

000
001
002
003
004
005
006
007
008
009
010
011012
013
014
015
016
017
018
019
020
021
022
023
024
025
026
027
028
029
030
031
032
033
034
035
036
037
038
039
040
041
042
043
044
045
046
047
048
049
050
051
052
053

What makes an object memorable?

Anonymous ICCV submission

Paper ID ****

054
055
056
057
058
059
060
061
062
063
064
065
066
067
068
069
070
071
072
073
074
075
076
077
078
079
080
081
082
083
084
085
086
087
088
089
090
091
092
093
094
095
096
097
098
099
100
101
102
103
104
105
106
107

Abstract

Recent work by Isola et. al. (2011) has demonstrated that memorability is an intrinsic property of images that is consistent across viewers and can be predicted accurately with current computer vision techniques. Despite progress, a clear understanding of the specific components of an image that drive memorability are still unknown. While previous studies such as Khosla et. al. (2012) have tried to investigate computationally the memorability of image regions within individual images, no behavioral study has systematically explored which memorability of image regions. Here we study which region from an image is memorable or forgettable. Using a large image database, we obtained the memorability scores of the different visual regions present in every image. In our task, participants viewed a series of images, each of which were displayed for 1.4 seconds. After the sequence was complete, participants similarly viewed a series of image regions and were asked to indicate whether each region was seen in the earlier sequence of full images.

1. Introduction

Consider the image and its corresponding objects in Figure 1. Even though the person on the right is comparable in size to the left person, he is remembered far less by humans (indicated by their memorability scores of 0.18 and 0.64 respectively). People tend to remember the fish in the center and the person on the left, even after 30 minutes have passed (memorability score = 0.64). Interestingly, despite vibrant colors and considerable size, the boat is also remembered far less by humans (memorability = 0.18).

NOTE: Big picture way to think about this (rough): A currently massive goal in computer science and artificial intelligence is to model the data driven inference processes that humans make use of for solving vision problems, such as object recognition and scene understanding. A subset of these processes make use of nearly all of the information in the scene, much of which is used implicitly but others are dependent on more explicitly remembered information in



Figure 1: Not all objects are equally remembered. Memorability scores of objects for the image in the top row obtained from our psychophysics experiment.

a scene, that is, the information that the brain has deemed worthy to create special storage for. For example, the recognition of edges in a scene is learned and used automatically, but if we ask a person to report the important information in a scene, only a few high level things are reported. More specifically, humans can only make judgements about elements of an image that they can remember...

Just like aesthetics, interestingness, and other metrics of image importance, memorability quantifies something about the utility of a photograph toward our everyday lives. For many practical tasks, memorability is an especially desirable property to maximize. For example, this may be the case when creating educational materials, logos, advertisements, book covers, websites, and much more. Understanding memorability, and being able to automatically predict it, lends itself to a wide variety of applications in each of these areas. **rewrite this to draw attention of reviewer to importance of image memorability.** Due to this, automatic prediction of intrinsic memorability of images using computer vision and machine learning techniques has received considerable attention in the recent years [5], [7], [4], [2], [8]. While these studies have shed light on what distinguishes the memorability of different images and the intrinsic and extrinsic properties that make those images memorable, the above example raises an interesting question: what exactly about an image is remembered? Despite progress in the computer vision literature on image memorability, a clear understanding of the memorability of the specific components of an image is still unknown. For example, not all objects in an image will be equally remembered by people and as the figure 1 seems to suggest, there exists significant

108 and interesting differences in memorability of objects in an
109 image. Furthermore, the memorability of complex images
110 may be principally driven by the memorability of its objects.
111 Can specific objects inside images be memorable to
112 all us and how can we better understand what makes those
113 objects more memorable?
114

115 In this paper, we systematically explore the memorability
116 of objects within individual images and shed light
117 on the various factors and properties that drive object
118 memorability by augmenting both the images and object
119 segmentations in the 850 existing images from PASCAL
120 2010 [3] dataset with memorability scores and class labels.
121 By exploring the connection between object memorability,
122 saliency, and image memorability, our paper makes several
123 important contributions.
124

