Recognizing Image Style: Supplementary Materials

Sergey Karayev[1]
Matthew Trentacoste[2]
Helen Han[1]
Aseem Agarwala[2]
Trevor Darrell[1]
Aaron Hertzmann[2]
Holger Winnemoeller[2]

[1] University of California, Berkeley
[2] Adobe

List of Figures

1 Top five most confident predictions on the Flickr Style test set: styles 1-8. . 3
2 Top five most confident predictions on the Flickr Style test set: styles 9-14. 4
3 Top five most confident predictions on the Flickr Style test set: styles 15-20. 5
4 Distribution of image style, genre, and date in the Wikipaintings dataset. . . 7
5 Confusion matrix of our best classifier on the AVA Style dataset. 11
6 Confusion matrix of our best classifier on the Flickr dataset. 12
7 Confusion matrix of our best classifier on the Wikipaintings dataset. . . . 13

List of Tables

1 Exact Flickr group names, and their sizes. 2
2 All per-class APs on all evaluated features on the AVA Style dataset. . . . 6
3 All per-class APs on all evaluated features on the Flickr dataset. 6
4 All per-class APs on all evaluated features on the Wikipaintings dataset. . . 8
5 Comparison of Flickr Style per-class accuracies for our method and Mech Turkers. 9
6 Per-class accuracies on the Wikipaintings dataset, using the MC-bit feature. 10

1 MTurk Study Details

Test images were grouped into 10 images per Human Interface Task (HIT). Each task asks the Turker to evaluate the style (e.g., "Is this image VINTAGE?") for each image. For each style, we provided a short blurb describing the style in words, and provided 12-15 hand-chosen positive and negative examples for each Flickr Group. Each HIT included 2 sentinels:

© 2014. The copyright of this document resides with its authors.
It may be distributed unchanged freely in print or electronic forms.

Table 1: Exact Flickr group names, and their sizes.

Style	Group names [num images]
Bokeh	Bokeh Photography (1/day) [187K]
Bright	Colour Mania [100K]
Depth of Field	Depth of Field [116K], Finest DoF [54K]
Detailed	Details aller Art - Details of all kind [22K], Detail pictures [5K]
Ethereal	Ethereal World [21K]
Geometric Composition	Geometric Beauty [168K]
Hazy	Misty hazy smokey [14K]
HDR	HDR ADDICTED [374K]
Horror	Horror [16K]
Long Exposure	Long Exposure [619K]
Macro	Closer and Closer Macro Photography [990K]
Melancholy	melancholy [106K]
Minimal	Less Is More... [44K]
Noir	Film Noir Mood [7K]
Romantic	Romantic Images [20K]
Serene	S̃erene ̃[68K]
Pastel	pastel and dreamy [120K], Pastel Soft tone [7K]
Sunny	Sun, sun and more sun [23K]
Texture	Texture [103K]
Vintage	Vintage Feelings [4K], Vintage & Retro [61K]

images which were very clearly positives and similar to the examples. HITs were rejected when Turkers got both sentinels wrong. Turkers were paid 0.10 per HIT, and were allowed to perform multiple hits. Manual inspection of the results indicate that the Turkers understood the task and were performing effectively. A few Turkers sent unsolicited feedback indicating that they were really enjoying the HITs (“some of the photos are beautiful”) and wanted to perform them as effectively as possible.

