

Supplementary **M**aterials

for

Eye-movements reveal semantic interference effects during the encoding of naturalistic scenes in long-term memory

^{a,b} Anastasiia Mikhailova (a.mikhailova@psicologia.ulisboa.pt)

^a Ana Raposo (alraposo@psicologia.ulisboa.pt)

^c Sergio Della Sala (sergio@ed.ac.uk)

^{a,d} Moreno I. Coco* (M.Coco@uel.ac.uk)

^a Faculdade de Psicologia, Universidade de Lisboa, Alameda da Universidade, 1649-013, Lisbon, Portugal

^b Instituto Superior Técnico, Universidade de Lisboa, Av. Rovisco Pais 1, 1049-001, Lisbon, Portugal

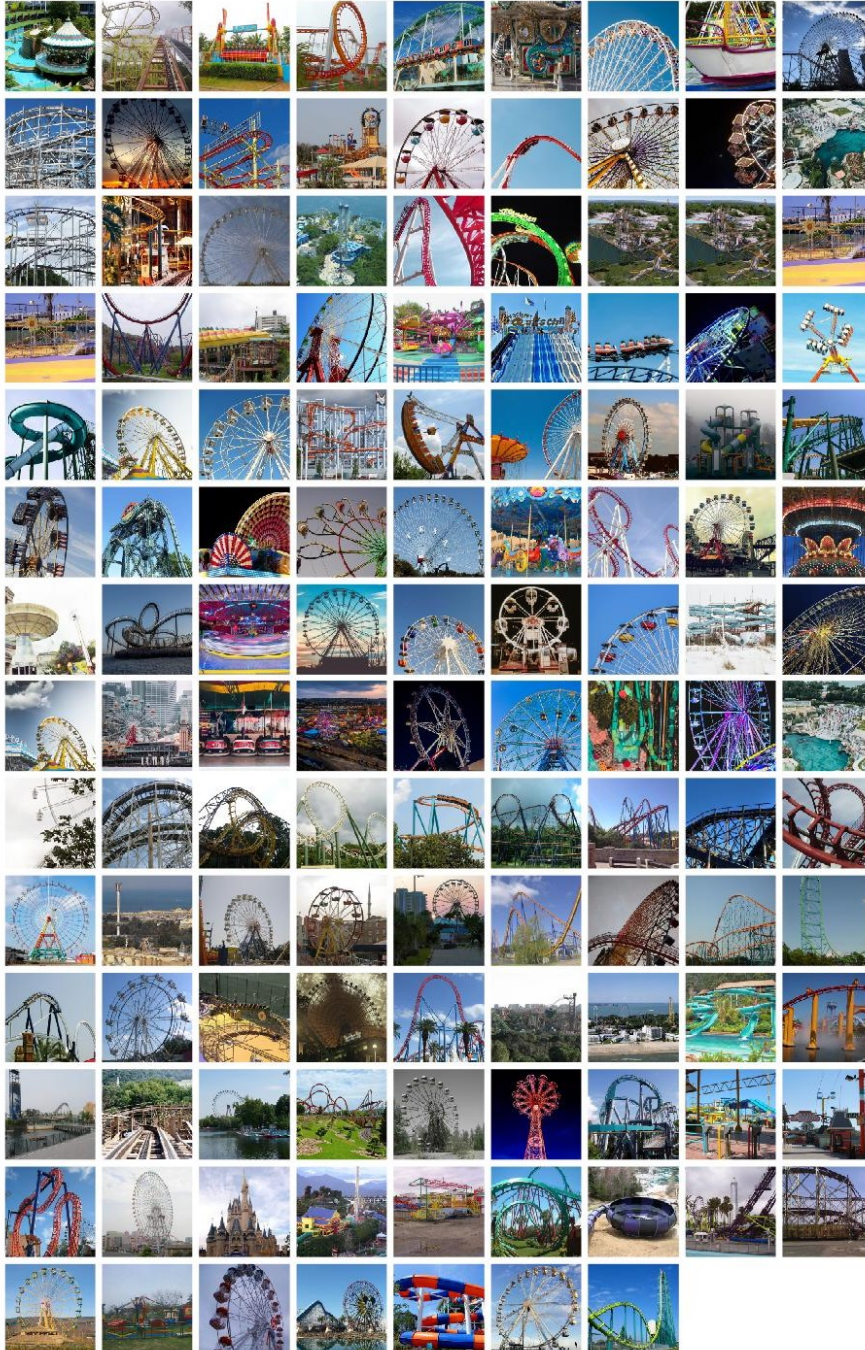
^c Human Cognitive Neuroscience, Psychology, University of Edinburgh, 7 George Square, EH8 9JZ, Edinburgh, UK

^d School of Psychology, University of East London, Water Lane, E15 4LZ, London, UK

