71	32.17
Day	87.64	87.56	**87.74**	87.69	86.52	87.39
Night	51.99	50.24	52.26	**52.35**	49.91	51.33
No_Visual_Time	72.42	72.06	73.01	**73.04**	70.95	72.35
Sunny	**30.47**	27.89	29.08	29.18	28.01	28.28
Sunset_Sunrise	69.66	69.20	69.78	69.72	68.84	68.99
Canvas	13.26	12.79	13.16	13.24	12.34	12.39
Still_Life	30.88	31.11	31.07	31.09	30.40	**31.30**
Macro	23.90	23.22	23.45	23.30	**24.33**	23.98
Portrait	51.96	51.98	52.52	**52.52**	50.23	52.08
Overexposed	3.54	3.36	3.41	3.49	3.79	3.76
Underexposed	37.09	36.62	36.92	36.58	35.37	36.62
Neutral_Illumination	97.50	97.44	97.53	97.51	97.46	97.50
Motion_Blur	9.28	9.66	**9.84**	9.67	9.79	9.69
Out_of_focus	5.79	4.77	**5.72**	5.72	4.26	5.08
Partly_Blurred	68.30	68.12	**69.01**	68.84	67.53	68.14
No_Blur	88.35	87.78	**88.38**	88.44	87.77	88.10
Single_Person	57.16	57.69	57.34	**57.40**	56.57	56.51
Small_Group	27.10	26.80	26.53	**26.55**	24.89	24.75
Big_Group	28.33	27.21	28.80	**28.76**	26.35	29.73
No_Persons	89.49	89.10	89.61	**89.62**	89.05	89.17
Animals	**51.23**	48.75	51.12	51.18	48.12	49.79
Food	**41.00**	37.06	40.71	40.93	36.61	38.75
Vehicle	32.69	32.32	**32.93**	32.87	32.40	32.13
Aesthetic_Impression	25.11	25.39	25.15	25.05	25.00	25.36
Overall_Quality	35.47	34.78	34.23	34.35	36.23	36.06
Fancy	19.64	21.07	19.79	19.68	21.07	21.60

