

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], [6], [4], [2], [7]. 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 In this paper, we systematically explore the memorability
115 of objects within individual images and shed light
116 on the various factors and properties that drive object
117 memorability by augmenting both the images and object
118 segmentations in the 850 existing images from PASCAL
119 2010 [3] dataset with memorability scores and class labels.
120 By exploring the connection between object memorability,
121 saliency, and image memorability, our paper makes several
122 important contributions.

123 Firstly, we show that just like image memorability, object
124 memorability is a property that is shared across subjects
125 and objects remembered by one person are also likely
126 to be remembered by others and vice versa. Secondly, we
127 show that there exists a strong correlation between visual
128 saliency and object memorability and demonstrate insights
129 when can visual saliency directly predict object memorability
130 and when does it fail to do so. While there have been
131 have a few studies that explore the connection between image
132 memorability and visual saliency [2], [11], our work is
133 the first to explore the connection between object memorability
134 and visual saliency. Third, we explore the connection
135 between image memorability and object memorability
136 and show that the most memorable object inside an image
137 can be a strong predictor of image memorability in certain
138 cases. Studying these questions, help not only understand
139 visual saliency, image and object memorability in more detail,
140 but it can also have important contributions to computer
141 vision. For example, understanding which regions and objects
142 in an image are memorable would enable us to modify
143 the memorability of images which can have applications in
144 advertising, user interface design etc. With this in mind,
145 as shown in the section 4, our proposed dataset serves as
146 a benchmark for evaluating object memorability model
147 algorithms and can help usher in future algorithms that try to
148 predict memorability maps.

149 1.1. Related works

150 **Image Memorability:** Describe Isola's first paper n
151 some insights that have been raised on image memorability
152 thus far. Also describe Khosla's comp model but we are
153 the first work to actually describe what humans actually
154 remember and don't

155 **Visual Saliency:** Talk about visual attention and models
156 that have been proposed. Also, talk about Pascal-S and how
157 it has helped reduce dataset bias

158 **Saliency and memorability:** discuss some results related
159 to saliency and image memorability.

160 and talk about our work plans on connecting and shedding
161 light on all these phenomena together.

162 2. Measuring Object Memorability

163 As a first step towards understanding memorability of
164 objects, we built an image database containing a variety of
165 objects from a diverse range of categories, and measured the
166 probability that every object in each image will be remembered
167 by a large group of subjects after a single viewing.
168 This helps provide ground truth memorability scores for the
169 objects inside the images and allows for a precise analysis
170 of the memorable elements within an image. For this task,
171 we utilized the PASCAL-S dataset [8], a fully segmented
172 dataset built on the validation set of the PASCAL VOC
173 2010 [3] segmentation challenge. For improved segmentation
174 purposes, we manually cleaned up and refined the segmentations
175 from this dataset. We removed all homogenous
176 non-object or background segments such as ground, grass,
177 floor, sky etc, as well as imperceptible object fragments and
178 excessively blurred regions. All remaining object segmentations
179 were tested for memorability. In the end, our final
180 dataset consisted of 850 images and 3412 object segmentations
181 i.e. on average each image consisted of approximately
182 4 object segments for which we gathered the ground truth
183 memorability on.