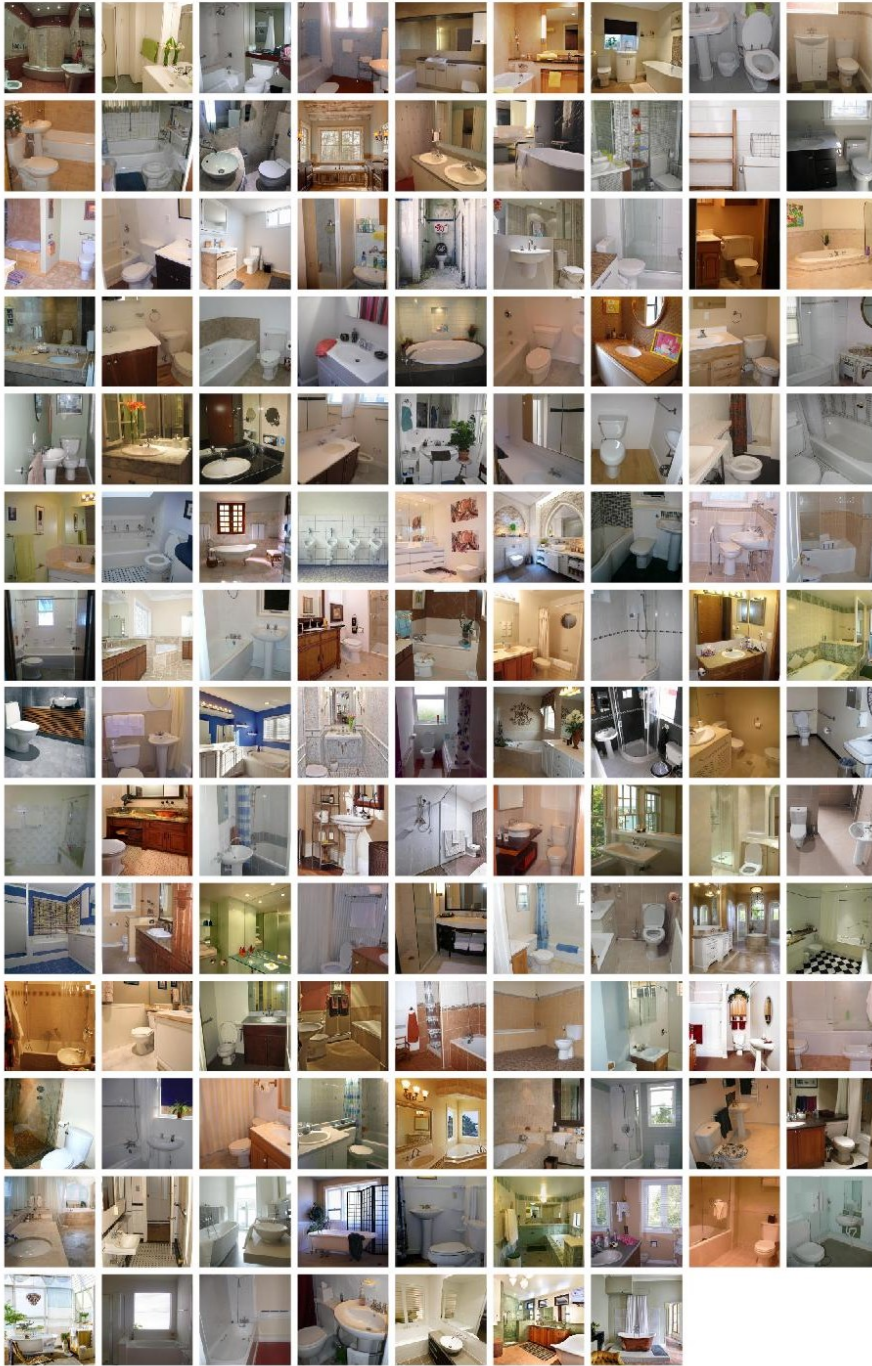
Supplementary S1

Miniatures of all images used in this study by category

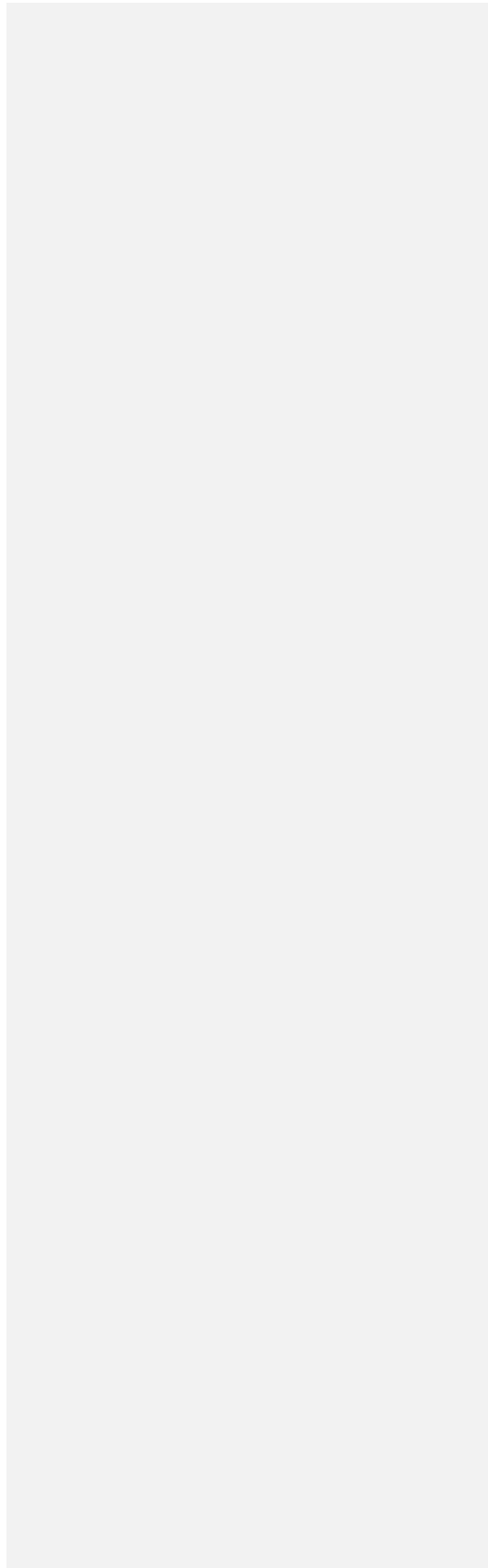
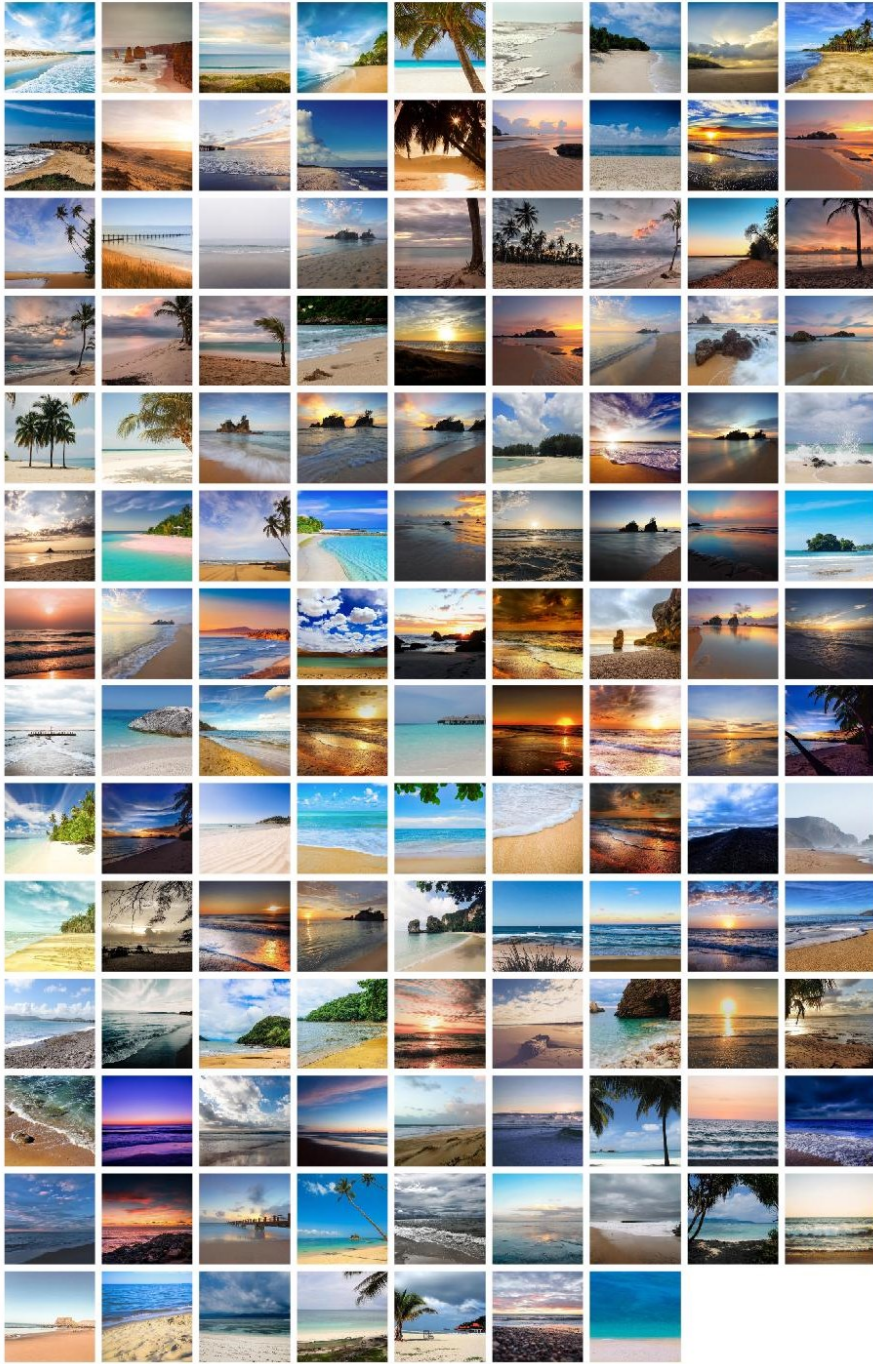
Amusement park