AP / methods	B+L1	B+L1.5	B+L2	C+L1	C+L1.0625	C+L2
average	42.47	43.33	43.36	41.00	41.23	40.66
Partylife	21.60	25.33	25.49	19.17	20.19	24.39
Familiy_Friends	38.49	39.59	39.67	35.21	35.58	35.55
Beach_Holidays	15.91	17.70	17.75	14.78	15.91	15.10
Building_Sights	33.91	35.31	35.35	33.06	32.73	33.96
Snow	13.21	14.89	15.07	8.27	8.88	11.58
Citylife	45.12	47.64	47.70	44.19	44.47	43.84
Landscape_Nature	62.34	63.11	63.00	62.23	61.65	60.48
Sports	4.12	4.69	4.71	2.95	2.96	**6.08**
Desert	12.52	12.84	12.90	9.71	8.30	5.93
Spring	19.04	17.69	17.70	18.09	16.77	14.75
Summer	44.02	43.85	43.89	38.12	38.22	35.55
Autumn	16.19	14.52	14.47	11.30	10.03	10.23
Winter	22.72	25.44	25.46	20.28	24.18	25.27
No_Visual_Season	93.39	93.73	93.74	93.02	93.06	92.70
Indoor	67.33	68.23	68.23	67.48	67.62	67.28
Outdoor	89.41	89.86	89.85	88.74	88.53	88.58
No_Visual_Place	38.50	39.17	39.21	39.29	39.63	37.78
Plants	61.99	63.66	63.61	59.77	59.66	59.50
Flowers	55.44	57.28	57.27	50.70	51.05	51.07
Trees	52.19	52.59	**52.63**	50.24	50.71	47.61
Sky	72.42	71.86	71.93	71.10	70.95	70.05
Clouds	72.51	72.98	73.09	72.32	71.57	70.10
Water	53.06	54.76	54.88	50.91	51.47	49.64
Lake	25.70	27.83	27.91	25.26	23.65	21.38
River	20.60	24.62	24.78	16.44	19.06	15.78
Sea	43.38	44.57	44.71	44.29	43.31	36.80
Mountains	29.76	32.14	32.24	30.09	31.19	29.71
Day	87.23	87.41	87.40	85.96	86.12	85.43
Night	49.72	51.52	51.68	51.14	51.17	49.71
No_Visual_Time	71.78	72.82	72.81	70.28	70.30	69.92
Sunny	27.01	27.76	27.79	26.40	26.87	27.07
Sunset_Sunrise	68.84	69.18	69.18	70.25	69.99	66.70
Canvas	12.06	12.30	12.34	11.57	13.38	**16.78**
Still_Life	31.18	31.18	31.14	28.67	29.92	29.73
Macro	23.08	23.89	23.83	21.18	22.30	23.66
Portrait	51.81	52.24	52.30	48.72	49.06	47.23
Overexposed	3.36	3.75	3.75	3.85	4.47	5.45
Underexposed	36.46	36.38	36.36	37.31	37.50	35.14
Neutral_Illumination	97.44	97.46	97.46	97.69	97.76	97.63
Motion_Blur	9.68	9.53	9.68	9.06	8.83	7.56
Out_of_focus	4.50	5.00	4.99	4.37	4.52	4.47
Partly_Blurred	67.86	68.70	68.64	66.20	66.02	65.50
No_Blur	87.64	88.08	88.06	87.30	87.23	86.82
Single_Person	57.36	56.73	56.69	55.60	55.22	55.38
Small_Group	25.73	24.54	24.49	22.75	23.30	24.92
Big_Group	27.29	29.49	29.69	21.07	21.65	22.14
No_Persons	88.77	89.24	89.27	88.60	88.23	88.76
Animals	47.99	49.83	49.84	42.64	43.19	42.12
Food	35.67	38.73	38.86	31.27	32.81	31.34
Vehicle	32.25	32.05	32.09	32.53	32.84	29.48
Aesthetic_Impression	25.31	25.22	25.23	24.84	24.79	25.07
Overall_Quality	34.80	36.04	35.99	34.85	34.79	35.60
Fancy	21.04	21.33	21.45	21.64	21.37	20.57