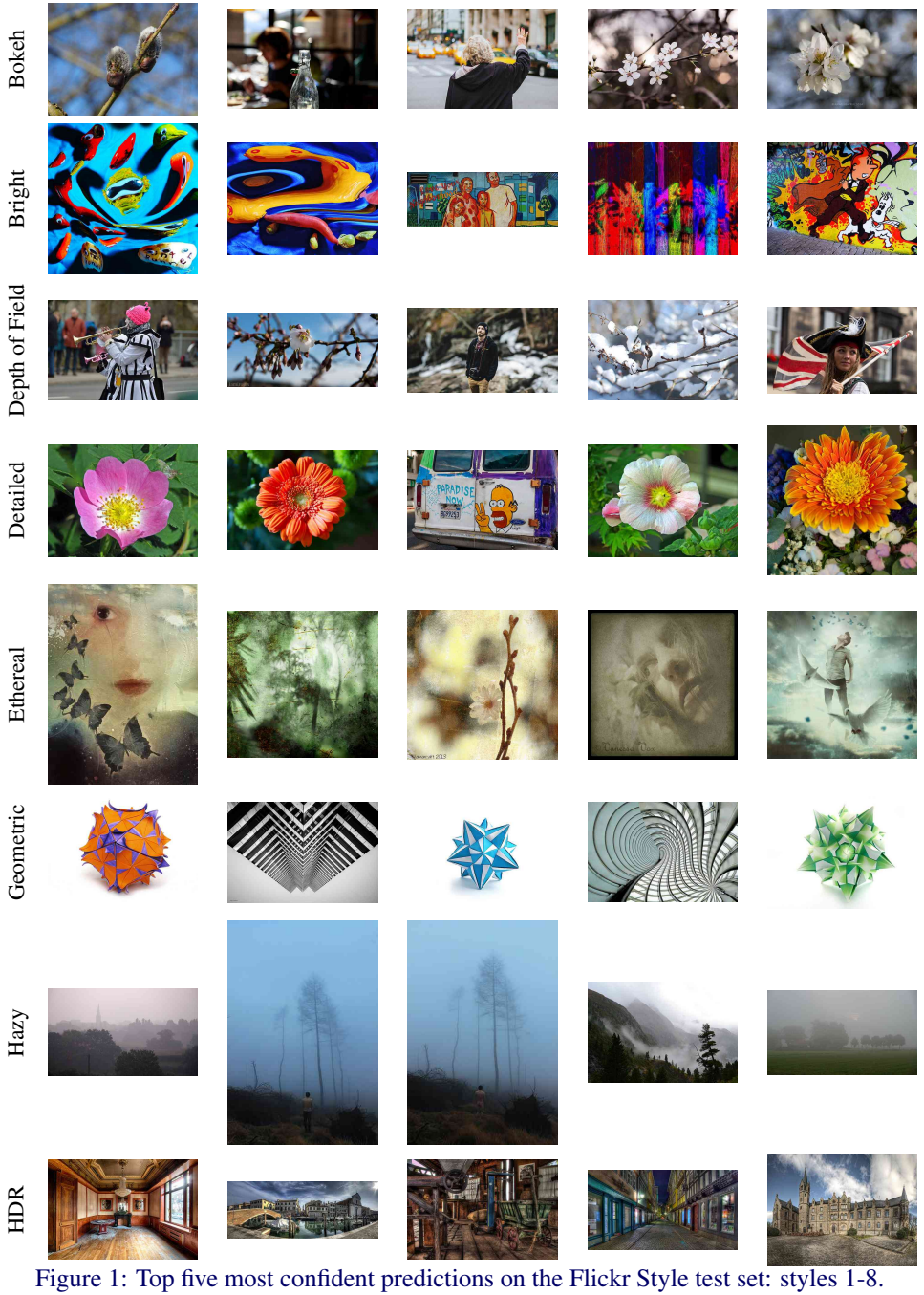

Figure 1: Top five most confident predictions on the Flickr Style test set: styles 1-8.

