468: Recognizing Image Style: Supplemental Materials

Anonymous

May 9, 2014

Contents

List of Figures

List of Tables

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: 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.

2 Image Features

In order to classify styles, we must choose appropriate image features. We hypothesize that image style may be related to many different features, including low-level statistics [LRF04], color choices, composition, and content. Hence, we test features that embody these different elements, including features from the object recognition literature. We evaluate single-feature performance, as well as second-stage fusion of multiple features.

L*a*b color histogram. Many of the Flickr styles exhibit strong dependence on color. For example, *Noir* images are nearly all black-and-white, while most *Horror* images are very dark, and *Vintage* images use old photographic colors. We use a standard color histogram feature, computed on the whole image. The 784-dimensional joint histogram in CIELAB color space has 4, 14, and 14 bins in the L*, a*, and b* channels, following Palermo et al. [PHE12], who showed this to be the best performing single feature for determining the date of historical color images.

GIST. The classic gist descriptor [OT01] is known to perform well for scene classification and retrieval of images visually similar at a low-resolution scale, and thus can represent image composition to some extent. We use the INRIA LEAR implementation, resizing images to 256 by 256 pixels and extracting a 960-dimensional color GIST feature.

Graph-based visual saliency. We also model composition with a visual attention feature [HKP06]. The feature is fast to compute and has been shown to predict human fixations in natural images basically as well as an individual human (humans are far better in aggregate, however). The 1024-dimensional feature is computed from images resized to 256 by 256 pixels.

Meta-class binary features. Image content can be predictive of individual styles, e.g., *Macro* images include many images of insects and flowers. The `mc-bit` feature [BT12] is a 15,000-dimensional bit vector feature learned as a non-linear combination of classifiers trained using existing features (e.g., SIFT, GIST, Self-Similarity) on thousands of random ImageNet synsets, including internal ILSVRC2010 nodes. In essence, MC-bit is a hand-crafted "deep" architecture, stacking classifiers and pooling operations on top of lower-level features.

Deep convolutional net. Current state-of-the-art results on ImageNet, the largest image classification challenge, have come from a deep convolutional network trained in a fully-supervised manner [KSH12]. We use the Caffe [Jia13] open-source implementation of the ImageNet-winning eght-layer convolutional network, trained on over a million images annotated with 1,000 ImageNet classes. We investigate using features from two different levels of the network, referred to as DeCAF_6 and DeCAF_7 (following [DJV$^+$13]). Both features are 4,000-dimensional and are close to the supervised signal, and are computed from images center-cropped and resized to 256 by 256 pixels.

Content classifiers. Following Dhar et al. [DOB11], who use high-level classifiers as features for their aesthetic rating prediction task, we evaluate using object classifier confidences as features. Specifically, we train classifiers for all 20 classes of the PASCAL VOC [EVW$^+$10] using the DeCAF_6 feature. The resulting classifiers are quite reliable, obtaining 0.7 mean AP on the VOC 2012.

We aggregate the data to train four classifiers for “animals”, “vehicles”, “indoor objects” and “people”. These aggregate classes are presumed to discriminate between vastly different types of images – types for which different style signals may apply. For example, a *Romantic* scene with people may be largely about the composition of the scene, whereas, *Romantic* scenes with vehicles may be largely described by color.

To enable our classifiers to learn content-dependent style, we can take the outer product of a feature channel with the four aggregate content classifiers.