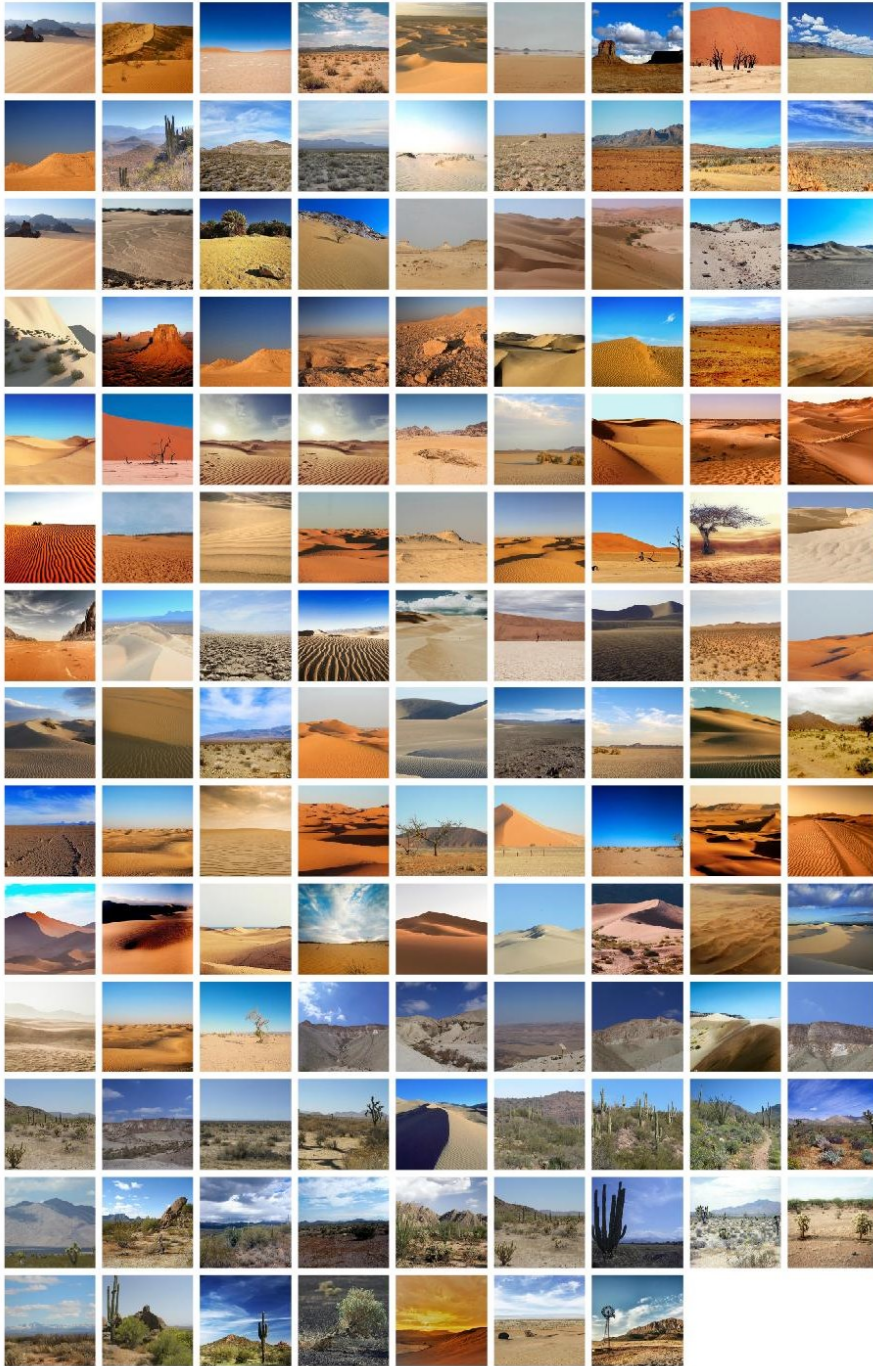
Bathroom



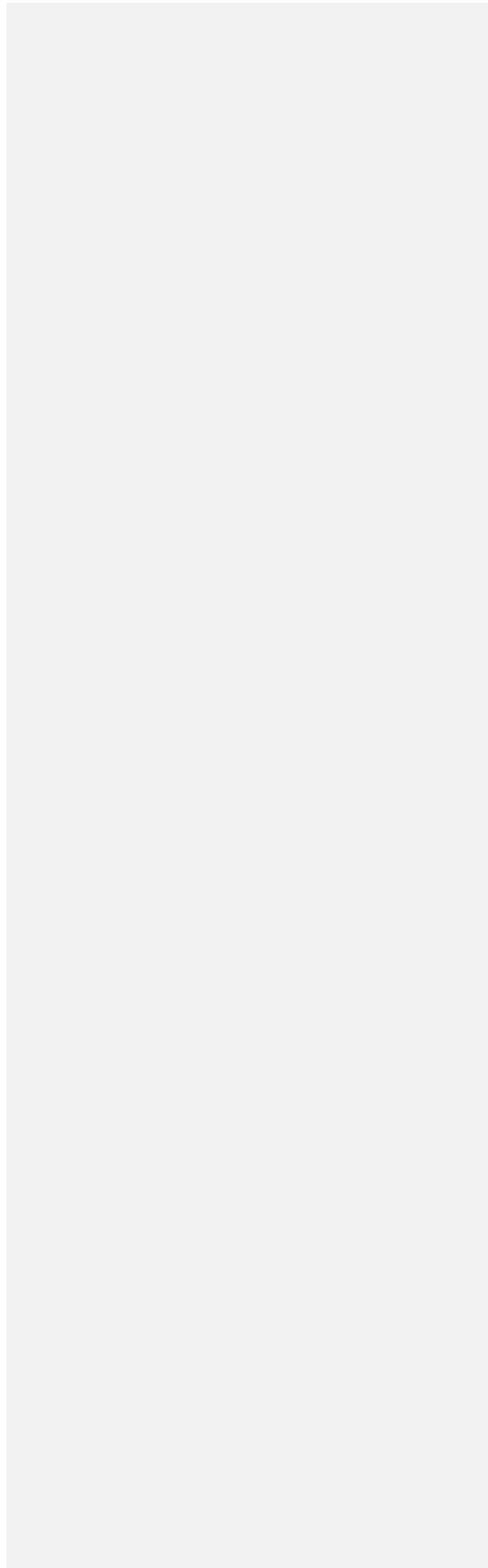
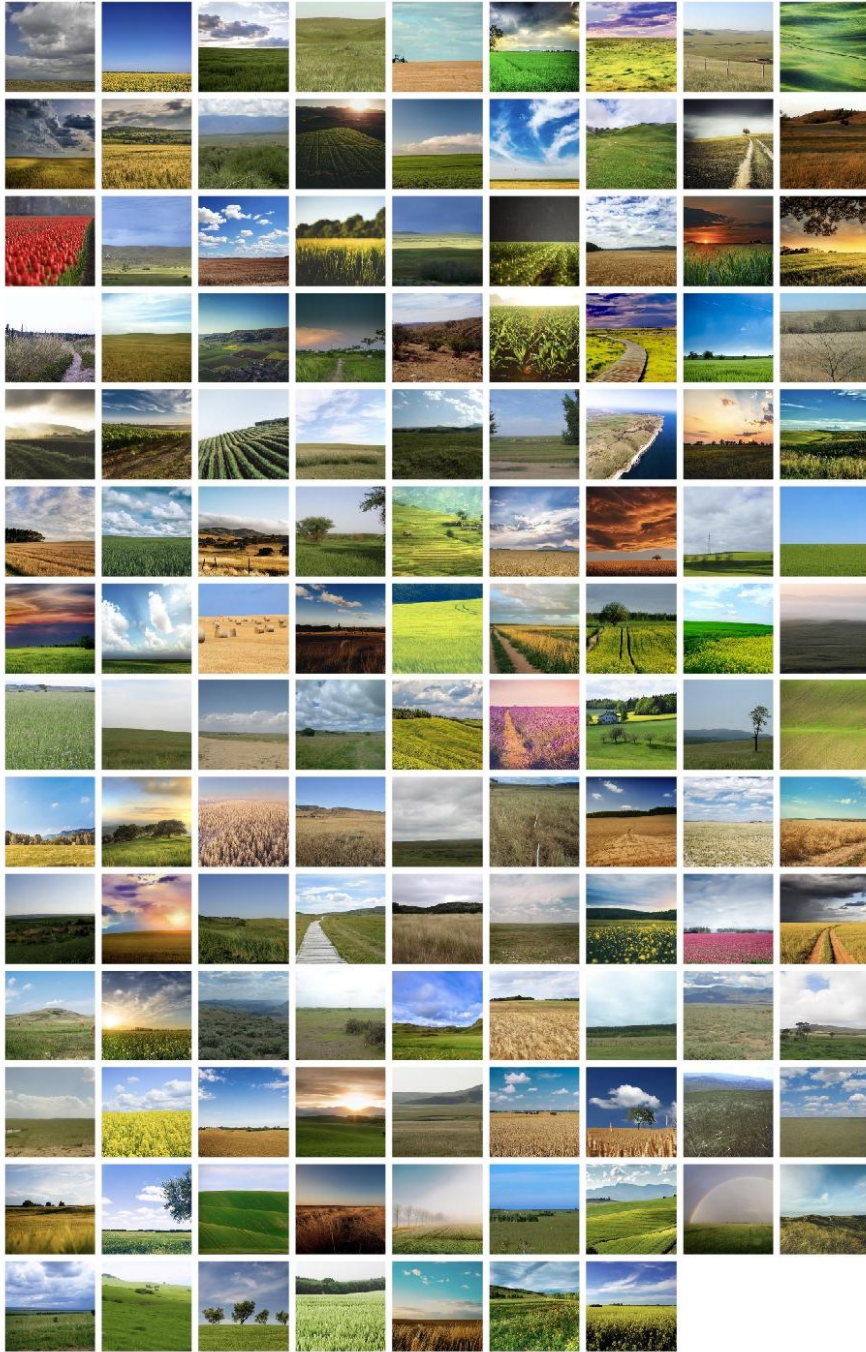
Beach