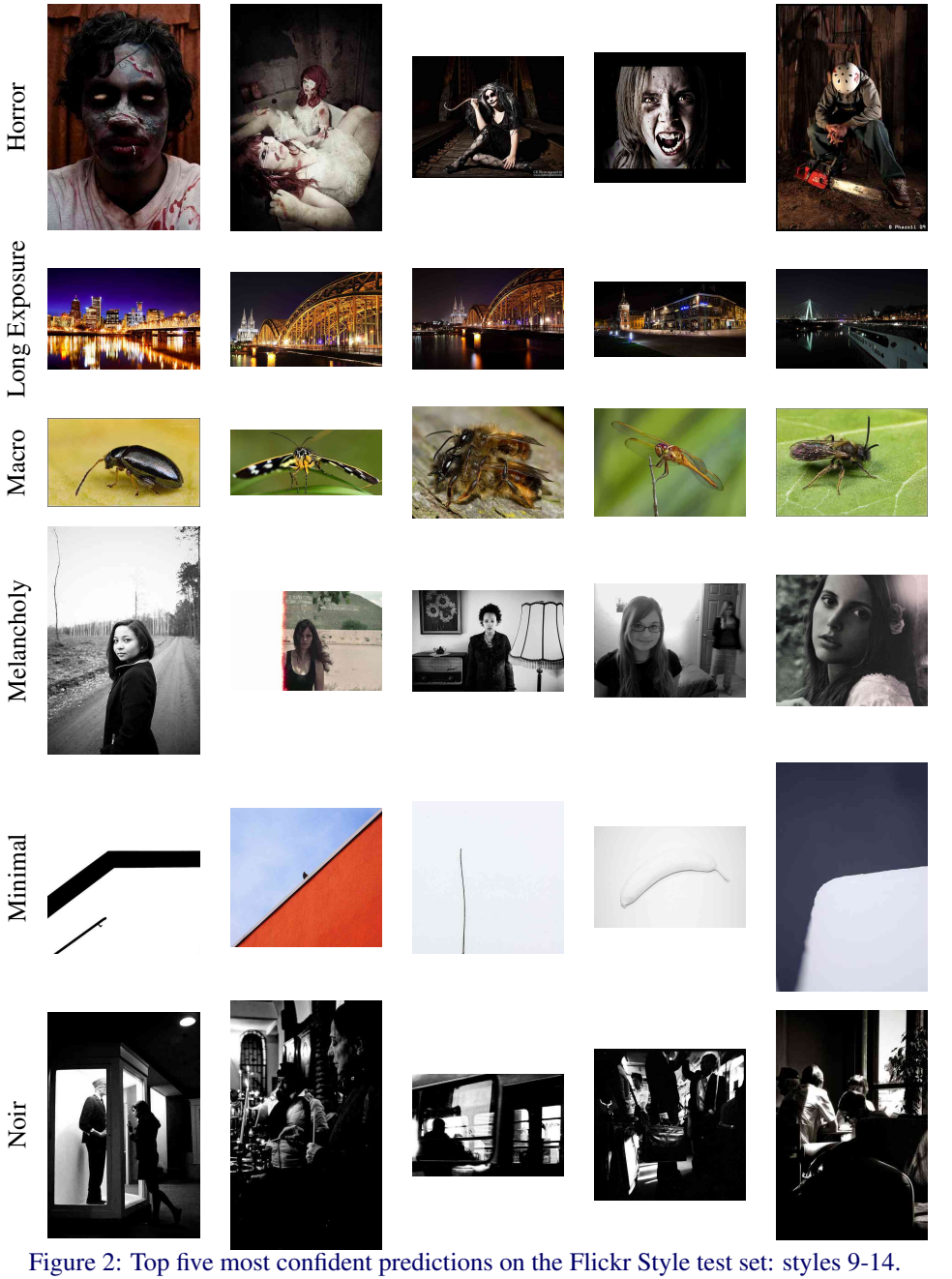

Figure 2: Top five most confident predictions on the Flickr Style test set: styles 9-14.