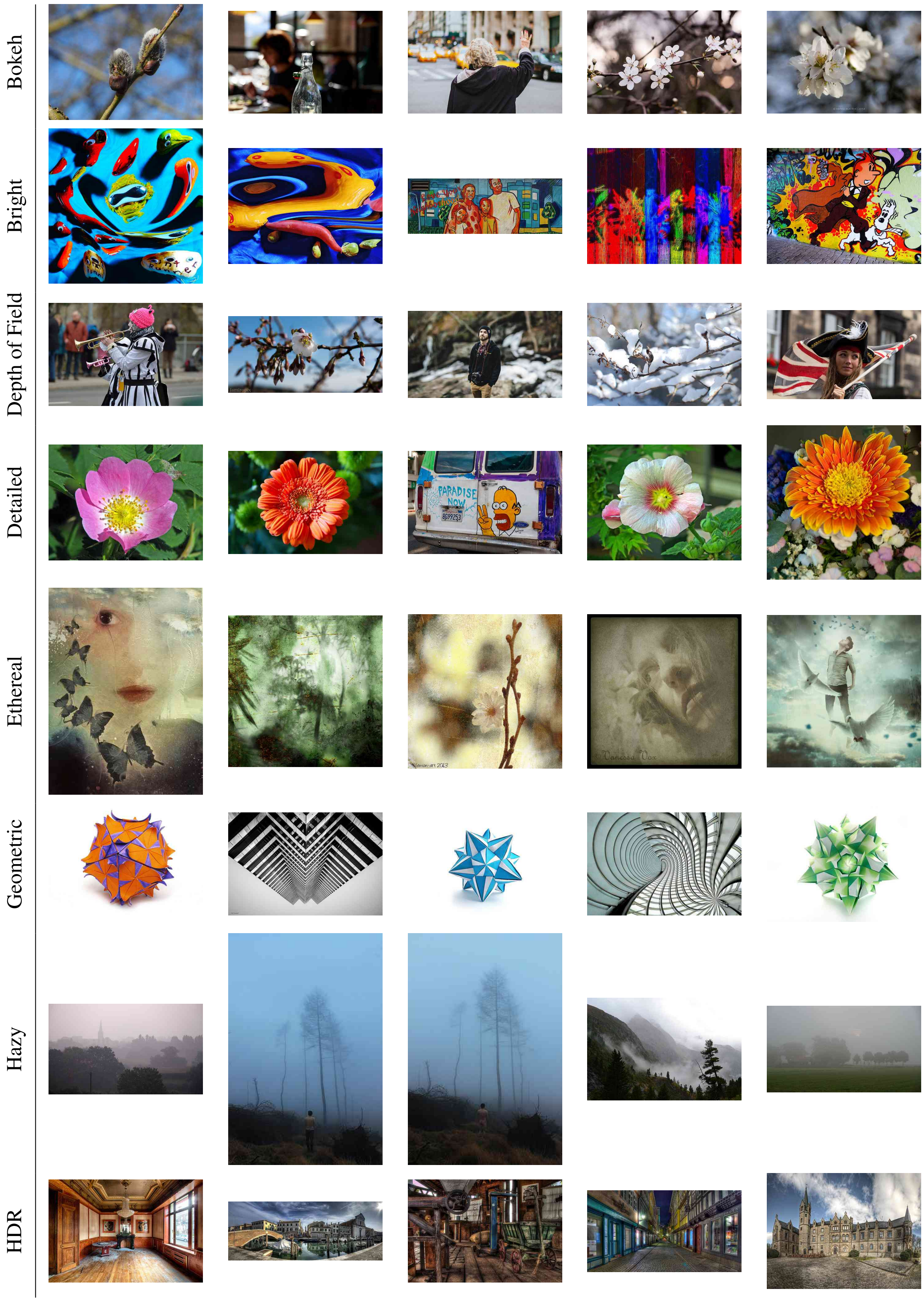

Figure 1: Top five most confident predictions on the Flickr Style test set: styles 1-8.