125 Firstly, we show that just like image memorability, object
126 memorability is a property that is shared across subjects and objects remembered by one person are also likely
127 to be remembered by others and vice versa. Secondly, we
128 show that there exists a strong correlation between visual
129 saliency and object memorability and demonstrate insights
130 when can visual saliency directly predict object memorability
131 and when does it fail to do so. While there have been
132 have a few studies that explore the connection between image
133 memorability and visual saliency [2], [13], our work is
134 the first to explore the connection between object memorability
135 and visual saliency. Third, we explore the connection
136 between image memorability and object memorability
137 and show that the most memorable object inside an image
138 can be a strong predictor of image memorability in certain
139 cases. Studying these questions, help not only understand
140 visual saliency, image and object memorability in more detail,
141 but it can also have important contributions to computer
142 vision. For example, understanding which regions and objects
143 in an image are memorable would enable us to modify
144 the memorability of images which can have applications in
145 advertising, user interface design etc. With this in mind,
146 as shown in the section 4, our proposed dataset serves as
147 a benchmark for evaluating object memorability model
148 algorithms and can help usher in future algorithms that try to
149 predict memorability maps.
150

1.1. Related works

151 In this section, we briefly discuss existing work related
152 to visual memory and image memorability. We also review
153 research related to visual saliency prediction and discuss the
154 relationship of memorability and visual attention.
155

156 **Image Memorability:** Describe Isola's first paper n
157 some insights that have been raised on image memorability
158 thus far. Also describe Khosla's comp model but we are
159 the first work to actually describe what humans actually
160 remember and don't
161

Visual Saliency: Talk about visual attention and models

162 that have been proposed. Also, talk about Pascal-S and how
163 it has helped reduce dataset bias
164

165 **Saliency and memorability:** discuss some results related
166 to saliency and image memorability.
167

168 and talk about our work plans on connecting and shedding
169 light on all these phenomena together.
170

2. Measuring Object Memorability

171 As a first step towards understanding memorability of
172 objects, we built an image database containing a variety of
173 objects from a diverse range of categories, and measured the
174 probability that every object in each image will be remembered
175 by a large group of subjects after a single viewing.
176 This helps provide ground truth memorability scores for the
177 objects inside the images and allows for a precise analysis
178 of the memorable elements within an image. For this task,
179 we utilized the PASCAL-S dataset [10], a fully segmented
180 dataset built on the validation set of the PASCAL VOC
181 2010 [3] segmentation challenge. For improved segmentation
182 purposes, we manually cleaned up and refined the segmentations
183 from this dataset. We removed all homogenous
184 non-object or background segments such as ground, grass,
185 floor, sky etc, as well as imperceptible object fragments and
186 excessively blurred regions. All remaining object segmentations
187 were tested for memorability. In the end, our final
188 dataset consisted of 850 images and 3412 object segmentations
189 i.e. on average each image consisted of approximately
190 4 object segments for which we gathered the ground truth
191 memorability on.
192

2.1. Object Memory Game

193 To measure the memorability of individual objects from
194 our dataset, we created an alternate version of the Visual
195 Memory Game through Amazon Mechanical Turk following
196 the basic design in [5], with the exception of a few key
197 differences. In our game, participants first viewed a
198 sequence of images one at a time, with a 1.5 second gap in
199 between image presentations. Subjects were asked to remember
200 the contents and objects inside those images as much as
201 they could. To ensure that subjects would not just only look
202 at the salient or center objects, subjects had unlimited time
203 to freely view the images. Once they were done viewing an
204 image, they could press any key to advance to the next image.
205 Following the initial image sequence, participants then
206 viewed a sequence of objects, their task then being to indicate
207 through a key press which of those objects was present
208 in one of the previously shown images. Each object was
209 displayed for 1.5 second, with a 1.5 second gap in between
210 the object sequences. Pairs of corresponding image and object
211 sequences were broken up into 10 blocks. Each block
212 consisted of 80 total stimuli (35 images and 45 objects), and
213 lasted approximately 3 minutes. At the end of each block,
214