184 2.1. Memory Game: Measuring Object Memorability

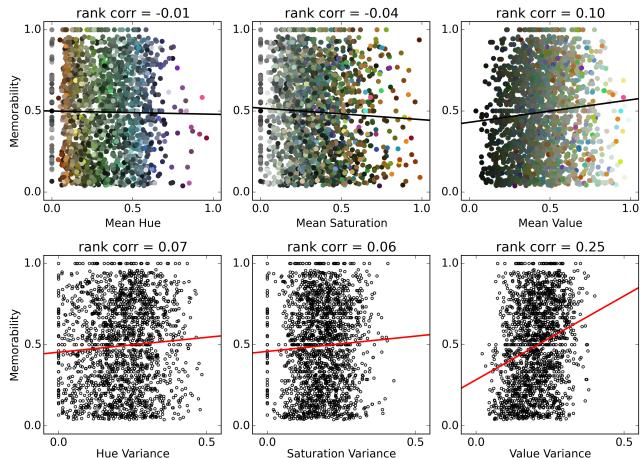
185 To measure the memorability of individual objects from
186 our dataset, we created an alternate version of the Visual
187 Memory Game through Amazon Mechanical Turk following
188 the basic design in [5], with the exception of a few key
189 differences. In our game, participants first viewed a
190 sequence of images one at a time, with a 1.5 second gap in
191 between image presentations. Subjects were asked to remember
192 the contents and objects inside those images as much as
193 they could. To ensure that subjects would not just only look
194 at the salient or center objects, subjects had unlimited time
195 to freely view the images. Once they were done viewing an
196 image, they could press any key to advance to the next image.
197 Following the initial image sequence, participants then
198 viewed a sequence of objects, their task then being to indicate
199 through a key press which of those objects was present
200 in one of the previously shown images. Each object was
201 displayed for 1.5 second, with a 1.5 second gap in between
202 the object sequences. Pairs of corresponding image and object
203 sequences were broken up into 10 blocks. Each block
204 consisted of 80 total stimuli (35 images and 45 objects), and
205 lasted approximately 3 minutes. At the end of each block,
206 the subject could take a short break. Overall, the experiment
207 took approximately 30 minutes to complete.

208 Unknown to the subjects, inside each block, each sequence
209 of images was pseudo-random and consisted of 3

216 'target' images taken from the Pascal-S dataset whose
 217 objects the participants were to later identify. The remaining
 218 images in the sequence consisted of 16 'filler' images and
 219 16 'familiar' images. The 'filler' images were randomly se-
 220 lected from the DUT-OMRON dataset [12] and the 'familiar'
 221 images were randomly sampled from the MSRA dataset
 222 proposed in [10]. Similarly, the object sequence was also
 223 pseudo-random and consisted of 3 'target' objects (1 object
 224 taken randomly from each previously shown target image).
 225 The remaining objects in the sequence consisted of 10 'con-
 226 trol' objects, 16 'filler' objects, and 16 'familiar' objects.
 227 The 'filler' objects were taken randomly from the 80 differ-
 228 ent object categories in the Microsoft COCO dataset [9] and
 229 the 'familiar' objects were the objects taken from the pre-
 230 viously displayed 'familiar' images in the image sequence.
 231 The fillers and familiars helped provide spacing between the
 232 target images and target objects, whereas the control objects
 233 allowed us to check if the subjects were paying attention to
 234 the task [1], [5]. While the fillers and familiars (both the
 235 images and objects) were taken from datasets resembling
 236 real world scenes and objects, the 'control' objects were ar-
 237 tificial stimuli randomly sampled from the dataset proposed
 238 in [1] and helped serve as a control to test the attentiveness
 239 of the subjects. The target images and the respective tar-
 240 get objects were spaced 70 – 79 stimuli apart, and familiar
 241 images and their respective objects were spaced 1 – 79 stim-
 242uli apart. All images and objects appeared only once, and
 243 each subject was tested on only one object from each tar-
 244 get image. Objects were centered within their parent frame
 245 and non-object pixels were set to grey. Participants were re-
 246 quired to complete the entire task, which included 10 blocks
 247 (overall time approximately 30 minutes), and could not par-
 248 ticipate in the experiment a second time. After collecting
 249 the data, we assigned a 'memorability score' to each target
 250 object in our dataset, defined as the percentage of correct
 251 detections by subjects. In all our analysis, we removed all
 252 subjects whose accuracy on the control objects was below
 253 70%. In the end, our analysis was performed on a total of
 254 1823 workers from Mechanical Turk ($> 95\%$ approval rate
 255 in Amazons system) and on average each object was scored
 256 by 16 subjects.
 257

2.2. Consistency Analysis

260 To assess human consistency in remembering objects,
 261 we repeatedly divided our entire subject pool into two equal
 262 halves and quantified the degree to which memorability
 263 scores for the two sets of subjects were in agreement using
 264 Spearmans rank correlation (ρ). We computed the average
 265 correlation over 25 of these random split iterations, yield-
 266 ing a final value of 0.76. This high consistency in object
 267 memorability indicates that, like full images, object memo-
 268 rability is a shared property across subjects. People tend
 269 to remember the same objects in images, and exhibit simi-



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
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323

Figure 2: Correlations between simple color features and object memorability. Color features were computed from HSV representations of the objects.