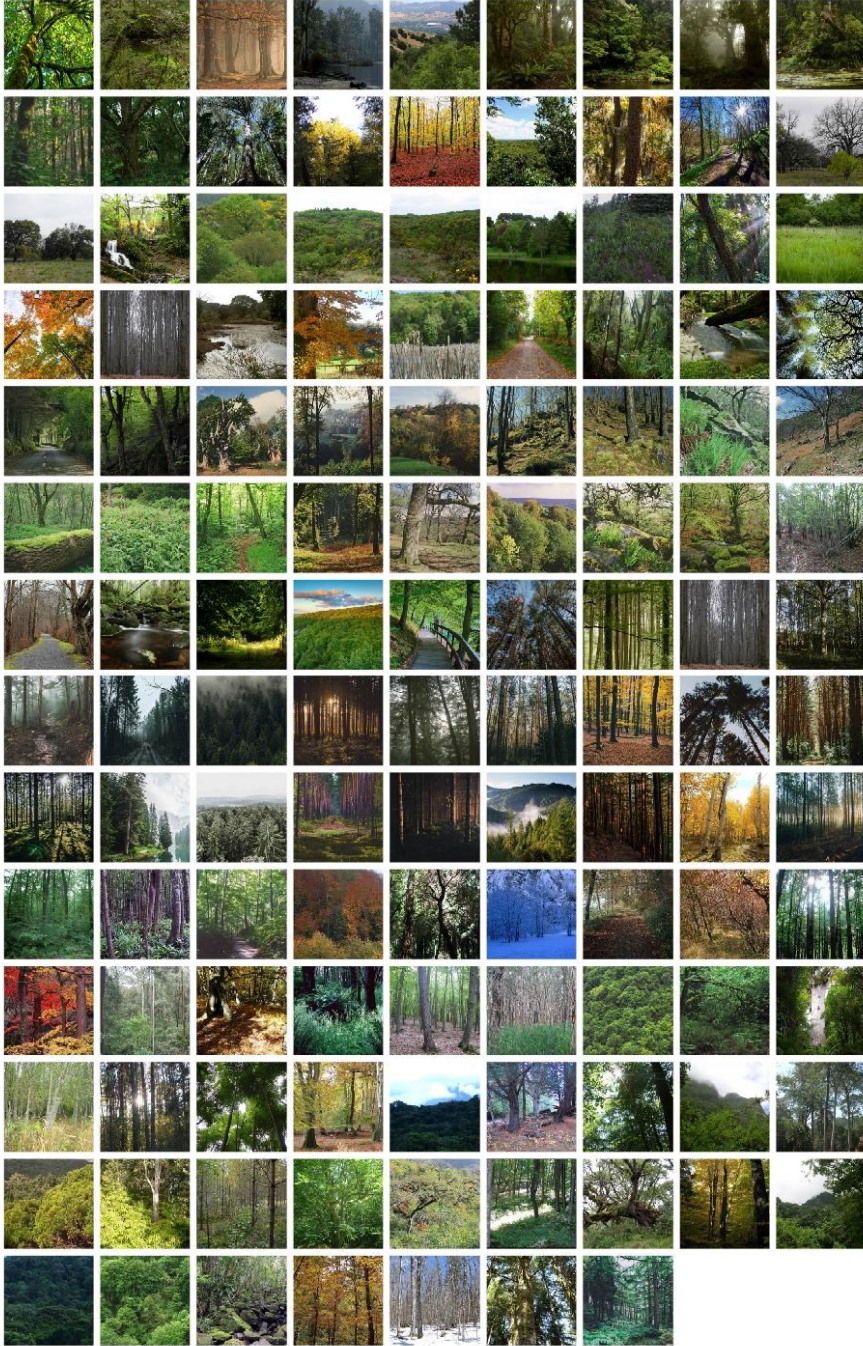
Desert



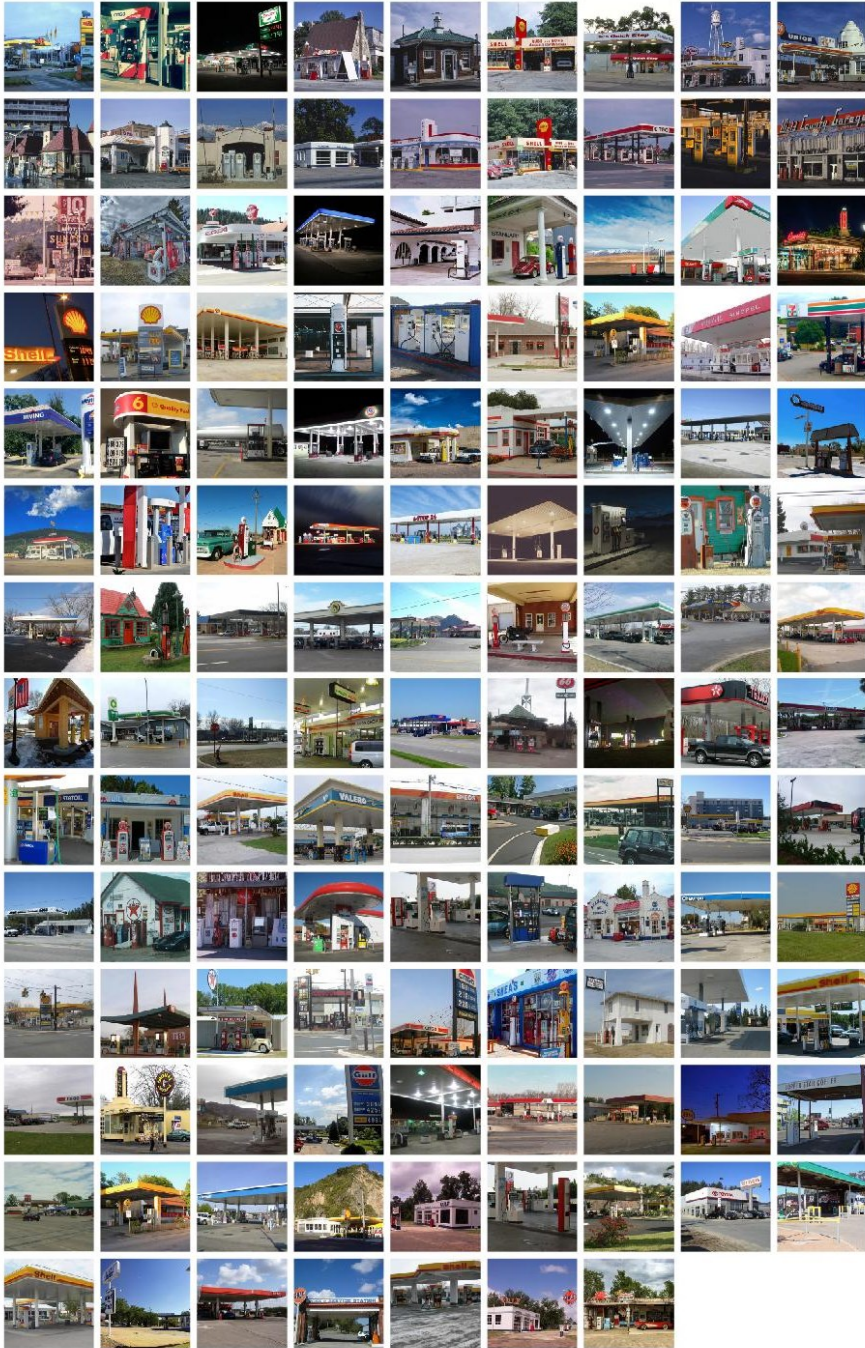
Field