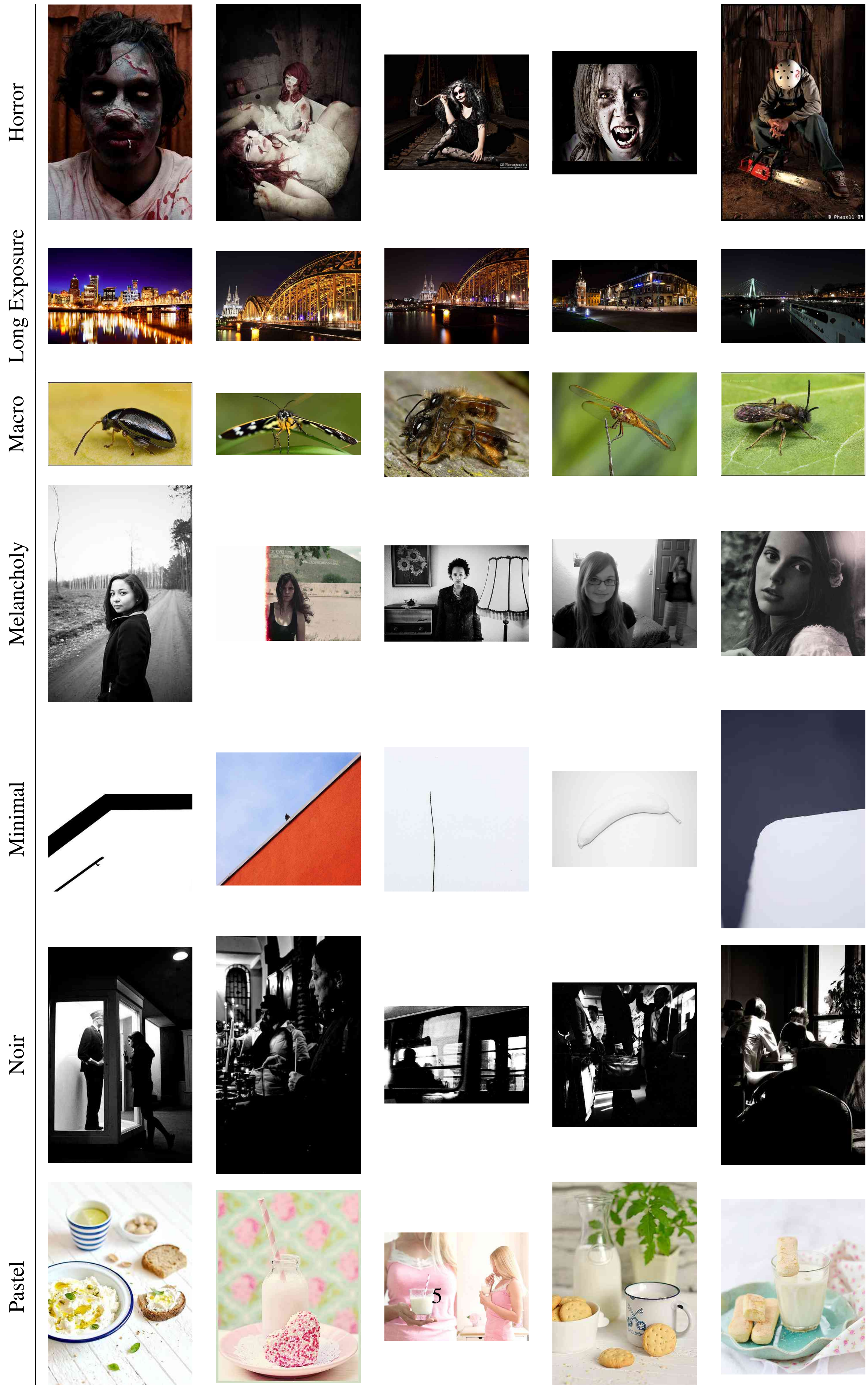

Figure 2: Top five most confident predictions on the Flickr Style test set: styles 9-15.