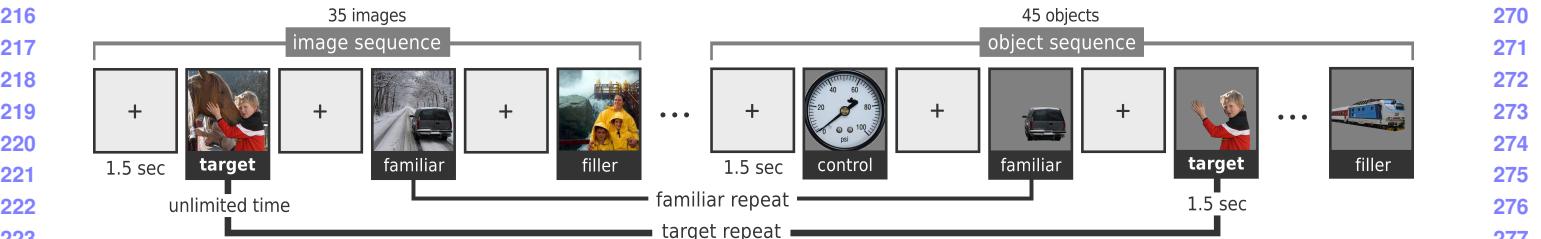


Figure 2: Main task. add-in later.

216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
the subject could take a short break. Overall, the experiment took approximately took 30 minutes to complete.

231 Unknown to the subjects, inside each block, each sequence of images was pseudo-random and consisted of 3 'target' images taken from the Pascal-S dataset whose objects the participants were to later identify. The remaining images in the sequence consisted of 16 'filler' images and 16 'familiar' images. The 'filler' images were randomly selected from the DUT-OMRON dataset [14] and the 'familiar' images were randomly sampled from the MSRA dataset proposed in [12]. Similarly, the object sequence was also pseudo-random and consisted of 3 'target' objects (1 object taken randomly from each previously shown target image). The remaining objects in the sequence consisted of 10 'control' objects, 16 'filler' objects, and 16 'familiar' objects. The 'filler' objects were taken randomly from the 80 different object categories in the Microsoft COCO dataset [11] and the 'familiar' objects were the objects taken from the previously displayed 'familiar' images in the image sequence. The fillers and familiars helped provide spacing between the target images and target objects, whereas the control objects allowed us to check if the subjects were paying attention to the task [11], [5]. While the fillers and familiars (both the images and objects) were taken from datasets resembling real world scenes and objects, the 'control' objects were artificial stimuli randomly sampled from the dataset proposed in [1] and helped serve as a control to test the attentiveness of the subjects. The target images and the respective target objects were spaced 70 – 79 stimuli apart, and familiar images and their respective objects were spaced 1 – 79 stimuli apart. All images and objects appeared only once, and each subject was tested on only one object from each target image. Objects were centered within their parent frame and non-object pixels were set to grey. Participants were required to complete the entire task, which included 10 blocks (overall time approximately 30 minutes), and could not participate in the experiment a second time. After collecting the data, we assigned a 'memorability score' to each target object in our dataset, defined as the percentage of correct detections by subjects. In all our analysis, we removed all subjects whose accuracy on the control objects was below 70%. In the end, our analysis was performed on a total of 1823 workers from Mechanical

270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
Turk ($> 95\%$ approval rate in Amazons system). The memorability score of an object corresponded to the number of subjects that correctly detected the repetition of that object. On average, each object was scored by 16 subjects and the average memorability score was 33% ($SD = 28\%$).

2.2. Consistency Analysis

298 To assess human consistency in remembering objects, 299 we repeatedly divided our entire subject pool into two equal 300 halves and quantified the degree to which memorability 301 scores for the two sets of subjects were in agreement using 302 Spearmans rank correlation (ρ). We computed the average 303 correlation over 25 of these random split iterations, yielding 304 a final value of 0.76. This high consistency in object 305 memorability indicates that, like full images, object memorability 306 is a shared property across subjects. People tend to 307 remember (and forget) the same objects in images, and 308 exhibit similar performance in doing so. Thus memorability 309 of objects in images can potentially be predicted with high 310 accuracy. In the next section, we study the various factors 311 that possibly drive object memorability in images.