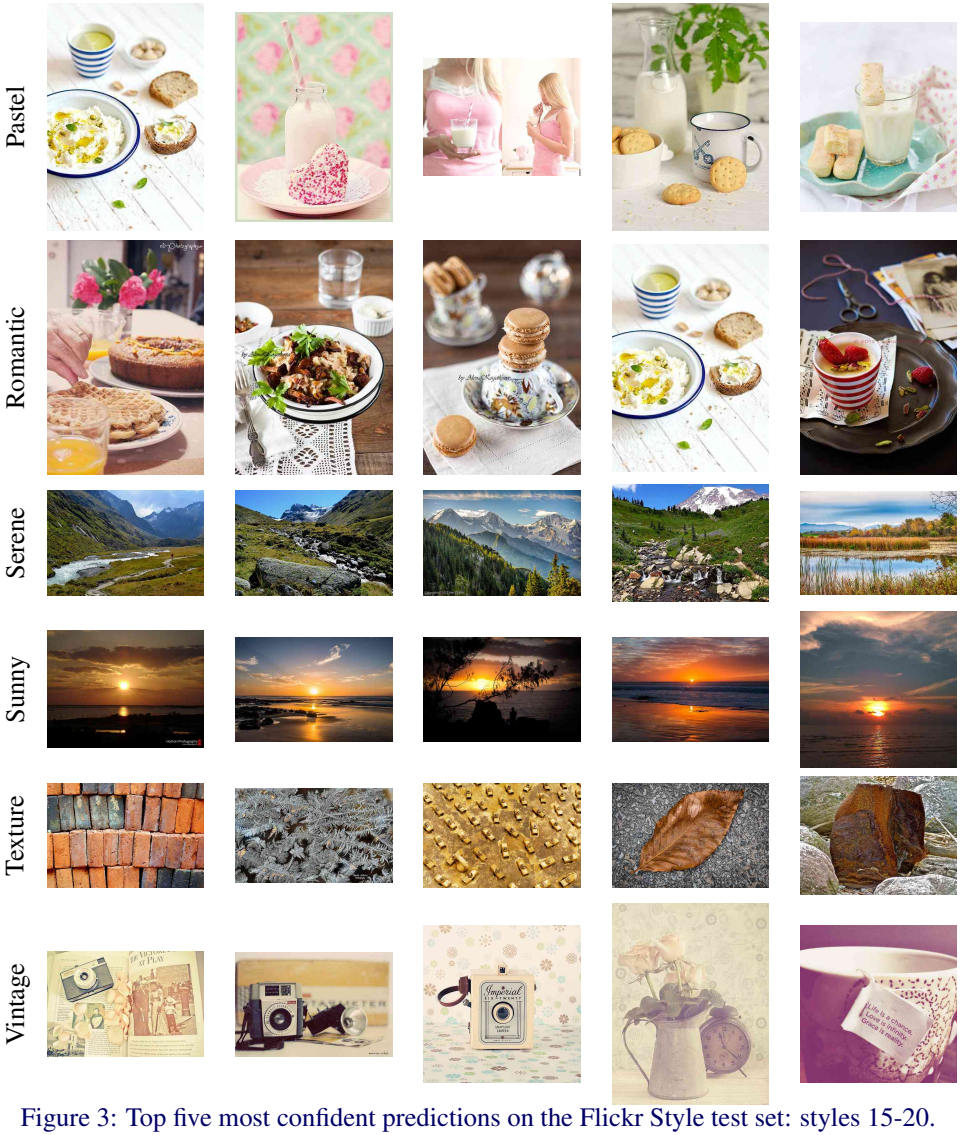

Figure 3: Top five most confident predictions on the Flickr Style test set: styles 15-20.

Table 2: All per-class APs on all evaluated features on the AVA Style dataset.

	Fusion	$DeCAF_6$	MC-bit	Murray	$DeCAF_5$	ImageNet	L*a*b*
Complementary_Colors	0.469	0.548	0.329	0.440	0.368	0.389	0.294
Duotones	0.676	0.737	0.612	0.510	0.363	0.383	0.582
HDR	0.669	0.594	0.624	0.640	0.494	0.335	0.194
Image_Grain	0.647	0.545	0.744	0.740	0.535	0.219	0.213
Light_On_White	0.908	0.915	0.802	0.730	0.805	0.508	0.867
Long_Exposure	0.453	0.431	0.420	0.430	0.208	0.242	0.232
Macro	0.478	0.427	0.413	0.500	0.376	0.438	0.230
Motion_Blur	0.478	0.467	0.458	0.400	0.327	0.186	0.117
Negative_Image	0.595	0.619	0.499	0.690	0.427	0.323	0.268
Rule_of_Thirds	0.352	0.353	0.236	0.300	0.269	0.244	0.188
Shallow_DOF	0.624	0.659	0.637	0.480	0.522	0.517	0.332
Silhouettes	0.791	0.801	0.801	0.720	0.609	0.401	0.261
Soft_Focus	0.312	0.354	0.290	0.390	0.225	0.170	0.127
Vanishing_Point	0.684	0.658	0.685	0.570	0.527	0.542	0.123
mean	0.581	0.579	0.539	0.539	0.432	0.350	0.288

Table 3: All per-class APs on all evaluated features on the Flickr dataset.

	Fusion x Content	$DeCAF_6$	MC-bit	$DeCAF_5$	Imagenet
Bokeh	0.281	0.262	0.248	0.253	-
Bright,_Energetic	0.355	0.331	0.250	0.313	0.231
Depth_of_Field	0.266	0.241	0.230	0.208	0.202
Detailed	0.289	0.277	0.279	0.277	-
Ethereal	0.418	0.365	0.328	0.356	0.190
Geometric_Composition	0.442	0.395	0.399	0.369	0.347
HDR	0.548	0.477	0.527	0.332	0.293
Hazy	0.565	0.506	0.489	0.386	0.330
Horror	0.479	0.464	0.304	0.337	0.286
Long_Exposure	0.469	0.388	0.426	0.300	0.254
Macro	0.684	0.683	0.620	0.588	0.640
Melancholy	0.178	0.157	0.169	0.096	0.131
Minimal	0.498	0.465	0.452	0.319	0.281
Noir	0.529	0.521	0.409	0.372	0.290
Romantic	0.200	0.206	0.162	0.140	0.185
Serene	0.209	0.191	0.219	0.142	0.175
Soft,_Pastel	0.309	0.317	0.267	0.269	0.272
Sunny	0.550	0.540	0.523	0.481	0.388
Vintage	0.421	0.385	0.348	0.309	0.268
mean	0.405	0.377	0.350	0.308	0.280