AP / methods	D+L1	D+L1.0625	D+L2	E+L1	E+L1.0625	E+L2
average	41.21	41.66	41.37	41.06	41.35	41.47
Partylife	17.50	19.70	22.45	18.34	19.98	21.11
Familiy_Friends	36.01	36.83	36.26	35.30	36.03	36.36
Beach_Holidays	12.21	14.56	15.28	13.02	13.76	12.45
Building_Sights	32.85	34.30	34.50	32.66	32.16	33.74
Snow	9.68	10.72	13.61	7.06	7.71	8.35
Citylife	43.98	45.20	45.44	43.87	43.99	46.98
Landscape_Nature	62.14	62.59	62.75	61.87	61.31	62.51
Sports	3.53	3.01	3.36	3.00	2.80	2.66
Desert	13.64	7.36	8.71	17.30	**18.14**	14.98
Spring	18.55	18.24	15.82	17.74	17.47	16.69
Summer	37.54	37.20	34.98	40.10	40.13	38.87
Autumn	9.86	9.83	9.50	11.57	10.47	10.46
Winter	19.03	21.60	23.65	19.58	20.84	21.23
No_Visual_Season	92.93	93.08	92.87	93.15	93.19	93.40
Indoor	67.01	67.86	67.85	68.23	68.15	68.70
Outdoor	88.75	88.88	88.86	89.01	88.74	89.07
No_Visual_Place	38.88	39.54	37.74	39.22	39.51	38.39
Plants	60.09	60.75	61.28	60.13	59.57	61.91
Flowers	51.07	51.90	52.34	52.61	53.51	53.94
Trees	50.89	51.10	50.56	49.59	49.50	50.39
Sky	71.19	71.22	71.12	71.62	71.33	71.02
Clouds	71.67	71.82	71.59	72.10	71.53	71.19
Water	50.45	51.20	49.63	50.04	50.60	48.63
Lake	31.18	**32.39**	27.44	25.86	26.82	25.78
River	16.73	17.97	19.58	17.88	20.34	23.36
Sea	43.97	43.29	40.31	44.22	44.45	41.00
Mountains	28.07	29.22	26.56	30.74	31.29	30.85
Day	85.93	86.09	85.62	85.99	86.09	85.97
Night	49.62	50.39	50.40	49.45	49.50	48.95
No_Visual_Time	69.97	70.35	70.72	70.16	70.20	71.11
Sunny	25.45	25.91	27.49	26.91	27.65	27.07
Sunset_Sunrise	70.41	**70.70**	69.84	70.42	70.59	70.42
Canvas	13.17	15.84	16.30	11.56	12.82	14.65
Still_Life	29.01	30.41	28.88	28.61	29.19	25.57
Macro	23.46	23.93	21.93	20.91	21.58	19.69
Portrait	50.27	50.28	48.92	48.97	49.18	49.61
Overexposed	3.75	4.39	5.01	3.71	4.36	**7.66**
Underexposed	36.58	36.64	35.03	35.60	36.03	**34.67**
Neutral_Illumination	97.68	97.73	97.65	97.45	97.49	**97.44**
Motion_Blur	10.92	10.46	10.39	9.08	8.78	7.48
Out_of_focus	4.56	4.62	4.38	4.38	4.30	4.07
Partly_Blurred	66.29	66.82	66.50	66.27	66.13	67.29
No_Blur	87.12	87.38	87.11	87.30	87.19	86.80
Single_Person	57.08	57.46	57.12	55.24	55.06	56.90
Small_Group	23.80	24.21	24.74	22.76	22.69	25.08
Big_Group	21.29	20.92	22.56	20.24	20.49	24.04
No_Persons	88.41	88.66	88.61	88.66	88.19	88.87
Animals	43.91	44.49	43.31	42.87	44.08	46.33
Food	31.78	34.27	33.60	31.10	32.82	30.10
Vehicle	31.97	32.35	31.29	31.48	32.57	32.43
Aesthetic_Impression	25.01	24.91	24.80	24.82	24.83	**25.66**
Overall_Quality	36.03	**36.23**	35.95	34.73	34.93	35.55
Fancy	21.05	20.99	20.40	**21.76**	21.64	20.61