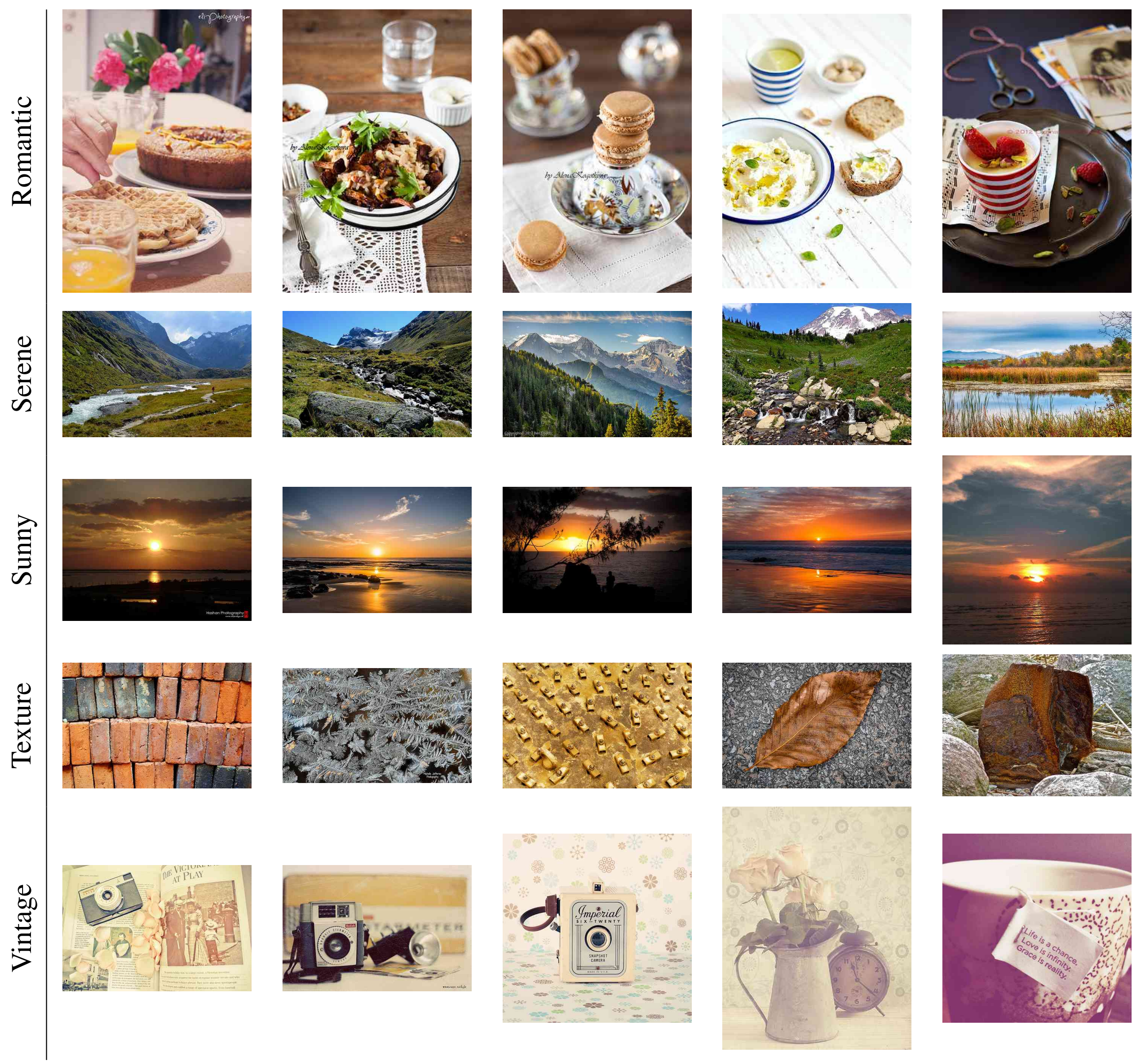

Figure 3: Top five most confident predictions on the Flickr Style test set: styles 16-20.

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]

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*	GIST	Saliency
Complementary_Colors	0.469	0.548	0.329	0.440	0.368	0.389	0.294	0.223	0.111
Duotones	0.676	0.737	0.612	0.510	0.363	0.383	0.582	0.255	0.233
HDR	0.669	0.594	0.624	0.640	0.494	0.335	0.194	0.124	0.101
Image_Grain	0.647	0.545	0.744	0.740	0.535	0.219	0.213	0.104	0.104
Light_On_White	0.908	0.915	0.802	0.730	0.805	0.508	0.867	0.704	0.172
Long_Exposure	0.453	0.431	0.420	0.430	0.208	0.242	0.232	0.159	0.147
Macro	0.478	0.427	0.413	0.500	0.376	0.438	0.230	0.269	0.161
Motion_Blur	0.478	0.467	0.458	0.400	0.327	0.186	0.117	0.114	0.122
Negative_Image	0.595	0.619	0.499	0.690	0.427	0.323	0.268	0.189	0.123
Rule_of_Thirds	0.352	0.353	0.236	0.300	0.269	0.244	0.188	0.167	0.228
Shallow_DOF	0.624	0.659	0.637	0.480	0.522	0.517	0.332	0.276	0.223
Silhouettes	0.791	0.801	0.801	0.720	0.609	0.401	0.261	0.263	0.130
Soft_Focus	0.312	0.354	0.290	0.390	0.225	0.170	0.127	0.126	0.114
Vanishing_Point	0.684	0.658	0.685	0.570	0.527	0.542	0.123	0.107	0.161
mean	0.581	0.579	0.539	0.539	0.432	0.350	0.288	0.220	0.152

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

References

[BT12] A. Bergamo and L. Torresani. Meta-class features for large-scale object categorization on a budget. In *CVPR*, 2012. 2

[DJV+13] Jeff Donahue, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, Trevor Darrell, Trevor Eecs, and Berkeley Edu. DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. Technical report, 2013. arXiv:1310.1531v1. 2

[DOB11] Sagnik Dhar, Vicente Ordonez, and Tamara L Berg. High Level Describable Attributes for Predicting Aesthetics and Interestingness. In *CVPR*, 2011. 2

[EVW+10] M Everingham, L Van Gool, C K I Williams, J Winn, and A Zisserman. The PASCAL VOC Challenge Results, 2010. 2

[HKP06] Jonathan Harel, Christof Koch, and Pietro Perona. Graph-Based Visual Saliency. In *NIPS*, 2006. 2

[Jia13] Yangqing Jia. Caffe: An open source convolutional architecture for fast feature embedding. http://caffe.berkeleyvision.org/, 2013. 2

[KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoff E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proc. NIPS*, 2012. 2

[LRF04] Siwei Lyu, Daniel Rockmore, and Hany Farid. A digital technique for art authentication. *PNAS*, 101(49), 2004. 2