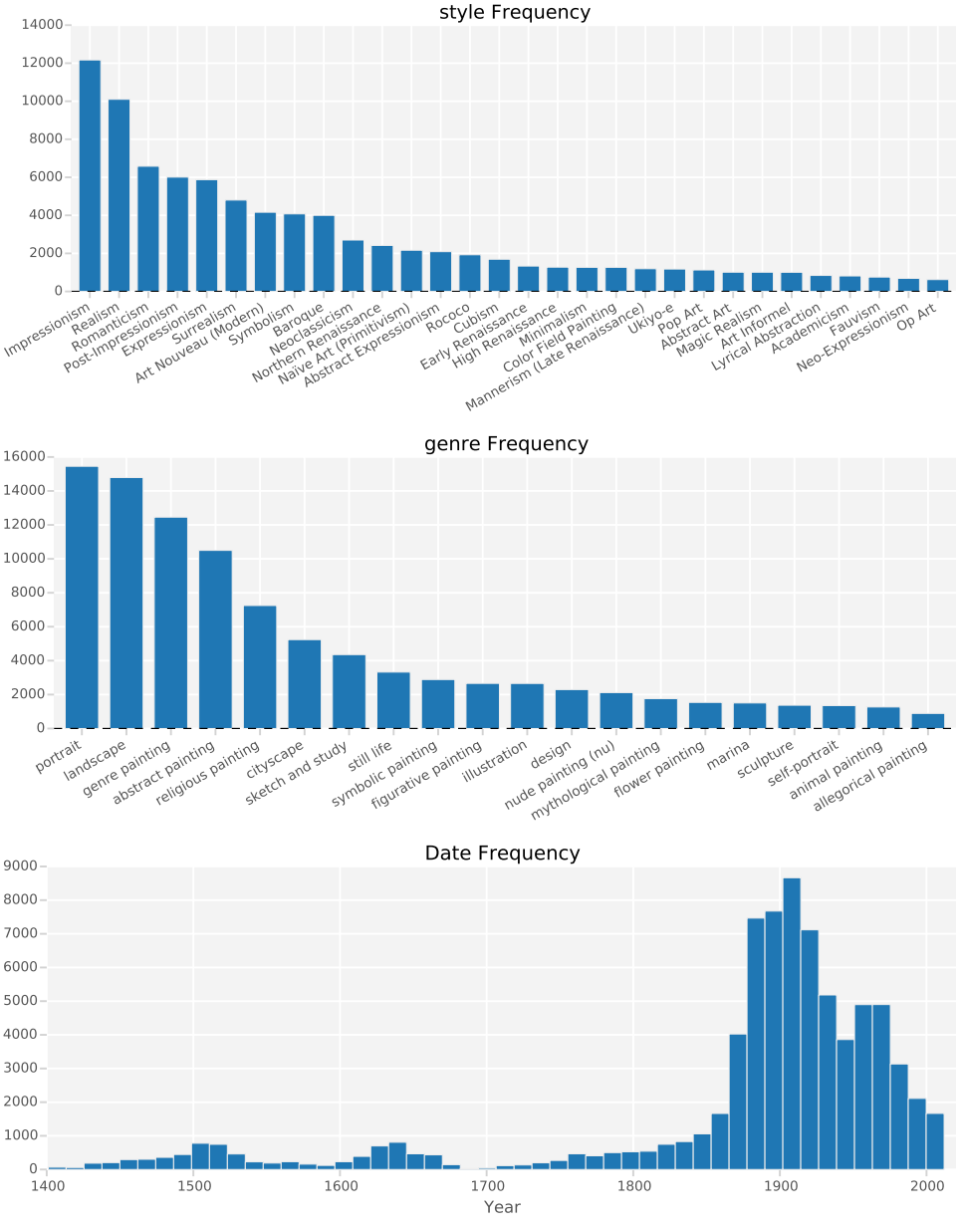

Figure 4: Distribution of image style, genre, and date in the Wikipaintings dataset.

Table 4: All per-class APs on all evaluated features on the Wikipaintings dataset.

	Fusion x Content	MC-bit	$DeCAF_6$	ImageNet
Abstract_Art	0.341	0.314	0.258	0.192
Abstract_Expressionism	0.351	0.340	0.243	0.159
Art_Informel	0.221	0.217	0.187	0.138
Art_Nouveau_(Modern)	0.421	0.402	0.197	0.096
Baroque	0.436	0.386	0.313	0.162
Color_Field_Painting	0.773	0.739	0.689	0.503
Cubism	0.495	0.488	0.400	0.193
Early_Renaissance	0.578	0.559	0.453	0.192
Expressionism	0.235	0.230	0.186	0.093
High_Renaissance	0.401	0.345	0.288	0.165
Impressionism	0.586	0.528	0.411	0.227
Magic_Realism	0.521	0.465	0.428	0.198
Mannerism_(Late_Renaissance)	0.505	0.439	0.356	0.171
Minimalism	0.660	0.614	0.604	0.449
Nave_Art_(Primitivism)	0.395	0.425	0.225	0.111
Neoclassicism	0.601	0.537	0.399	0.179
Northern_Renaissance	0.560	0.478	0.433	0.119
Pop_Art	0.441	0.398	0.281	0.163
Post-Impressionism	0.348	0.348	0.292	0.135
Realism	0.408	0.309	0.266	0.159
Rococo	0.616	0.548	0.467	0.242
Romanticism	0.392	0.389	0.343	0.185
Surrealism	0.262	0.247	0.134	0.099
Symbolism	0.390	0.390	0.260	0.172
Ukiyo-e	0.895	0.894	0.788	0.260
mean	0.473	0.441	0.356	0.191