3. Understanding Object Memorability

303 In this section, we aim to better understand object memorability 304 and the factors that make an object more memorable 305 or forgettable to humans. We first investigate the role that 306 simple color features play in determining object memorability. 307

3.1. Can simple features explain memorability?

312 While simple image features are traditionally poor predictors 313 of memorability in full images [5], and with good 314 reason [9], do they play any role in determining object 315 memorability? We decomposed each image into its hue, 316 saturation, and value components and calculated the mean 317 and standard deviation of each channel. Mean value ($\rho = 318 0.1$) and variance in value ($\rho = 0.25$) were weakly correlated 319 with object memorability suggesting that brighter and 320 higher contrast objects may be more memorable (Figure 3). 321 On the other hand, essentially no relationship was found 322 between memorability and either hue or saturation (Figure 3). 323 This deviates slightly from the findings in [5] that

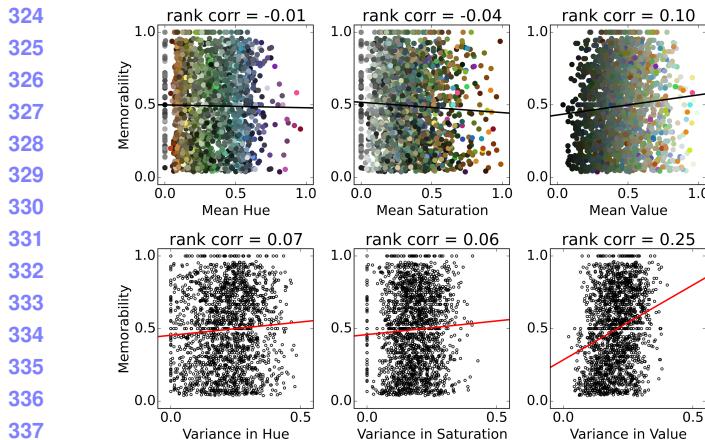


Figure 3: **Correlations between simple color features and object memorability.** Color features were computed from HSV representations of the objects.

showed mean hue to be weakly predictive of image memorability. However, this makes sense since the effect was speculated to be due to the blue and green outdoor landscapes being less memorable than warmly colored human faces and indoor scenes. While our dataset contained plenty of indoor objects and people, outdoor scene-related image regions such as sky and ground were not included as objects. Taken together, these results show that, like image memorability, basic pixel statistics do not play a significant role in determining the memorability of objects in images.

3.2. What is the role of saliency in memorability?

Intuitively, the regions within an image that are most salient are likely to have a higher probability of being remembered, since they will draw the attention of viewers and a majority of a viewer's eye fixations will be spent looking at those regions. On the other hand, it is conceivable that some visually appealing regions will not be memorable, especially since aesthetic images are known to be less memorable [5], [4]. When can visual saliency predict object memorability and what are the possible differences between these two phenomena? Quantifying the precise relationship between saliency and memorability will be paramount towards understanding object memorability in greater depth.

To this aim, we utilized the eye fixation dataset made available for the Pascal-S dataset in [10]. With this dataset in hand, we first calculated the number of unique fixation points within the area of each object and computed the correlation between this metric and the object's memorability score (Figure 5 a). We found this correlation to be positive and considerably high ($\rho = 0.71$), suggesting that fixation count and visual saliency may drive object memorability considerably. However, the large concentration of points on the bottom left part of scatter plot in Figure 5 a suggests that part of the reason for this high correlation is that ob-

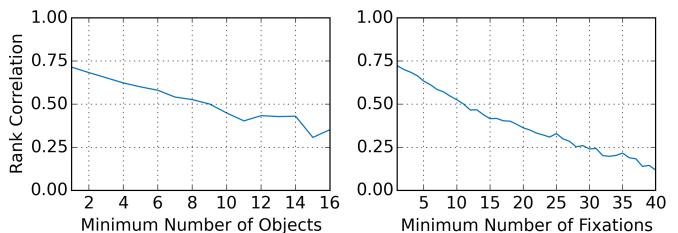


Figure 4: **Correlation between object memorability and number of fixations.** add-in later.

jects that have not been viewed at all have essentially no memorability. Indeed, only objects that have been seen can be remembered. In addition, the points toward the top left appear to decrease in trend. Looking deeper, Figure 4 plots the change in correlation between object memorability and fixations as the minimum number of fixations inside objects increases. The downward monotonic trend indicates that as the number of fixations inside an object increases, the predictive ability diminishes significantly. In addition, Figure 4 plots the correlation between object memorability and number of fixations as a function of total number of objects in an image. Similar to the previous trend, as the number of objects in an image increases, the correlation between saliency i.e. number of fixations and memorability score decreases sharply. This finding is in agreement with the two remaining scatter plots in Figure 5 b (shows that the memorability of an object decreases in the presence of many other objects) and Figure 5 c (shows that number of fixations decreases with the number of objects). This makes intuitive sense since people have more to look at in an image when more objects are present, and so they may look less at any one object, especially if they compete for saliency, and therefore may have a more difficult time remembering those objects.