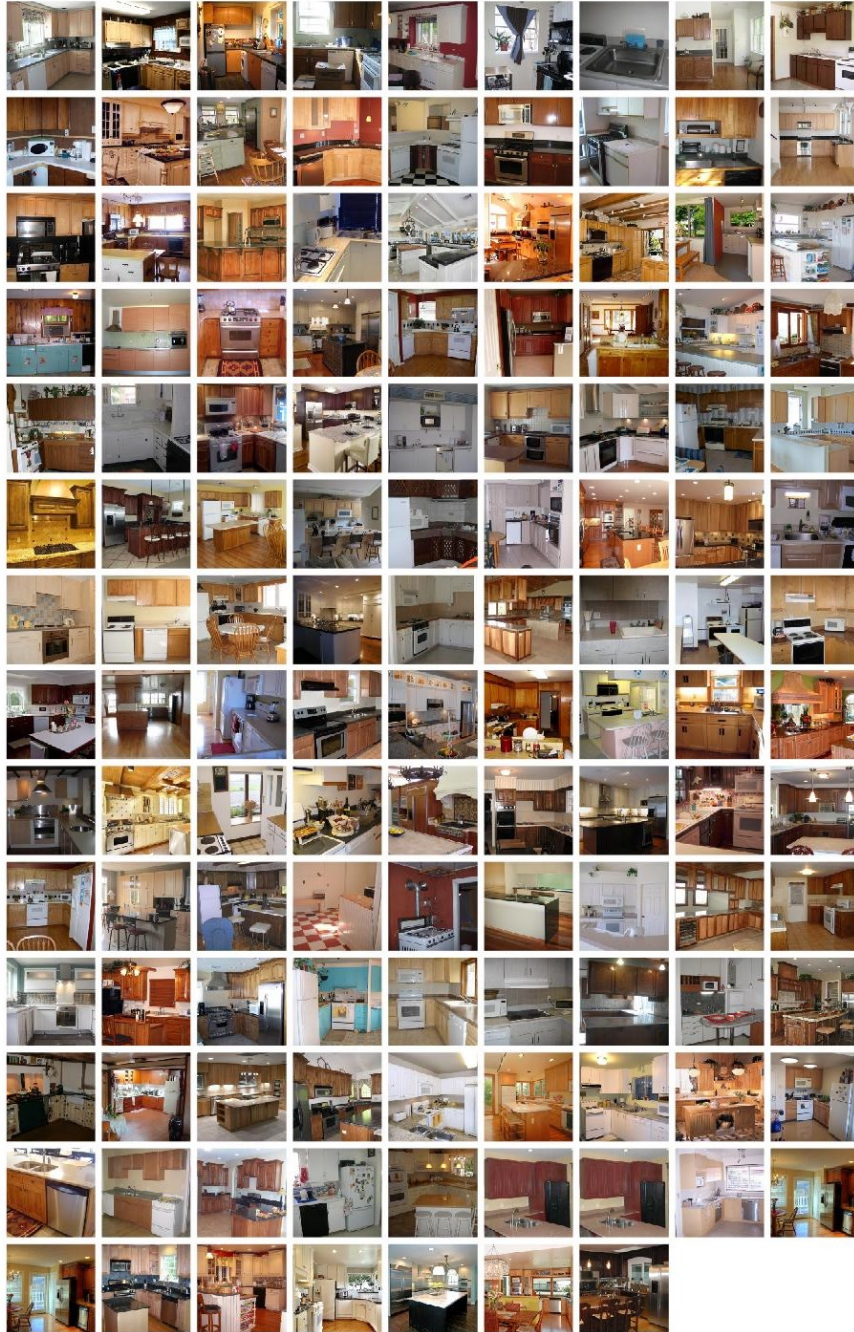
Forest



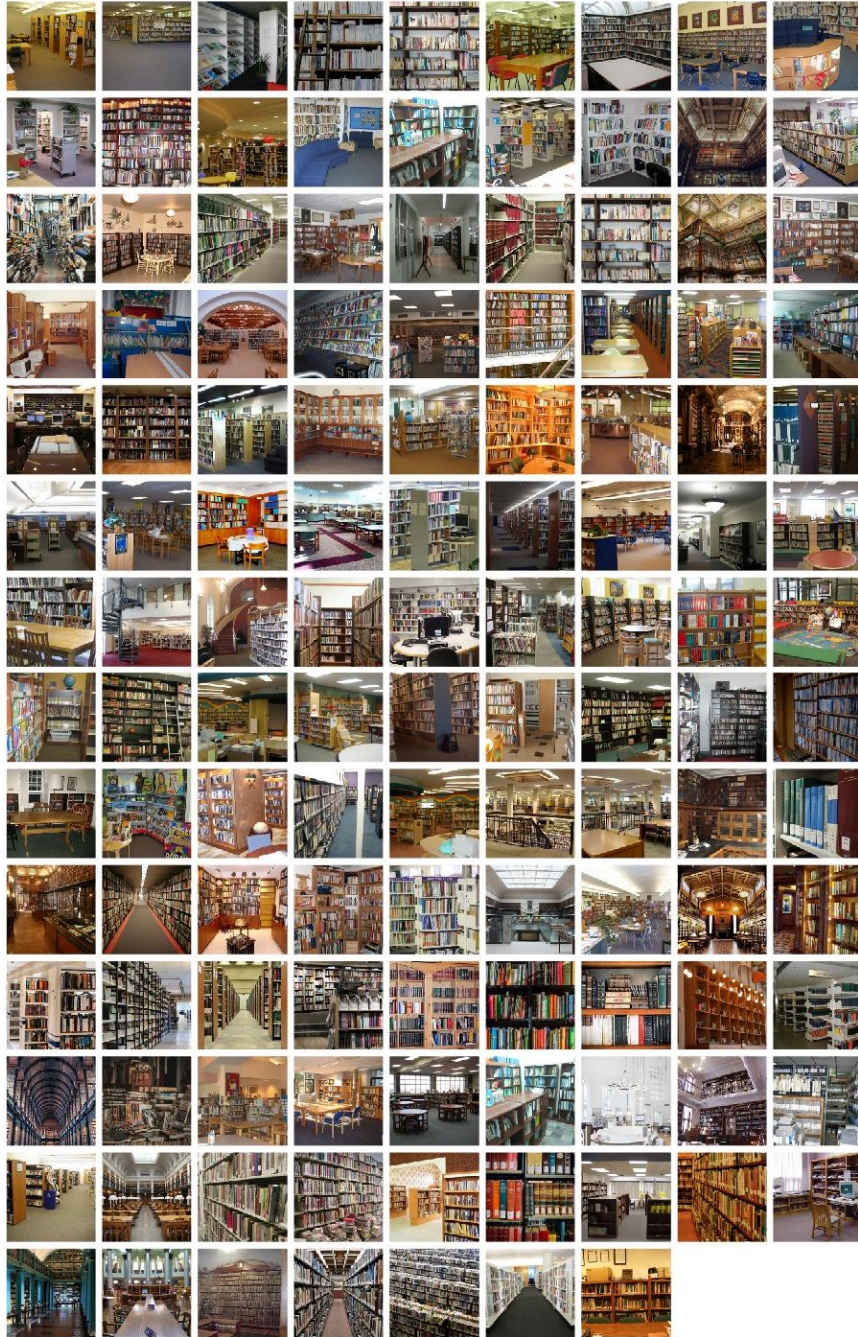
Gas station