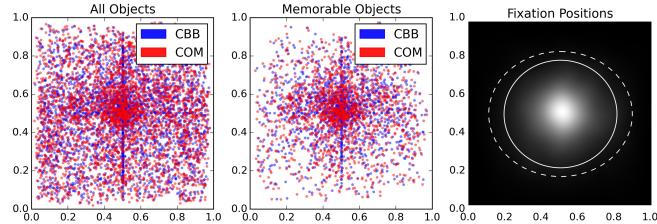
lar performance in doing so. Thus memorability of objects in images can potentially be predicted with high accuracy. Our results confirm that human consistency in remembering particular objects is very high and is as strong as that of images.

3. Understanding Object Memorability

In this section, we aim to better understand object memorability and the factors that make an object more memorable or forgettable to humans. We first investigate if simple color and geometric factors like simple object features like color etc can predict if an object would be memorable or not.

3.1. Can simple color features explain object memorability?

While simple image features are traditionally poor predictors of memorability in full images [Oliva, CVPR 2011], and with good reason [cite [11] from Oliva, CVPR 2011]], it is important to verify that this finding generalizes to individual objects. To do this, we first examined a number of simple color statistics. We decomposed each image into its hue, saturation, and value components and calculated the mean and standard deviation of each channel. Essentially no relationship existed between memorability and either mean hue ($\rho = 0$) or mean saturation ($\rho = 0.02$). This deviates somewhat from the findings related to images that show hue to be weakly predictive of memorability. However, this makes sense since the effect has been 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. This may

324
325
326
327
328
329
330
331
332
333Figure 3: **Positions.** add-in later.

334

335 explain why the effect was not present in our dataset. Looking
 336 at the variance of each channel, we see that variance in
 337 hue ($\rho = 0.16$) and saturation ($\rho = 0.15$) are weakly cor-
 338 related with memorability, while value has a medium cor-
 339 relation ($\rho = 0.38$). This suggests that variation in the color
 340 hue and color purity of an object contribute somewhat to
 341 object memorability. This may be because people are mem-
 342 orable regardless of the color of their clothing, or simply
 343 that color inhomogeneity draws ones attention. The find-
 344 ing that variation in value contributes considerably to object
 345 memorability indicates that high contrast objects are more
 346 memorable. Although the correlation seems high compared
 347 to the performance of past simple features in past research,
 348 the scatterplot in FIGURE X suggests that it may not pro-
 349 duce reliable predictions. Next, we computed image size,
 350 calculated by the number of pixels that make up the object
 351 normalized by the total number of pixels in the parent im-
 352 age. Not surprisingly, this metric correlated strongly with
 353 object memorability ($\rho = 0.53$). This makes sense given
 354 that large objects, especially that take up the majority of the
 355 parent image frame, are more likely to be seen and identi-
 356 fied.

357

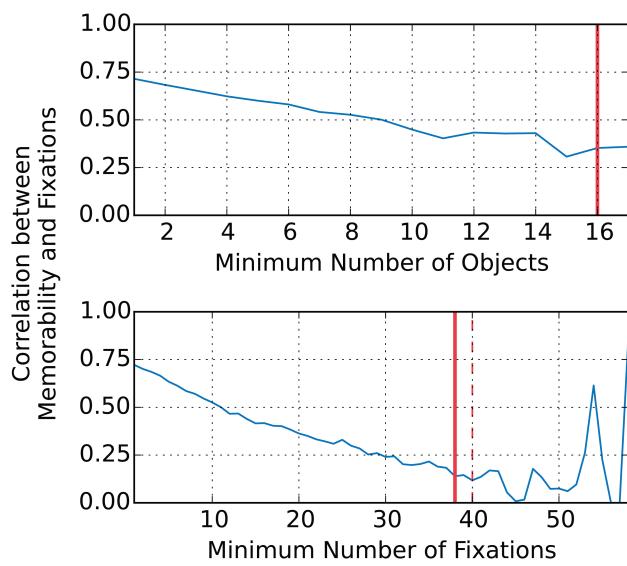
3.2. What is the relationship between visual saliency and object memorability?