To sum up, saliency is a surprisingly good index of object memorability in simple contexts where there are few objects in the image, or when an object has few interesting points, but it is a much weaker predictor of object memorability in complex scenes containing multiple objects that have many points of interest (Figure 7).

Center Bias: Figure 6 elucidates another example

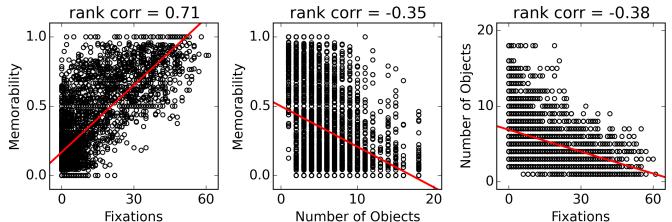
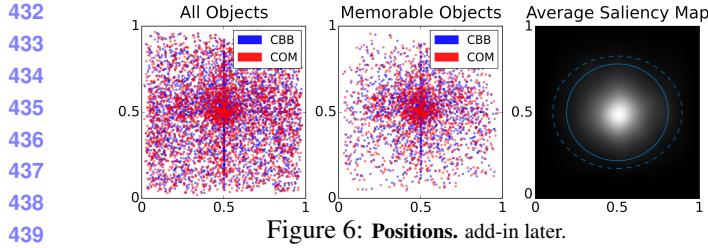
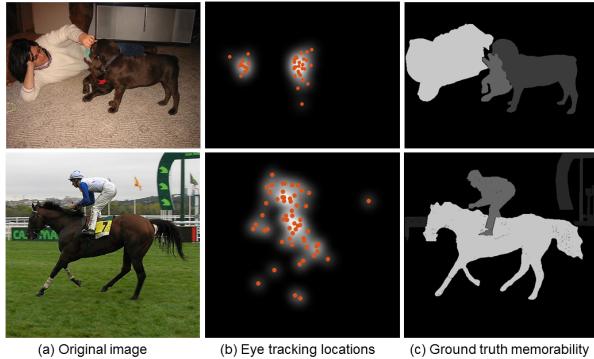


Figure 5: **Correlation between object memorability and number of fixations.** add-in later.

324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431

Figure 6: **Positions.** add-in later.Figure 7: **Saliency Fail cases.** add-in later.

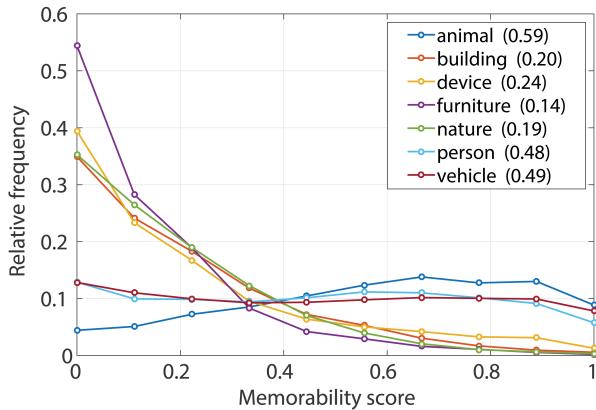
where saliency and memorability diverge. While previous studies related to visual saliency have showed that saliency is heavily influenced by center bias [6], [15], Figure 6 shows that memorability exhibits comparatively less center bias than saliency. This is most apparent when considering the difference in the solid ellipse in the right plot (shows where 95% of fixation positions are located), and the dashed ellipse (shows where the 95% of the above-median memorable objects are located).

3.3. How do object categories affect memorability?

In the previous sections, we showed that simple features have little predictive power over object memorability and explored the relationship between visual saliency and object memorability. In this section, we explore how the category of an object influences the probability that the object will be remembered.

3.3.1 Are some classes more memorable than others?

For this analysis, we first assigned three in-house annotators the task of assigning class labels to each object segmentation in our dataset. The annotators were given the original image (for reference) and the object segmentation and asked to assign a single category to the segment out of 7 possible categories: animal, building, device, furniture, nature, person, and vehicle. We choose these high-level categories such that a wide range of object classes could be covered under these categories. For example, device included object segments such as utensils, bottles, televisions, computers etc, nature included segments like trees, mountains, flow-

Figure 8: **Average memorability per object class.** Figure showing some object classes are more memorable than others.

ers, and vehicle contained segments like cars, bikes, buses, airplanes etc.