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

[OT01] Aude Oliva and Antonio Torralba. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *IJCV*, 42(3):145–175, 2001. 2

[PHE12] Frank Palermo, James Hays, and Alexei A Efros. Dating Historical Color Images. In *ECCV*, 2012. 2

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 accuracy, Flickr g.t.	Our accuracy, Flickr g.t.	Our accuracy, 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 accuracy, Flickr g.t.	Our accuracy, MTurk g.t.	% change going 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 accuracy, Flickr g.t.	MTurk accuracy, Flickr g.t.	Accuracy diff. between us and MTurk
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

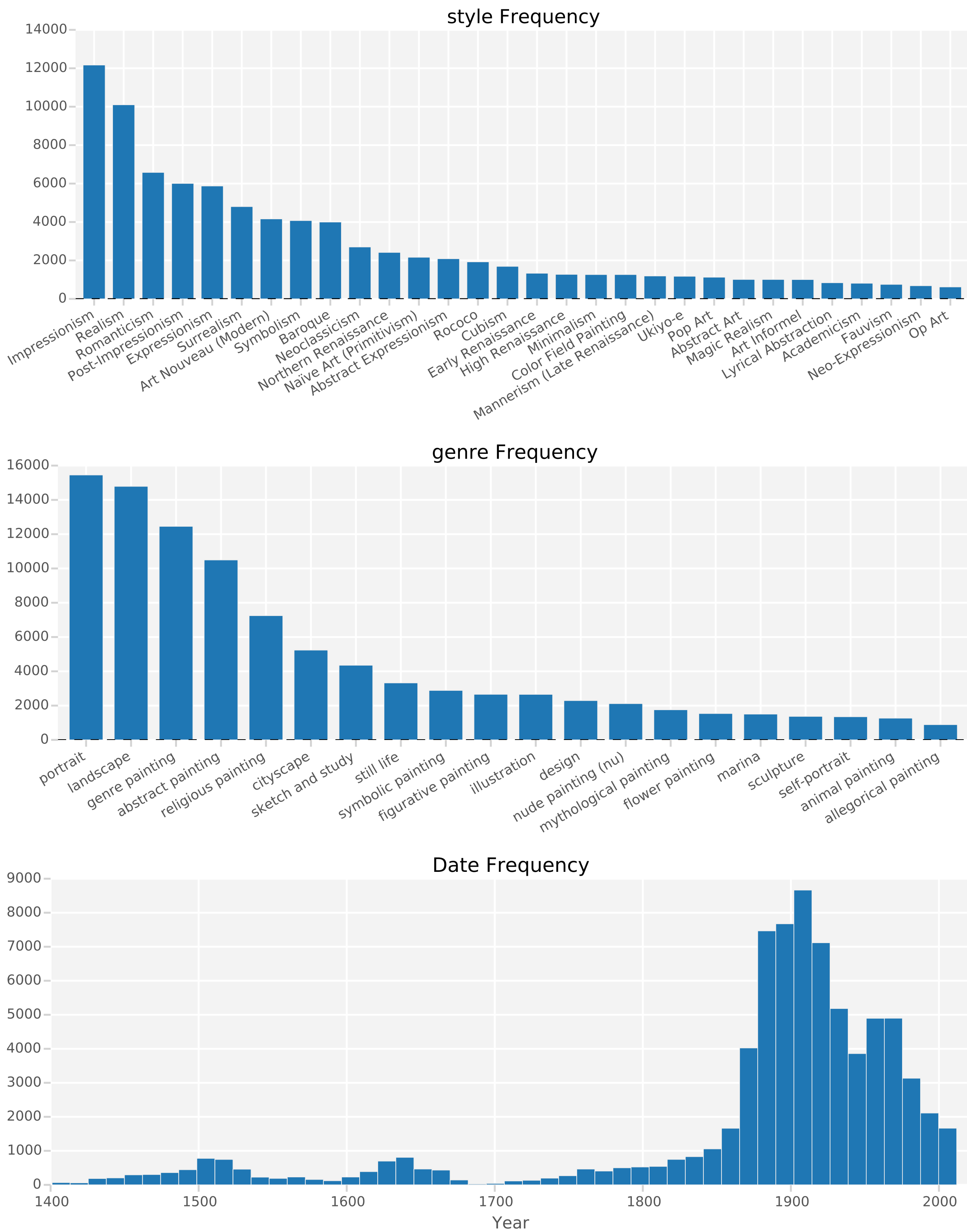

Figure 4: Distribution of image style, genre, and date in the Wikipaintings dataset.