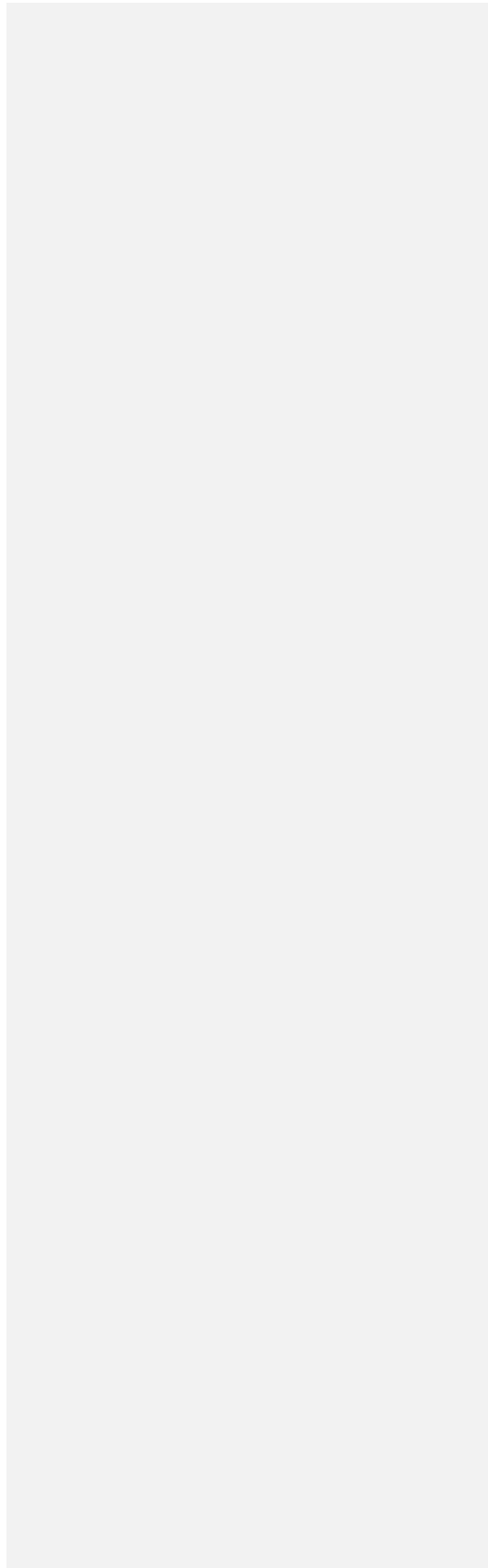
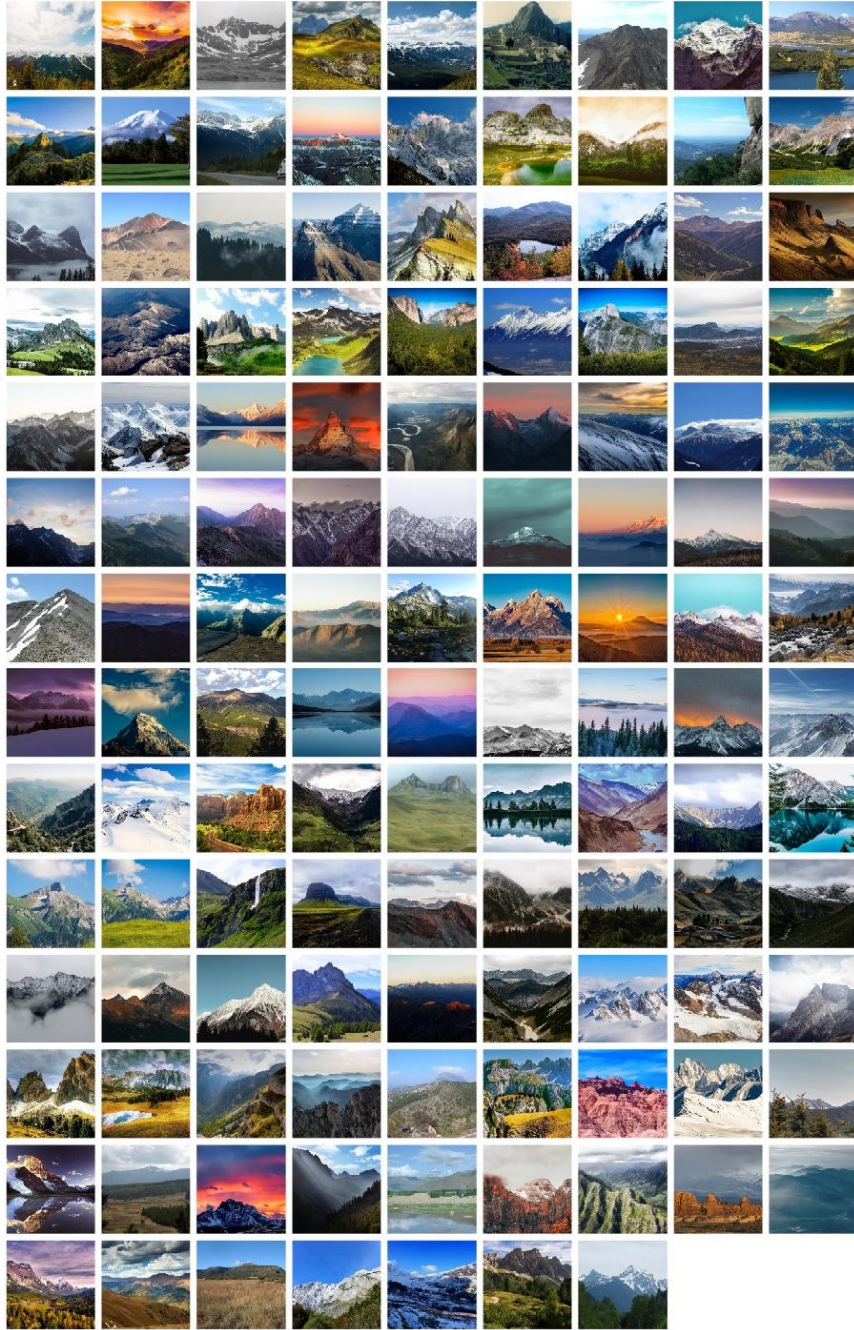
Kitchen