Figure 8 shows the distribution of the memorability scores for all 7 object classes in our dataset. This visualisation gives a sense of how the memorability changes across different object categories. Animal, person, and vehicle are all highly memorable classes each associated with an average memorability score greater than or close to 0.5. Interestingly, all other object categories have an average memorability score lower than 0.25, indicating that humans do not remember objects from these categories very well. In particular, furniture is the least memorable object class with an average memorability score of only 0.14. This could be possibly due to the fact that most objects from classes like furniture, nature, and building either appear mostly in the background or are occluded which likely decreases their memorability significantly. By contrast, objects from the animal, person, and vehicle classes appear mostly in the foreground, leading to a higher memorability score on average. Interestingly, the topmost memorable objects from building, furniture, and nature tend to have an average memorability score in the range of 0.4 – 0.8, whereas the topmost memorable objects from classes person, animal and vehicle have an average memorability higher than 0.90. This is particularly interesting as these top objects are not occluded and most of them tend to appear in the foreground. While the differences in the memorability of different classes could be driven primarily due to factors like occlusion, size, background/foreground, or photographic bias, the distribution in figure 8 suggests that humans remember some object classes such as person, animal, and vehicle irrespective of external nuisance factors and these object classes are *intrinsically* more memorable than others.

486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539

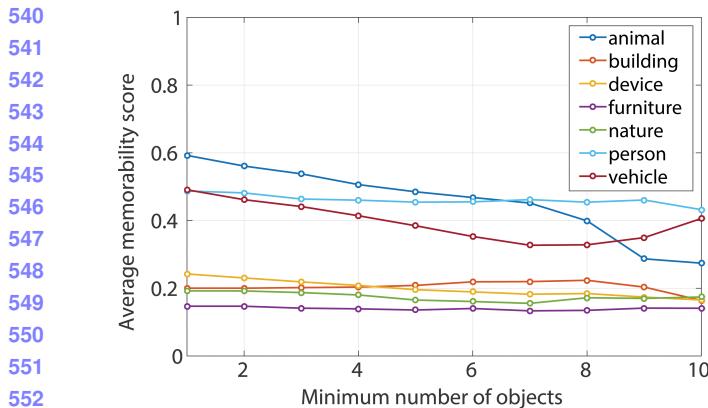


Figure 9: Correlation between object class and number of objects. add-in later.

3.3.2 Why are some objects not memorable in a class?

As demonstrated above, some object classes (i.e. animal, person, vehicle) are more memorable than others. However, not all objects in a class are equally memorable. The examples in figure ?? show the most memorable and least memorable objects for each object class. Across classes, non-memorable objects tend to be those that are occluded and obstructed by other objects. What other possible factors could influence the memorability of an object within a class? Among the various possible factors, we explored how object memorability within a particular class is influenced by a) the number of objects in an image and b) the presence of other object classes.

Number of objects: We first examined how the memorability of each object class is affected by the number of objects inside an image. Figure 9 shows the change in average memorability of the different object classes with respect to the increase in number of objects in an image. Results indicate that the number of objects present in an image is an important factor in determining memorability. For example, as the number of objects in an image increases, the memorability of animals and vehicles decreases sharply, likely as a result of competition for attention, or decreased spotlight on a single subject of the composition. Interestingly, the memorability of the person class does not change significantly with an increase in number of objects. This suggests that people are not only one of the most memorable object classes, but are also more robust to the presence of clutter in images. This may be because single people in images steal all of the attention of the viewer, but how do they behave in the presence of other people? To answer this, we turn to the question of interclass memorability next.