AUC / methods	A+avg	A+L1	A+L1.5	A+L2	B+L1.0625	B+avg
average	82.22	81.69	82.19	82.17	81.56	82.03
Partylife	81.69	80.53	**81.93**	81.83	80.32	81.53
Familiy_Friends	**81.91**	80.25	81.57	81.61	80.35	81.53
Beach_Holidays	**90.47**	88.12	90.37	90.41	87.80	88.92
Building_Sights	**84.04**	83.16	84.01	84.03	83.46	84.00
Snow	84.57	85.18	85.08	84.88	84.85	85.77
Citylife	**82.93**	82.00	82.67	82.68	82.35	82.63
Landscape_Nature	**89.23**	88.67	89.12	89.13	88.94	89.01
Sports	66.34	67.72	66.44	66.37	64.14	68.03
Desert	88.84	87.31	88.79	88.80	88.17	89.08
Spring	**83.08**	82.20	82.88	82.92	81.15	82.35
Summer	**81.18**	79.60	80.83	80.88	79.49	80.30
Autumn	**83.96**	84.13	84.00	83.98	83.79	84.24
Winter	**83.78**	83.47	84.00	83.94	83.30	83.58
No_Visual_Season	**81.43**	80.05	81.07	81.09	80.30	80.89
Indoor	**85.19**	84.57	84.97	84.97	84.85	84.90
Outdoor	**87.93**	87.88	88.50	88.51	87.51	87.73
No_Visual_Place	**75.93**	75.49	75.57	75.52	75.60	75.87
Plants	86.34	86.16	**86.47**	86.43	85.97	86.11
Flowers	91.21	90.95	91.21	91.21	90.27	**91.26**
Trees	88.34	88.24	88.41	88.35	88.54	88.25
Sky	**92.11**	92.07	92.07	92.03	91.92	92.04
Clouds	94.60	94.48	94.62	**94.62**	94.05	94.47
Water	89.04	88.65	**89.06**	89.05	88.08	88.85
Lake	80.33	82.74	80.50	80.45	82.94	80.97
River	93.27	93.45	93.54	93.45	93.07	93.46
Sea	92.60	92.17	92.54	92.56	92.39	92.35
Mountains	**89.21**	87.31	89.11	89.16	88.38	88.47
Day	85.03	84.95	**85.26**	85.24	83.88	84.77
Night	88.37	88.01	88.38	88.36	87.98	88.46
No_Visual_Time	83.54	83.27	**83.70**	83.68	82.56	83.41
Sunny	**72.16**	70.01	71.59	71.65	70.66	70.54
Sunset_Sunrise	95.02	95.03	95.07	95.05	94.79	94.85
Canvas	82.83	82.63	82.74	82.77	82.53	82.31
Still_Life	78.53	78.58	78.52	78.49	**78.70**	78.64
Macro	**79.79**	79.30	79.73	79.70	77.81	79.21
Portrait	83.40	82.68	83.44	83.42	82.23	**83.47**
Overexposed	73.46	70.40	73.39	73.41	71.90	73.63
Underexposed	83.67	83.65	83.79	83.72	84.06	83.62
Neutral_Illumination	79.59	79.34	79.66	79.58	79.46	79.49
Motion_Blur	72.85	72.44	**73.06**	72.97	74.26	73.71
Out_of_focus	74.88	70.81	**74.71**	74.79	70.90	73.75
Partly_Blurred	81.38	81.15	**81.43**	81.34	80.73	81.35
No_Blur	82.38	81.86	**82.54**	82.50	81.77	82.11
Single_Person	**79.84**	79.47	79.54	79.52	79.30	79.83
Small_Group	**72.53**	71.16	72.14	72.18	71.61	71.44
Big_Group	**85.71**	86.28	85.72	85.73	84.30	84.72
No_Persons	**83.33**	82.88	83.57	83.60	82.49	82.89
Animals	**84.03**	82.98	83.89	83.91	82.57	83.18
Food	**88.26**	87.37	88.18	88.21	87.28	88.09
Vehicle	83.51	**84.18**	83.61	83.58	83.24	83.24
Aesthetic_Impression	58.80	58.85	58.54	58.48	59.29	58.84
Overall_Quality	60.84	60.26	59.94	59.88	60.70	**60.91**
Fancy	54.22	55.20	54.40	54.29	**55.55**	54.78