Table 5: Comparison of Flickr Style per-class accuracies for our method and Mech Turkers. We first give the full results table, then show the signficant deviations between human and machine performance, and between using Flickr and MTurk ground truth.

	MTurk acc., Flickr g.t.	Our acc., Flickr g.t.	Our acc., MTurk g.t.
Bright	69.10	73.38	73.63
Depth of Field	68.92	68.50	81.05
Detailed	65.47	75.25	68.44
Ethereal	76.92	80.62	77.95
Geometric Composition	81.52	77.75	80.31
HDR	71.84	82.00	76.96
Hazy	83.49	80.75	81.64
Horror	89.85	84.25	81.64
Long Exposure	73.12	84.19	76.79
Macro	92.25	86.56	88.39
Melancholy	67.77	70.88	71.25
Minimal	79.71	83.75	78.57
Noir	81.35	85.25	85.88
Pastel	66.94	74.56	75.47
Romantic	60.91	68.00	66.25
Serene	69.49	70.44	76.80
Sunny	84.48	84.56	79.94
Vintage	68.77	75.50	67.80
Mean	75.11	78.12	77.15

	Our acc., Flickr g.t.	Our acc., MTurk g.t.	% change from Flickr to MTurk g.t.
Vintage	75.50	67.80	-10.19
Detailed	75.25	68.44	-9.05
Long Exposure	84.19	76.79	-8.79
Minimal	83.75	78.57	-6.18
HDR	82.00	76.96	-6.15
Sunny	84.56	79.94	-5.46
Serene	70.44	76.80	9.03
Depth of Field	68.50	81.05	18.32

	Our acc., Flickr g.t.	MTurk acc., Flickr g.t.	Acc. difference
Horror	84.25	90.42	-6.17
Macro	86.56	91.71	-5.15
Romantic	68.00	61.04	6.96
Pastel	74.56	66.87	7.69
HDR	82.00	72.79	9.21
Long Exposure	84.19	73.83	10.35
Detailed	75.25	63.30	11.95

Table 6: Per-class accuracies on the Wikipaintings dataset, using the MC-bit feature.

Style	Accuracy	Style	Accuracy
Symbolism	71.24	Impressionism	82.15
Expressionism	72.03	Northern Renaissance	82.32
Art Nouveau (Modern)	72.77	High Renaissance	82.90
Nave Art (Primitivism)	72.95	Mannerism (Late Renaissance)	83.04
Surrealism	74.44	Pop Art	83.33
Post-Impressionism	74.51	Early Renaissance	84.69
Romanticism	75.86	Abstract Art	85.10
Realism	75.88	Cubism	86.85
Magic Realism	78.54	Rococo	87.33
Neoclassicism	80.18	Ukiyo-e	93.18
Abstract Expressionism	81.25	Minimalism	94.21
Baroque	81.45	Color Field Painting	95.58
Art Informel	82.09		