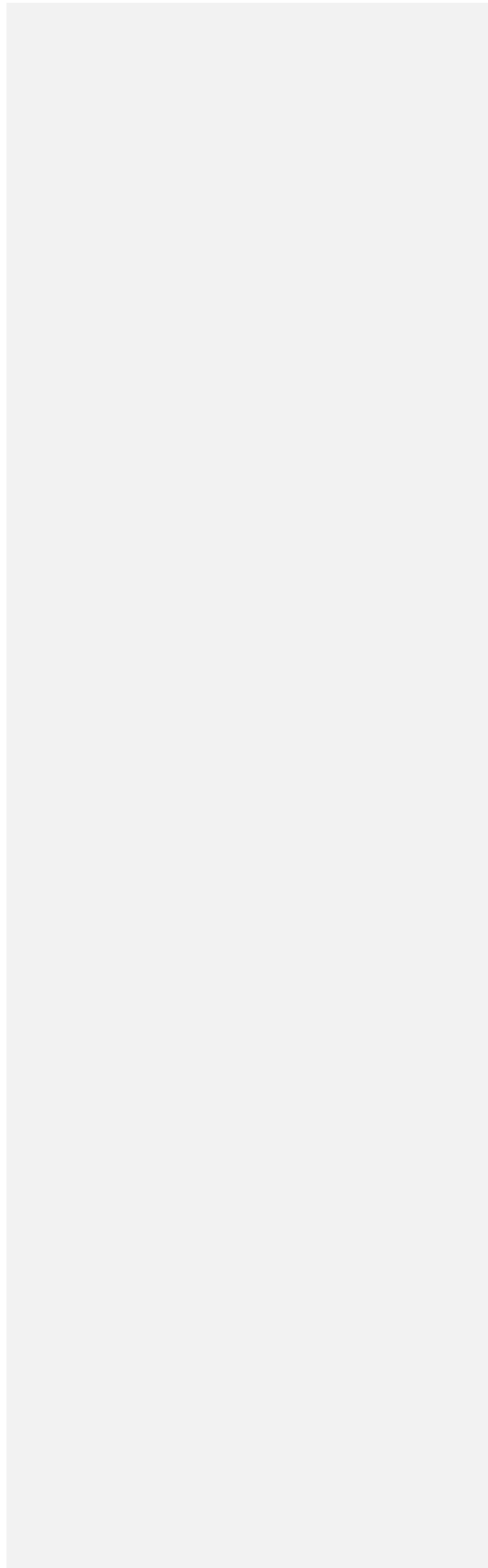
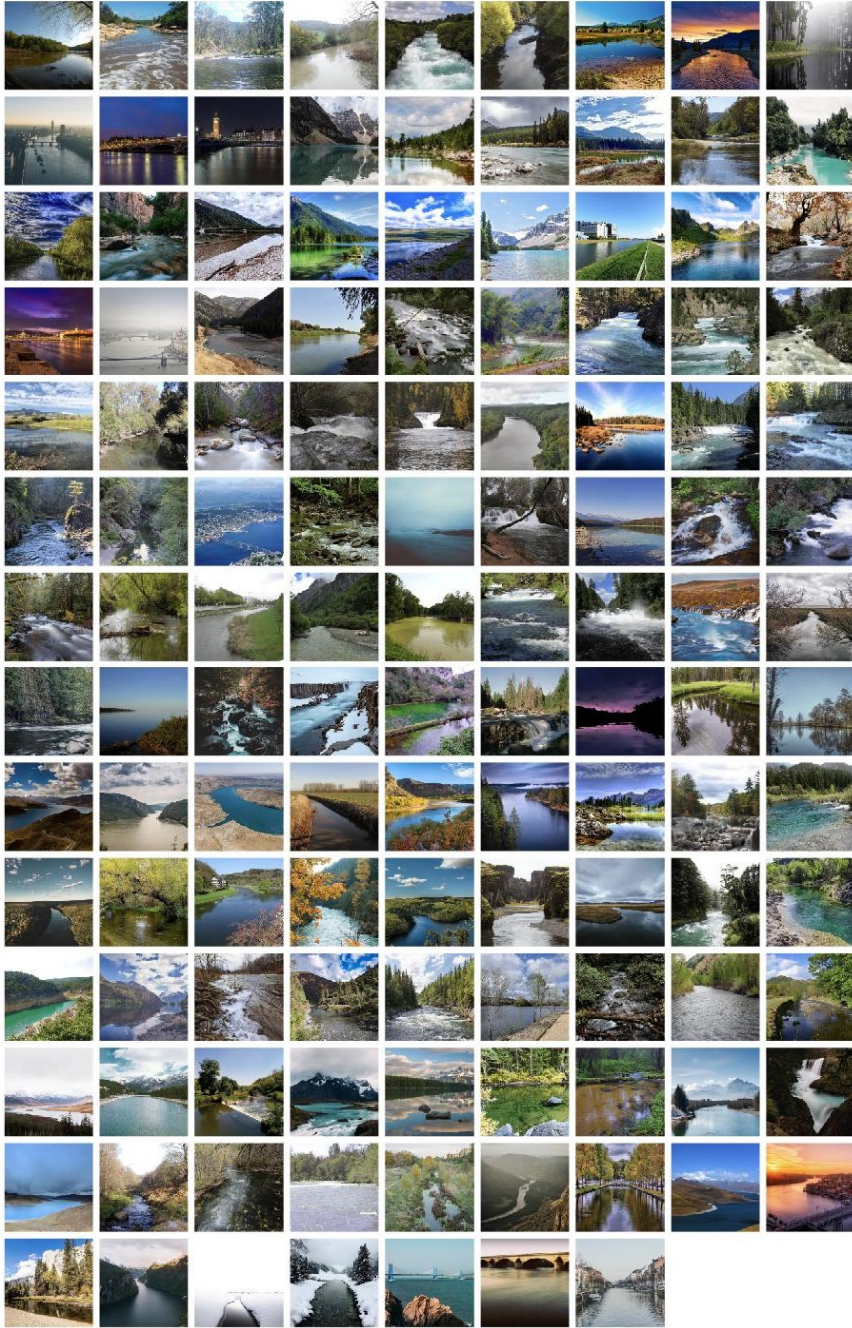
Library



Mountain



River

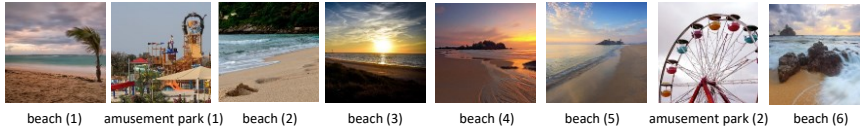


Supplementary S2

Illustration of semantic interference

Figure S2:

Semantic interference as continuous variable:



Semantic interference as frequency:



Figure Fig. S2: Example of semantic interference for ~~8~~-eight scenes in two semantic categories: amusement park and beach. Top: (continuous) semantic interference defined as the rolling number of scenes from the same semantic category that participants were administered (e.g., 1, 2, 3 ... 20). Bottom: (categorical) semantic interference as frequency level per category (4, 20, 40, 60).