361

Note: add in small intro about past work

362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377

To examine the relationship between visual saliency and object memorability, we used two indexes of saliency to de-
 363 scribe each object. The first was calculated as the number
 364 of unique fixation points within the area of each object. The
 365 correlation between object memorability and number of fixa-
 366 tions within those objects was quite high ($\rho = 0.71$). How-
 367 ever, this correlation decreases significantly as the number
 368 of total fixations within the object increases. In figure X,
 369 we show the correlation between object memorability and
 370 number of fixations for subsets of the objects that have at
 371 least some minimum number of fixations. On the leftmost
 372 side of the plot, the correlation is at it's highest since
 373 the set of objects with a fixation count greater than or equal
 374 to zero includes all objects and so the correlation is identical to
 375 FIGURE FIX-CORR. However, as the minimum number of
 376 fixations increases, we look only at objects that contain an

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
431Figure 4: **Correlation between object memorability and number of fixations.** add-in later.

increasingly higher number of distinct fixation points, and the correlation decreases. By the time we reach a minimum fixation count of around 35, the correlation falls to around 0.2. This may mean that the predictive value of saliency decreases when there are many areas of interest within a single object. (NOTE: We may want to show examples of images with lots of fixation points that were not well predicted. We could just plot the fixation locations as points on the image) The reason for this decline may be that...[make arguments here] Alternatively, FIGURE Z shows a similar relationship in which the correlation between memorability and fixation count is plotted as a function of minimum number of objects present in the parent image of the object. In a similar pattern, we see that as the minimum number of objects increases, the correlation between fixation count and object memorability decreases. This follows from the fact that fixation count and number of parent image objects are negatively correlated ($\rho = -0.38$), meaning that if there are more objects in an image, there are less fixation points in any one object, which may a consequence of needing to distribute attention across the objects in the scene. In both scenarios, the ability to predict object memorability sharply decreases.

The second metric we used... (put map method here?)

3.3. What is the relationship between object memorability and object categories?

In the previous sections, we showed that simple features have little predictive power over object memorability and there exists a weak relationship between visual saliency and object memorability. In this section, we show that object

432 memorability is heavily influenced by it's category and in-
 433 vestigate the possible reasons behind the same.
 434
 435
 436

437 3.3.1 Are some object classes more memorable than 438 others?

440 Our dataset contains 3414 object segmentations from the
 441 Pascal-S dataset. For this analysis, we first assigned three
 442 in-house annotators the task of assigning class labels to each
 443 object segmentation in our dataset. The annotators were
 444 given the original image (for reference) and the object seg-
 445 mentation and asked to assign 1 category to the segment out
 446 of 7 possible categories: animal, building, device, furniture,
 447 nature, person, and vehicle. We choose these high-level cat-
 448 egories such that a wide range of object classes could be
 449 covered under these categories. For example, device in-
 450 cluded object segments like utensils, bottle, tv, computer
 451 etc, nature included segments like tree, mountain, flowers,
 452 and vehicle covered segments like car, bike, bus, aeroplane
 453 etc. **add a small table here which shows stats of the object
 454 classes n their frequency in the dataset.**

455 Figure 5 shows the average memorability score for the all
 456 7 different object classes in our dataset. This visualisation
 457 gives a sense of how the memorability of different object
 458 classes differs. Animal, person, and vehicle are all highly
 459 memorable classes with each of them having an average
 460 memorability score greater than 0.5. Interestingly, all the
 461 other object categories have an average memorability score
 462 lower than 0.25 indicating that humans do not remember
 463 objects from these categories very well. In particular, fur-
 464 niture has an average memorability score of only 0.18 and
 465 is the least memorable object class. This could be possibly
 466 due to the fact that most objects from classes like furniture,
 467 nature, and building either appear mostly in the background
 468 or are occluded which could decrease their memorability
 469 significantly. On the contrary, objects from animal, person,
 470 and vehicle classes appear mostly in the foreground leading
 471 to a higher memorability score on average. Table 1 shows
 472 the average memorability of the top 20 most memorable ob-
 473 jects for each object class. This analysis is particularly in-
 474 teresting as these top objects are not occluded and most of
 475 them tend to appear in the foreground. The top 20 most
 476 memorable objects from person, animal and vehicle have
 477 memorability higher than 0.90 whereas the average memo-
 478 rability of the top 20 objects from building, furniture, and
 479 nature is lesser than 0.70. While the differences in the mem-
 480 orability of different classes could be driven primarily due
 481 to factors like occlusion, size, background/foreground, the
 482 results in table 1 suggest that memorability could be an in-
 483 trinsic property of an object class and some object classes
 484 like person, animal, vehicle are in general intrinsically more
 485 memorable than classes like furniture, nature etc.