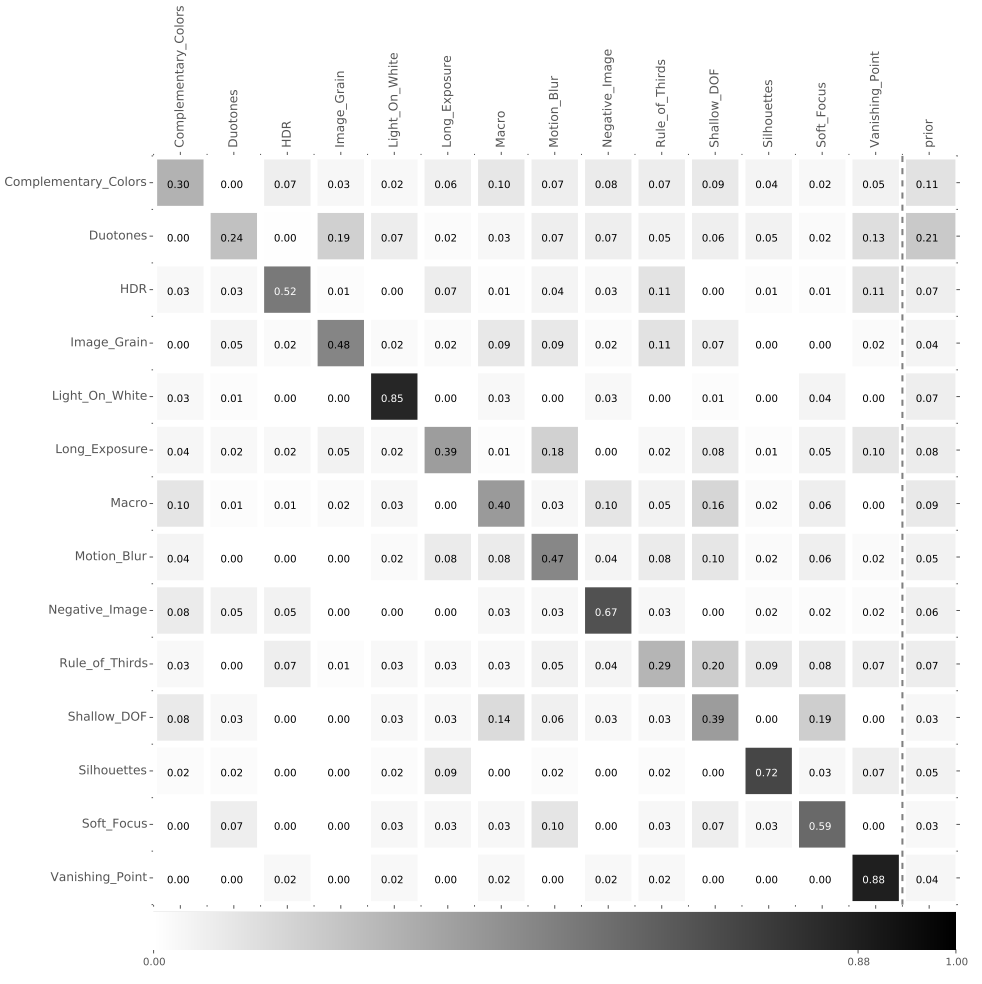

Figure 5: Confusion matrix of our best classifier (Late-fusion $\times$ Content) on the AVA Style dataset. The right-most "prior" column reflects the distribution of ground-truth labels in the test set. The confusions are mostly understandable: "Soft Focus" vs. "Depth of Field" for example.