AUC / methods	B+L1	B+L1.5	B+L2	C+L1	C+L1.0625	C+L2
average	81.49	81.96	81.97	80.50	80.70	80.04
Partylife	80.34	81.60	81.58	76.27	77.11	77.80
Familiy_Friends	80.09	81.25	81.28	77.97	78.09	77.05
Beach_Holidays	87.86	88.85	88.87	84.98	83.97	80.45
Building_Sights	83.03	83.86	83.87	82.59	81.99	82.06
Snow	85.19	85.96	85.89	84.53	85.83	85.30
Citylife	81.85	82.36	82.36	81.03	81.08	80.22
Landscape_Nature	88.58	88.90	88.89	88.70	88.43	88.26
Sports	67.75	67.88	67.96	65.52	65.92	**68.97**
Desert	87.33	89.03	89.08	85.94	84.73	78.61
Spring	81.50	82.17	82.23	79.86	79.77	78.35
Summer	79.02	79.94	79.97	76.77	77.31	75.93
Autumn	84.10	84.18	84.20	79.36	78.75	77.31
Winter	83.22	83.75	83.71	84.12	84.67	82.98
No_Visual_Season	79.73	80.58	80.60	78.56	78.93	78.20
Indoor	84.49	84.66	84.64	84.12	83.94	83.42
Outdoor	87.59	88.20	88.22	87.03	86.72	87.00
No_Visual_Place	75.04	75.52	75.50	75.10	75.35	74.12
Plants	85.85	86.16	86.15	84.41	84.42	84.22
Flowers	90.77	91.20	91.21	89.90	89.76	89.00
Trees	88.20	88.33	88.28	87.97	87.92	86.78
Sky	91.91	91.94	91.93	91.91	91.81	91.39
Clouds	94.28	94.44	94.46	94.07	93.96	93.98
Water	88.47	88.87	88.89	87.28	87.27	85.92
Lake	82.73	81.06	81.03	83.03	83.75	83.52
River	93.49	**93.61**	93.55	91.43	91.77	89.30
Sea	92.03	92.27	92.29	92.87	92.79	90.59
Mountains	87.28	88.42	88.44	87.35	88.19	87.64
Day	84.62	84.93	84.92	83.27	83.33	82.78
Night	87.98	**88.47**	88.47	88.02	88.24	87.33
No_Visual_Time	82.98	83.43	83.43	82.01	81.95	81.69
Sunny	69.09	70.08	70.08	68.85	69.60	68.20
Sunset_Sunrise	94.78	94.90	94.89	95.06	94.97	94.35
Canvas	82.70	82.36	82.33	81.48	82.15	82.79
Still_Life	78.60	78.62	78.59	76.33	76.60	75.78
Macro	78.80	79.07	79.05	76.65	77.69	78.66
Portrait	82.55	83.41	83.46	79.65	79.94	79.95
Overexposed	70.42	73.43	73.53	71.22	74.01	76.61
Underexposed	83.54	83.58	83.56	85.34	85.82	84.76
Neutral_Illumination	79.30	79.41	79.42	80.91	81.27	80.21
Motion_Blur	72.45	73.75	73.73	72.21	72.78	71.16
Out_of_focus	70.52	73.48	73.59	69.84	70.20	70.40
Partly_Blurred	81.03	81.27	81.20	80.04	79.94	79.45
No_Blur	81.68	82.17	82.14	81.19	80.95	80.50
Single_Person	79.39	79.47	79.46	78.19	77.87	78.01
Small_Group	70.56	71.17	71.20	69.19	69.62	69.50
Big_Group	85.93	84.86	84.80	80.78	81.79	81.40
No_Persons	82.37	83.09	83.12	81.28	80.83	81.21
Animals	82.53	83.08	83.10	79.62	79.76	79.46
Food	87.16	88.04	88.06	85.27	85.71	85.50
Vehicle	83.99	83.36	83.31	82.76	82.76	81.66
Aesthetic_Impression	58.85	58.72	58.69	59.16	**59.43**	59.05
Overall_Quality	60.26	60.17	60.16	60.73	60.86	59.67
Fancy	55.20	54.83	54.77	54.73	54.68	53.65