Object Class	Memorability of top 20 objects
Animal	0.94
Building	0.65
Device	0.86
Furniture	0.63
Nature	0.67
Person	0.93
Vehicle	0.96

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

Table 1: Average memorability of top 20 objects for each class

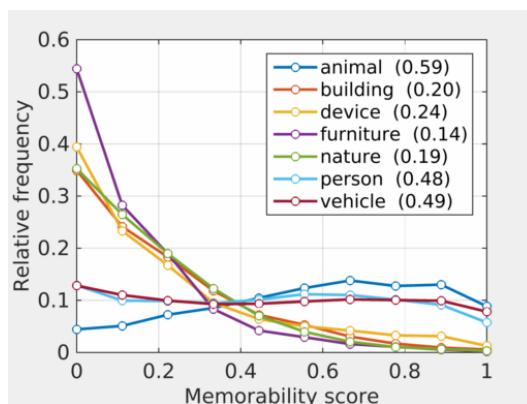


Figure 5: Average memorability per object class. Figure showing some object classes are more memorable than others.

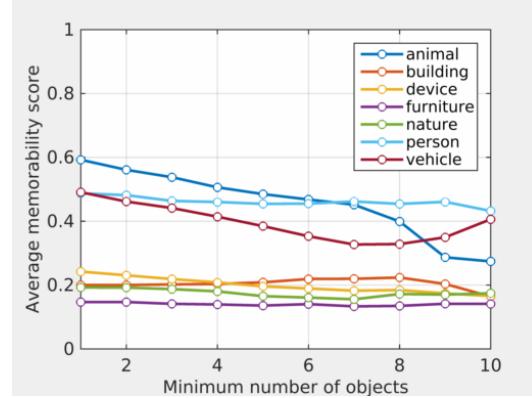
3.3.2 Why are some objects not memorable within a class?

As demonstrated above, some object classes like animal, person, vehicle are more than others either due to their intrinsic property or external factors like occlusion, foreground/background etc. However, not all objects in a class are equally memorable. The examples in figure 6 show the most memorable and least memorable objects for each object class. A common property between these objects is that the least memorable objects within each object class tend to be occluded and obstructed by other objects. What other possible factors could influence the memorability of an object within a class?

Relationship with number of objects in an image: We first tested how the memorability of each object class is affected by the number of objects inside an image. We took in all images that had at least 'n' objects (where $1 \leq n \leq 10$) inside them and computed the average memorability of each object class. Figure 7 shows the change in average memorability of the different object classes with respect to the increase in number of objects in an image. This figure shows that the number of objects present in an image is an important factor that results in the variation in the memorability of objects within a class. As the number of



Figure 6: Qual Results. Figure showing top and bottom objects for each class.

Figure 7: Correlation between object class and number of objects.
add-in later.

objects in an image increase, the memorability of the object classes decreases. In particular, the memorability of objects belonging to highly memorable classes like animal, person, and vehicle decreases significantly with the increase in number of objects in an image. Intuitively, this could mean that with an increase in number of objects, humans cannot concentrate on a single object in an image and have a difficulty remembering objects in those images. Thus, if a person, animal or vehicle is present in a highly cluttered image, humans can not remember those objects very well.