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		

	Complementary_Colors	Duotones	HDR	Image_Grain	Light_On_White	Long_Exposure	Macro	Motion_Blur	Negative_Image	Rule_of_Thirds	Shallow_DOF	Silhouettes	Soft_Focus	Vanishing_Point	prior
Complementary_Colors	0.30	0.00	0.07	0.03	0.02	0.06	0.10	0.07	0.08	0.07	0.09	0.04	0.02	0.05	0.11
Duotones	0.00	0.24	0.00	0.19	0.07	0.02	0.03	0.07	0.07	0.05	0.06	0.05	0.02	0.13	0.21
HDR	0.03	0.03	0.52	0.01	0.00	0.07	0.01	0.04	0.03	0.11	0.00	0.01	0.01	0.11	0.07
Image_Grain	0.00	0.05	0.02	0.48	0.02	0.02	0.09	0.09	0.02	0.11	0.07	0.00	0.00	0.02	0.04
Light_On_White	0.03	0.01	0.00	0.00	0.85	0.00	0.03	0.00	0.03	0.00	0.01	0.00	0.04	0.00	0.07
Long_Exposure	0.04	0.02	0.02	0.05	0.02	0.39	0.01	0.18	0.00	0.02	0.08	0.01	0.05	0.10	0.08
Macro	0.10	0.01	0.01	0.02	0.03	0.00	0.40	0.03	0.10	0.05	0.16	0.02	0.06	0.00	0.09
Motion_Blur	0.04	0.00	0.00	0.00	0.02	0.08	0.08	0.47	0.04	0.08	0.10	0.02	0.06	0.02	0.05
Negative_Image	0.08	0.05	0.05	0.00	0.00	0.00	0.03	0.03	0.67	0.03	0.00	0.02	0.02	0.02	0.06
Rule_of_Thirds	0.03	0.00	0.07	0.01	0.03	0.03	0.03	0.05	0.04	0.29	0.20	0.09	0.08	0.07	0.07
Shallow_DOF	0.08	0.03	0.00	0.00	0.03	0.03	0.14	0.06	0.03	0.03	0.39	0.00	0.19	0.00	0.03
Silhouettes	0.02	0.02	0.00	0.00	0.02	0.09	0.00	0.02	0.00	0.02	0.00	0.72	0.03	0.07	0.05
Soft_Focus	0.00	0.07	0.00	0.00	0.03	0.03	0.03	0.10	0.00	0.03	0.07	0.03	0.59	0.00	0.03
Vanishing_Point	0.00	0.00	0.02	0.00	0.02	0.00	0.02	0.00	0.02	0.02	0.00	0.00	0.00	0.88	0.04

0.00 0.88 1.00

Figure 5: Confusion matrix of our best classifier (Late-fusion × 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.