AUC / methods	D+L1	D+L1.0625	D+L2	E+L1	E+L1.0625	E+L2
average	80.75	81.11	80.78	80.50	80.64	80.66
Partylife	78.10	78.76	78.56	76.60	77.75	79.03
Familiy_Friends	78.36	78.88	78.13	78.21	78.49	78.31
Beach_Holidays	85.76	85.48	83.74	85.66	85.61	88.17
Building_Sights	82.47	83.22	82.80	82.41	81.89	82.32
Snow	82.96	84.19	**86.24**	83.35	83.55	82.85
Citylife	81.26	81.66	81.30	81.18	81.24	81.45
Landscape_Nature	88.50	88.77	88.97	88.55	88.22	88.60
Sports	65.47	64.94	66.17	65.50	65.22	64.08
Desert	86.14	87.32	87.62	**89.50**	88.11	87.25
Spring	79.13	78.45	75.68	79.43	79.45	78.60
Summer	76.48	76.63	75.55	77.67	77.99	77.94
Autumn	80.63	80.08	79.74	79.59	79.34	78.75
Winter	81.91	82.97	83.15	83.59	83.95	83.25
No_Visual_Season	78.20	78.63	78.30	78.93	79.05	79.40
Indoor	84.10	84.34	83.87	84.49	84.37	84.32
Outdoor	86.89	87.04	87.10	87.22	86.91	87.29
No_Visual_Place	74.59	75.28	74.93	75.19	75.49	74.98
Plants	84.62	84.94	84.88	84.60	84.55	85.64
Flowers	89.98	90.08	89.20	89.75	89.84	89.68
Trees	88.36	**88.66**	88.35	87.79	87.72	87.74
Sky	91.81	91.86	91.75	91.77	91.61	91.45
Clouds	93.79	93.92	93.76	93.96	93.81	93.59
Water	87.21	87.40	86.47	87.18	87.29	86.11
Lake	84.61	85.14	**86.23**	82.28	82.50	83.34
River	92.65	92.67	90.84	91.69	92.24	91.99
Sea	92.86	92.81	91.48	92.96	**93.04**	91.68
Mountains	86.72	87.60	87.38	88.42	89.13	88.85
Day	83.08	83.31	82.91	83.09	83.18	83.06
Night	87.62	88.00	87.82	86.90	87.02	86.18
No_Visual_Time	81.80	82.22	81.92	81.91	81.94	82.08
Sunny	68.86	69.54	69.57	69.18	69.89	69.80
Sunset_Sunrise	95.20	95.29	95.09	**95.59**	95.55	95.55
Canvas	84.35	**84.69**	84.36	81.23	80.76	80.34
Still_Life	78.05	78.28	77.26	76.35	76.78	77.21
Macro	76.89	77.08	76.62	76.12	76.28	75.30
Portrait	81.02	81.09	80.50	79.83	80.01	81.46
Overexposed	70.06	73.56	**77.29**	71.25	74.35	80.14
Underexposed	85.32	85.57	**84.99**	83.80	84.23	83.22
Neutral_Illumination	80.82	81.16	**80.26**	79.28	79.75	79.16
Motion_Blur	71.74	72.23	71.49	72.16	72.68	70.19
Out_of_focus	71.58	72.28	71.77	69.67	69.75	71.36
Partly_Blurred	79.98	80.32	80.01	80.02	79.85	80.25
No_Blur	81.08	81.46	81.32	81.16	80.85	81.03
Single_Person	78.75	78.97	78.81	78.17	77.93	78.44
Small_Group	70.83	70.96	70.30	69.40	69.81	70.31
Big_Group	83.51	83.33	82.76	79.97	81.32	81.48
No_Persons	81.11	81.56	81.48	81.56	81.02	82.26
Animals	80.73	80.77	80.48	79.70	79.96	80.37
Food	86.12	86.82	86.88	85.18	85.27	84.32
Vehicle	83.38	83.35	82.60	82.72	82.87	82.58
Aesthetic_Impression	58.96	59.28	58.72	59.10	59.25	58.60
Overall_Quality	60.42	60.61	59.99	60.65	60.62	60.29
Fancy	55.11	55.19	53.92	54.80	54.80	53.60

Acknowledgements

This work was supported in part by Federal Ministry of Economics and Technology of Germany under the project THESEUS (01MQ07018).

References

[1] Anna Bosch, Andrew Zisserman, and Xavier Muñoz. Representing shape with a spatial pyramid kernel. In *Proceedings of the 6th ACM international conference on Image and video retrieval (CIVR '07)*, pages 401–408, 2007.

[2] G. Csurka, C. Bray, C. Dance, and L. Fan. Visual categorization with bags of keypoints. In *Workshop on Statistical Learning in Computer Vision, ECCV*, pages 1–22, Prague, Czech Republic, May 2004.

[3] M. Everingham, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The PASCAL Visual Object Classes Challenge 2008 (VOC2008) Results. http://www.pascal-network.org/challenges/VOC/voc2008/workshop/index.html, 2008.

[4] Marius Kloft, Ulf Brefeld, Pavel Laskov, and Sören Sonnenburg. Non-sparse multiple kernel learning, 2008. Abstract of the NIPS workshop on Kernel Learning: Automatic Selection of Optimal Kernels.

[5] Svetlana Lazebnik, Cordelia Schmid, and Jean Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, volume 2, pages 2169–2178, New York, USA, 2006.

[6] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, 2004.

[7] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27(10):1615–1630, 2005.

[8] S. Sonnenburg. The shogun toolbox, http://www.shogun-toolbox.org/.

[9] S. Sonnenburg, G. Rätsch, C. Schäfer, and B. Schölkopf. Large scale multiple kernel learning. *Journal of Machine Learning Research*, 7:1531–1565, 07 2006.

[10] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (in press), 2010.