Relationship of inter-class memorability: While the previous analysis sheds light on how the memorability of an object class is effected by factors such as occlusion, and number of objects in an image, how does the presence of a particular object class effect the memorability of another object class? To investigate this, for each possible pair of object categories we took all images that contained at least one object from both the categories and computed the average memorability scores for the two object categories. This way we were able to capture the relationship between pairs of object categories and how their presence effects their respective memorability scores. Figure 8 plots these results and for each object class, shows how it's memorability is effected in the presence of other object classes. The red mark on each line shows the over all average memorability of that class. Firstly, the memorability of low memorable classes i.e. nature, furniture, device, and building is not effected a lot in the presence of other object categories and their memorability tends to be low throughout. The mem-

orability of person class also remains close to it's average memorability score and tends to be unaffected by the presence of most object categories. However, the memorability of person does get effected in the presence of buildings as the average memorability of person class drops below 0.4 for images containing both persons and buildings. This could be due to the fact that there exists images wherein the building is quite large and the persons in those images are usually smaller or less significant in contrast to the building (also illustrated in figure 9). The memorability of the animal class remains close to it's average memorability in presence of most classes, however it drops significantly to below 0.4 in presence of the person class (also shown in figure 9). The memorability of vehicle class is strongly affected by the presence of other object categories. In particular, it's memorability drops significantly in presence of person and animal classes also shown in figure 9.

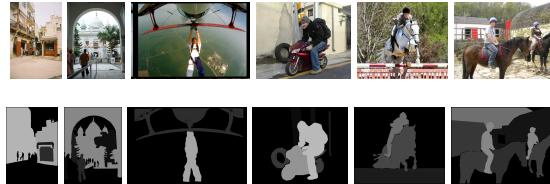
Figure 8: inter-class object memorability relationship.
add-in later.

Figure 9: Qual Results. Figure showing how memorability of different classes is effected in presence of other classes. Bottom row is the memorability map

Taken together, the results from this section show that the

memorability an object within a particular class is greatly effected by occlusion, number of objects present in an image and presence of other object classes in the image.

3.4. What is the relationship between object and image memorability?

[Note: only the skeleton of this section is added, this still needs significant work]

While in the previous sections, we have investigated the various factors that could effect object memorability, an important question remains. Now that we know what objects people remember and the factor that influence their memorability, how does the memorability of objects effect the overall image memorability? To shed light on the relationship between image memorability and object memorability, we conducted another large-scale experiment on Amazon Mechanical Turk for the images in our dataset to gather their respective memorability scores. For this experiment, we followed the exact paradigm as the memory game experiment proposed in [5]. [add details about the fillers, add details about the experiment including number of workers, number of times each image was viewed by workers and explain the consistency analysis]

After collecting the image memorability scores, we conducted several analysis to answer - if an image is highly memorable, what happens to memorability of objects inside the image (and vice versa)? For each image in the dataset, we firstly computed the maximum object memorability within the image. Figure ?? shows the high correlation of maximum object memorability with image memorability ($\rho = 0.43$). Although there is a high correlation between maximum object and image memorability, it still doesn't predict image memorability completely. However, table ?? also demonstrates that maximum object memorability is a very strong predictor of whether an image is highly memorable or low memorability. If we only take the topmost 20 and bottommost 20 images from the dataset, then their average max object memorability is highly correlated with the image memorability scores. Similar trend is observed for 50 most and 100 most top and bottom memorable images. This is also shown via the figure XX [show the qualitative examples here to show memorable images contain memorable objects and non-memorable images do not]. This shows although simply the max object memorability cannot predict whether an image is medium memorable, max object memorability serves as a strong features in predicting if an image is highly memorable or not. Images that are lowly memorable, is possibly due to the fact that they do not contain a single highly memorable object [refer to qual figure again].

4. Predicting Object Memorability

here the benchmarking and baseline code comes up

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

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
 - [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, 7
 - [6] 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
 - [7] 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
 - [8] Y. Li, X. Hou, C. Koch, J. M. Rehg, and A. L. Yuille. The secrets of salient object segmentation. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 280–287. IEEE, 2014. 2
 - [9] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft coco: Common objects in context. In *Computer Vision–ECCV 2014*, pages 740–755. Springer, 2014. 3
 - [10] T. Liu, Z. Yuan, J. Sun, J. Wang, N. Zheng, X. Tang, and H.-Y. Shum. Learning to detect a salient object. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 33(2):353–367, 2011. 3
 - [11] M. Mancas and O. Le Meur. Memorability of natural scenes: the role of attention. In *ICIP*, 2013. 2
 - [12] C. Yang, L. Zhang, H. Lu, X. Ruan, and M.-H. Yang. Saliency detection via graph-based manifold ranking. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 3166–3173. IEEE, 2013. 3