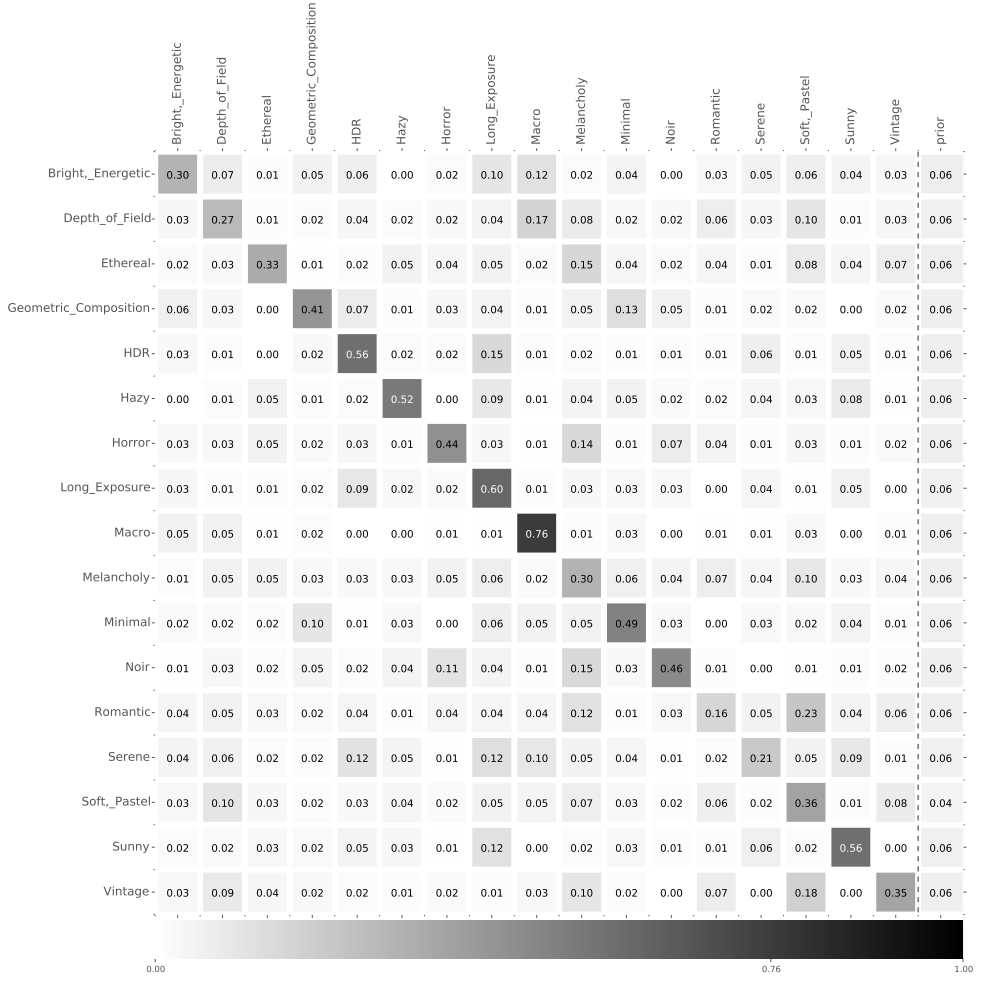

Figure 6: Confusion matrix of our best classifier (Late-fusion × Content) on the Flickr dataset.

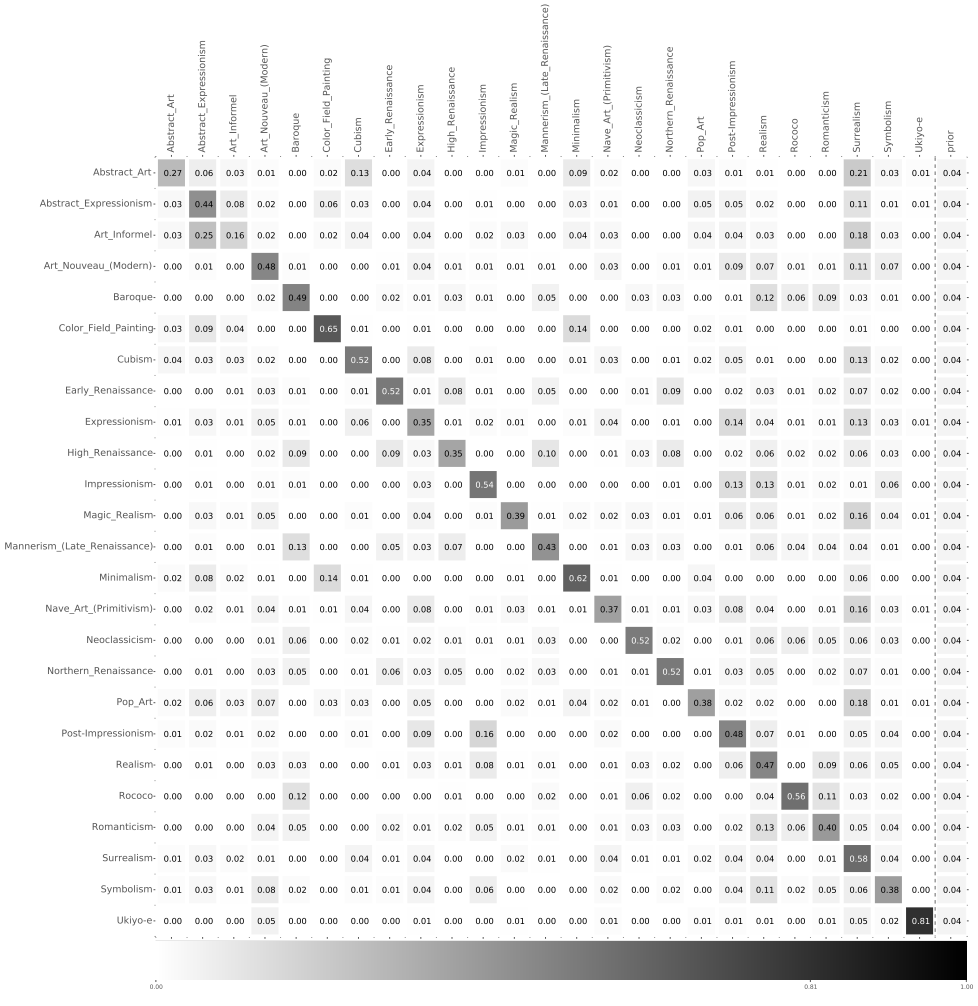

Figure 7: Confusion matrix of our best classifier (Late-fusion $\times$ Content) on the Wikipaintings dataset.