Inter-class memorability: How does the presence of a particular object class influence the memorability of another object class? For all pairwise combinations of object category, we gathered all images that contained at least

one object from both categories and computed the change in the average memorability scores for the two object categories. Figure 10 plots these data and visualizes how the memorability of each object class is affected by the presence of other object classes. The first thing to note is that the values of low-memorability classes (i.e. nature, furniture, device, and building) are not greatly affected by the presence of other object categories. Instead, their memorability tends to remain low across all contexts. The memorability of the animal class remains close to its high average memorability score in presence of most classes, but drops significantly in the presence of other animals, vehicles, and people. The memorability of people also remains close to its average memorability score and tends to be unaffected by the presence of most object categories (including other people). However, the memorability of a person decreases in the presence of vehicles and buildings. This could be due to the fact that people in images containing vehicles or buildings are usually zoomed out and are usually smaller in size (also illustrated in figure 11). The memorability of the vehicle class is strongly affected by the presence of other object categories. In particular, its memorability drops significantly in the presence of another vehicle, people, and animals. Taken together, when an animal, vehicle or a person is present in the same image, the memorability of all three classes usually goes down. However, this pattern of change in memorability varies by class, leading to interesting results. For example, when a vehicle and animal are present in the same image, the animal is generally more memorable, even though the memorability of both of these classes drops significantly. When a vehicle or an animal co-occurs with a person, the person is generally more memorable (also shown in Figure 11).

3.4. How are object & image memorability related?

We now know what objects people remember and the factors that influence their memorability, but to what ex-

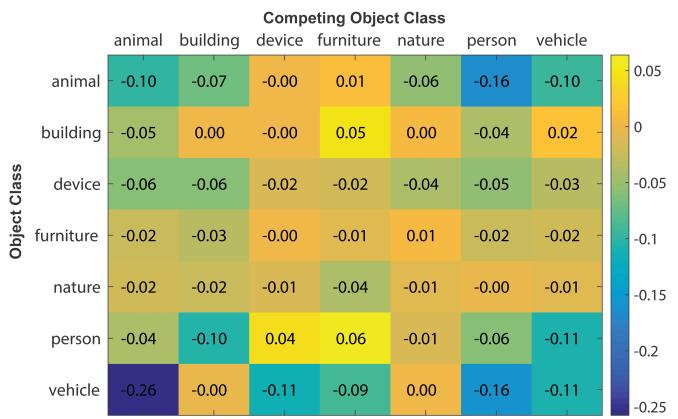


Figure 10: inter-class object memorability relationship. add-in later.

594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647



Figure 11: **Qual Results.** Figure showing how memorability of different classes is effected in presence of other classes. Bottom row is the memorability map

tent does the memorability of individual objects affect the overall memorability of a scene? Moreover, if an image is highly memorable, what can we say about the memorability of the objects inside those images (and vice versa)? To shed light on this question, we conducted a second large-scale experiment on Amazon Mechanical Turk for all images in our dataset to gather their respective image memorability scores. For this experiment, we followed the exact paradigm as the memory game experiment proposed in [5]. A series of images from our dataset and Microsoft COCO dataset [11] (i.e. the 'filler' images) were flashed for 1 second each, and subjects were instructed to press a key whenever they detected a repeat presentation of an image. A total of 350 workers participated in this experiment with each image being viewed 80 times on average. The rank correlation after averaging over 25 random split half trials was found to be 0.70, providing evidence for consistency in the image memorability scores.

Utilizing results from both experiments, we computed the correlation between the the scores of the single most memorable object in each image (from Experiment 1) and the overall memorability score of each image (from Experiment 2). We found this correlation to be moderately high ($\rho = 0.4$), suggesting that the most highly memorable object in the image plays a crucial role in determining the overall memorability of an image. To investigate this finding in relation to extreme cases only, we performed the same analysis as above on a subset of the data containing the topmost 100 memorable images and the bottommost 100 memorable images. The correlation between maximum object memorability and image memorability for this subset of the images increased significantly ($\rho = 0.62$), meaning maximum object memorability serves as a strong indicator of whether an image is *highly* memorable or non-memorable. That is, images that are highly memorable contain at least one highly memorable object, and images with low memorability usually do not contain a single highly memorable object (also shown in Figure 12).

It seems that maximum object memorability is highly explanatory, but does this behavior generalize across object classes? We further computed the correlation between maximum object and image memorability for each individ-

Animal	Building	Device	Furniture	Nature	Person	Vehicle	All
0.38	0.22	0.47	0.53	0.64	0.54	0.30	0.40

Table 1: **Max object memorability and image memorability.** add-in later.

ual object class. Results shown in Table 1 show that certain object classes are more strongly correlated than others. For example, images containing animals, buildings, or vehicles as the most memorable objects tend to have varying degree of image memorability (indicated by their lower ρ values). On the other hand, classes like device, furniture, nature, and person are strongly correlated with image memorability, indicating that if an image's most memorable object belongs to one of these classes, the object memorability score is strongly predictive of the image memorability score. We can imagine scenarios in which this information would be potentially useful. For example, in the case where vision systems are tasked to predict scene memorability, a *single* object and its class can serve as a strong prior in predicting image memorability.