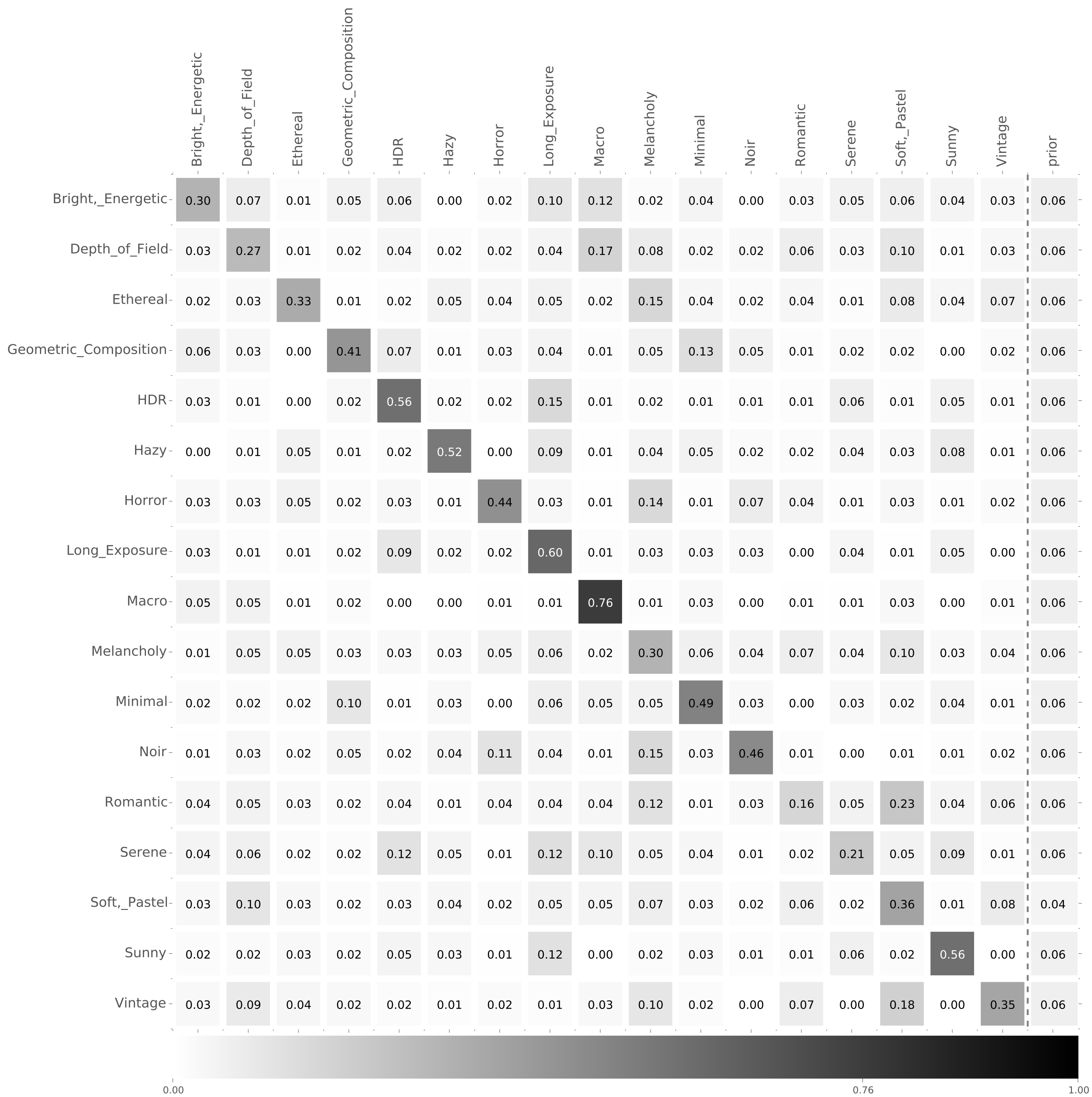

Figure 6: Confusion matrix of our best classifier (Late-fusion × Content) on the Flickr dataset.

	Abstract_Art	Abstract_Expressionism	Art_Informel	Art_Nouveau_(Modern)	Baroque	Color_Field_Painting	Cubism	Early_Renaissance	Expressionism	High_Renaissance	Impressionism	Magic_Realism	Mannerism_(Late_Renaissance)	Minimalism	Nave_Art_(Primitivism)	Neoclassicism	Northern_Renaissance	Pop_Art	Post-Impressionism	Realism	Rococo	Romanticism	Surrealism	Symbolism	Ukiyo-e	prior
Abstract_Art	0.27	0.06	0.03	0.01	0.00	0.02	0.13	0.00	0.04	0.00	0.00	0.01	0.00	0.09	0.02	0.00	0.00	0.03	0.01	0.01	0.00	0.00	0.21	0.03	0.01	0.04
Abstract_Expressionism	0.03	0.44	0.08	0.02	0.00	0.06	0.03	0.00	0.04	0.00	0.01	0.00	0.00	0.03	0.01	0.00	0.00	0.05	0.05	0.02	0.00	0.00	0.11	0.01	0.01	0.04
Art_Informel	0.03	0.25	0.16	0.02	0.00	0.02	0.04	0.00	0.04	0.00	0.02	0.03	0.00	0.04	0.03	0.00	0.00	0.04	0.04	0.03	0.00	0.00	0.18	0.03	0.00	0.04
Art_Nouveau_(Modern)	0.00	0.01	0.00	0.48	0.01	0.00	0.00	0.01	0.04	0.01	0.01	0.01	0.01	0.00	0.03	0.00	0.01	0.01	0.09	0.07	0.01	0.01	0.11	0.07	0.00	0.04
Baroque	0.00	0.00	0.00	0.02	0.49	0.00	0.00	0.02	0.01	0.03	0.01	0.00	0.05	0.00	0.00	0.03	0.03	0.00	0.01	0.12	0.06	0.09	0.03	0.01	0.00	0.04
Color_Field_Painting	0.03	0.09	0.04	0.00	0.00	0.65	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.14	0.00	0.00	0.00	0.02	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.04
Cubism	0.04	0.03	0.03	0.02	0.00	0.00	0.52	0.00	0.08	0.00	0.01	0.00	0.00	0.01	0.03	0.00	0.01	0.02	0.05	0.01	0.00	0.00	0.13	0.02	0.00	0.04
Early_Renaissance	0.00	0.00	0.01	0.03	0.01	0.00	0.01	0.52	0.01	0.08	0.01	0.00	0.05	0.00	0.00	0.01	0.09	0.00	0.02	0.03	0.01	0.02	0.07	0.02	0.00	0.04
Expressionism	0.01	0.03	0.01	0.05	0.01	0.00	0.06	0.00	0.35	0.01	0.02	0.01	0.00	0.01	0.04	0.00	0.01	0.00	0.14	0.04	0.01	0.01	0.13	0.03	0.01	0.04
High_Renaissance	0.00	0.01	0.00	0.02	0.09	0.00	0.00	0.09	0.03	0.35	0.00	0.00	0.10	0.00	0.01	0.03	0.08	0.00	0.02	0.06	0.02	0.02	0.06	0.03	0.00	0.04
Impressionism	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.00	0.03	0.00	0.54	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.13	0.13	0.01	0.02	0.01	0.06	0.00	0.04
Magic_Realism	0.00	0.03	0.01	0.05	0.00	0.00	0.01	0.00	0.04	0.00	0.01	0.39	0.01	0.02	0.02	0.03	0.01	0.01	0.06	0.06	0.01	0.02	0.16	0.04	0.01	0.04
Mannerism_(Late_Renaissance)	0.00	0.01	0.00	0.01	0.13	0.00	0.00	0.05	0.03	0.07	0.00	0.00	0.43	0.00	0.01	0.03	0.03	0.00	0.01	0.06	0.04	0.04	0.04	0.01	0.00	0.04
Minimalism	0.02	0.08	0.02	0.01	0.00	0.14	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.62	0.01	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.04
Nave_Art_(Primitivism)	0.00	0.02	0.01	0.04	0.01	0.01	0.04	0.00	0.08	0.00	0.01	0.03	0.01	0.01	0.37	0.01	0.01	0.03	0.08	0.04	0.00	0.01	0.16	0.03	0.01	0.04
Neoclassicism	0.00	0.00	0.00	0.01	0.06	0.00	0.02	0.01	0.02	0.01	0.01	0.01	0.03	0.00	0.00	0.52	0.02	0.00	0.01	0.06	0.06	0.05	0.06	0.03	0.00	0.04
Northern_Renaissance	0.00	0.01	0.00	0.03	0.05	0.00	0.01	0.06	0.03	0.05	0.00	0.02	0.03	0.00	0.01	0.01	0.52	0.01	0.03	0.05	0.00	0.02	0.07	0.01	0.00	0.04
Pop_Art	0.02	0.06	0.03	0.07	0.00	0.03	0.03	0.00	0.05	0.00	0.00	0.02	0.01	0.04	0.02	0.01	0.00	0.38	0.02	0.02	0.00	0.00	0.18	0.01	0.01	0.04
Post-Impressionism	0.01	0.02	0.01	0.02	0.00	0.00	0.01	0.00	0.09	0.00	0.16	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.48	0.07	0.01	0.00	0.05	0.04	0.00	0.04
Realism	0.00	0.01	0.00	0.03	0.03	0.00	0.00	0.01	0.03	0.01	0.08	0.01	0.01	0.00	0.01	0.03	0.02	0.00	0.06	0.47	0.00	0.09	0.06	0.05	0.00	0.04
Rococo	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.01	0.06	0.02	0.00	0.00	0.04	0.56	0.11	0.03	0.02	0.00	0.04
Romanticism	0.00	0.00	0.00	0.04	0.05	0.00	0.00	0.02	0.01	0.02	0.05	0.01	0.01	0.00	0.01	0.03	0.03	0.00	0.02	0.13	0.06	0.40	0.05	0.04	0.00	0.04
Surrealism	0.01	0.03	0.02	0.01	0.00	0.00	0.04	0.01	0.04	0.00	0.00	0.02	0.01	0.00	0.04	0.01	0.01	0.02	0.04	0.04	0.00	0.01	0.58	0.04	0.00	0.04
Symbolism	0.01	0.03	0.01	0.08	0.02	0.00	0.01	0.01	0.04	0.00	0.06	0.00	0.00	0.00	0.02	0.00	0.02	0.00	0.04	0.11	0.02	0.05	0.06	0.38	0.00	0.04
Ukiyo-e	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.00	0.01	0.05	0.02	0.81	0.04

0.00 0.81 1.00

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

Figure 8: Correlation of PASCAL content classifers (columns) against ground truth Flickr style labels (rows). Note, for example, the prevalance of vehicles in HDR and Long Exposure images, and of people in Romantic images.