4. Predicting Object Memorability

here the benchmarking and baseline code comes up

References

- [1] T. F. Brady, T. Konkle, G. A. Alvarez, and A. Oliva. Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences*, 105(38):14325–14329, 2008. 3
- [2] Z. Bylinskii, P. Isola, C. Bainbridge, A. Torralba, and A. Oliva. Intrinsic and extrinsic effects on image memorability. *Vision research*, 2015. 1, 2
- [3] M. Everingham and J. Winn. The pascal visual object classes challenge 2010 (voc2010) development kit, 2010. 2
- [4] P. Isola, J. Xiao, D. Parikh, A. Torralba, and A. Oliva. What makes a photograph memorable? *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(7):1469–1482, 2014. 1, 4
- [5] P. Isola, J. Xiao, A. Torralba, and A. Oliva. What makes an image memorable? In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 145–152. IEEE, 2011. 1, 2, 3, 4, 7
- [6] T. Judd, K. Ehinger, F. Durand, and A. Torralba. Learning to predict where humans look. In *Computer Vision, 2009 IEEE 12th international conference on*, pages 2106–2113. IEEE, 2009. 5
- [7] A. Khosla, J. Xiao, A. Torralba, and A. Oliva. Memorability of image regions. In *Advances in Neural Information Processing Systems*, pages 305–313, 2012. 1
- [8] J. Kim, S. Yoon, and V. Pavlovic. Relative spatial features for image memorability. In *Proceedings of the 21st ACM international conference on Multimedia*, pages 761–764. ACM, 2013. 1

- 756 [9] T. Konkle, T. F. Brady, G. A. Alvarez, and A. Oliva. Conceptual
757 distinctiveness supports detailed visual long-term memory for real-world objects. *Journal of Experimental Psychology: General*, 139(3):558, 2010. 3 810
758 811
759 812
760 [10] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille. The
761 secrets of salient object segmentation. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*,
762 pages 280–287. IEEE, 2014. 2, 4 813
763 814
764 [11] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common
765 objects in context. In *Computer Vision–ECCV 2014*, pages 740–755. Springer, 2014. 3, 7 815
766 816
767 817
768 [12] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and
769 H.-Y. Shum. Learning to detect a salient object. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*,
770 33(2):353–367, 2011. 3 818
771 819
772 [13] M. Mancas and O. Le Meur. Memorability of natural scenes:
773 the role of attention. In *ICIP*, 2013. 2 820
774 821
775 [14] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang.
776 Saliency detection via graph-based manifold ranking. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3166–3173. IEEE, 2013. 3 822
777 823
778 [15] L. Zhang, M. H. Tong, T. K. Marks, H. Shan, and G. W. Cottrell.
779 Sun: A bayesian framework for saliency using natural
780 statistics. *Journal of vision*, 8(7):32, 2008. 5 824
781 825
782 826
783 827
784 828
785 829
786 830
787 831
788 832
789 833
790 834
791 835
792 836
793 837
794 838
795 839
796 840
797 841
798 842
799 843
800 844
801 845
802 846
803 847
804 848
805 849
806 850
807 851
808 852
809 853
810 854
811 855
812 856
813 857
814 858
815 859
816 860
817 861
818 862
819 863

864												918	
865												919	
866												920	
867												921	
868												922	
869												923	
870												924	
871												925	
872												926	
873												927	
874												928	
875												929	
876												930	
877												931	
878												932	
879												933	
880												934	
881												935	
882												936	
883												937	
884												938	
885												939	
886	0.94	1.00	0.93	0.75	0.92	0.80	0.92	0.94	0.92	0.80	0.91	0.76	940
887													941
888													942
889													943
890													944
891	0.47	0.56	0.48	0.38	0.50	0.38	0.50	0.61	0.51	0.50	0.52	0.39	945
892													946
893													947
894													948
895													949
896													950
897													951
898													952
899													953
900													954
901													955
902													956
903													957
904													958
905													959
906													960
907													961
908													962
909													963
910													964
911													965
912													966
913													967
914													968
915													969
916													970
917													971

Figure 12: Qual image-object results. add-in later.