Formatted: Font: Bold, Not Italic

Formatted: Font: Bold, Not Italic

Supplementary S3

Recognition accuracy predicted as a function of eye-movement measures and semantic interference

Figure S3

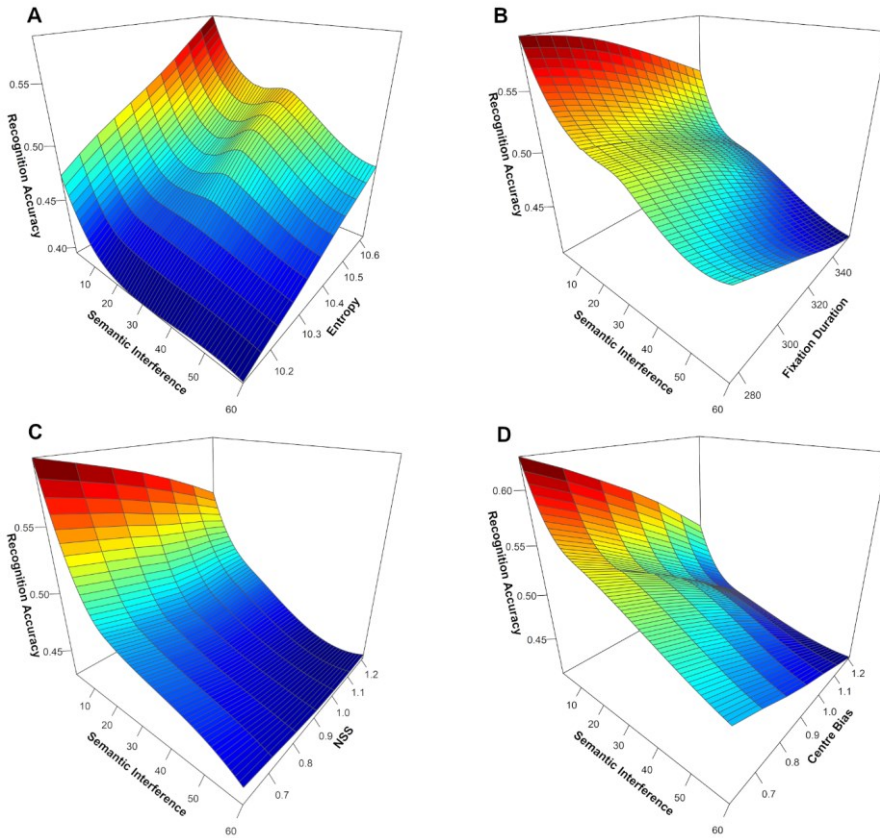


Figure Fig. S3: 3D visualization of the predicted values obtained using a local polynomial regression (loess function in R) fitting recognition accuracy as a function of semantic interference and each eye-movement measure (i.e., fixation entropy, fixation duration, NSS [(saliency)], NSS (centre-bias)-

Formatted: Font: Bold, Not Italic

Formatted: Font: Bold, Not Italic

Supplementary S4

Center bias, visual saliency, and memorability: ~~R~~eplicating and extending Lyu et al., 2020.

Formatted: Font: Not Bold

Formatted: Font: Not Bold

In this section, we replicate the analysis conducted by Lyu and colleagues (2020) on the relation between central bias and memorability of naturalistic scenes while extending it to include oculo-motor responses. We used the Graph-Based Visual Saliency (GBVS) algorithm with default settings (Harel et al., 2006) to compute the saliency map of each scene, which was normalized to sum to 1 (Fig. S4b). A central weight map (Fig. S4c) was also calculated using a Gaussian kernel with $\sigma = 10\%$ of the image height (i.e., 80 pixels). Then, each GBVS saliency map was multiplied by the central weight map to obtain a GBVS central bias map (Fig. S4d). From this map, in line with Lyu and colleagues (2020), we obtained a single central bias score, which conceptually represents how visually salient the center of the image is. We normalized center bias scores to range between 0 and 1 by dividing all scores for the maximum value observed in the scene set (referred as *GBVS central bias* in the Table S4). As this measure used by Lyu et al. (2020) does not incorporate eye-movement responses, we also computed the Normalized Scanpath Saliency (NSS; Peters et al., 2005) correspondence between fixation positions and the GBVS visual saliency maps, adjusted by center proximity (*NSS: fixation/GBVS adjusted by central bias*) or not (*NSS: fixation/GBVS saliency only*).

We applied the same analytical strategy described in the main text, whereby the three different GBVS measures (DVs) are predicted as a function of semantic interference and recognition memory (IVs) using linear-mixed effect models.

Our analysis shows that both NSS scores are significantly predicted by recognition memory whereby a higher reliance on low-level visual saliency implies a poorer recognition performance. None of the three metrics was instead significantly associated with semantic interference (see Table S4). These results align with those presented in the main text, whereby attending low-level visual information, as computed using FES, negatively impacted on scene memorability. However, an important difference between using GBVS or FES is that the former did not show a significant effect of semantic interference (although the trend was in the right direction), whereas the latter did. One possible explanation of this discrepancy is that FES better predicts viewing behavior than GBVS (see Bylinskii et al., 2019, for an evaluation of different saliency models). On center bias, as defined by Lyu et al. (2020) instead, we did not find any significant association with recognition memory, in line with their results.

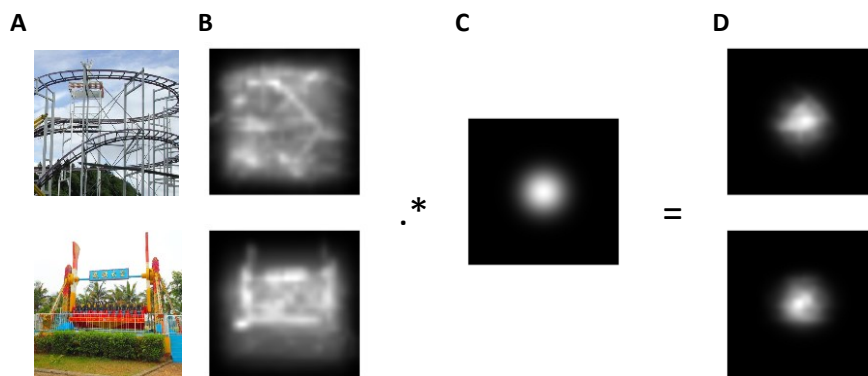
Taken together, this analysis confirms that low-level visual saliency is significantly associated with scene memorability, with and without center bias adjustments, and that this result emerges when oculo-motor responses are considered.

Table S4

Predictors	NSS: fixation/GBVS saliency only			NSS: fixation/GBVS adjusted by central bias			GBVS central bias		
	Estimates	SE	t-value	Estimates	SE	t-value	Estimates	SE	t-value
(Intercept)	0.87	0.04	19.88***	1.01	0.06	16.11***	0.42	0.007	56.25***
SI	0.01	0.005	1.45	0.01	0.008	1.57	-0.001	0.001	-0.07
Accuracy	-0.02	0.005	-4.89***	-0.07	0.008	-8.31***	-0.001	0.001	-0.49
SI:Accuracy	0.002	0.005	0.34	0.002	0.007	0.31	0.001	0.001	0.66

(*) $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Fig. S4 a Original images. **b** GBVS saliency maps. **c** Central bias weight map. **d** GBVS central bias maps (see Fig. 2 in Lyu, et al., 2020, for a similar visualization)



References

Bylinskii, Z., Judd, T., Oliva, A., Torralba, A., & Durand, F. (2019). What do different evaluation metrics tell us about saliency models? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(3), 740–757.

<https://doi.org/10.1109/TPAMI.2018.2815601>

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. *Advances in Neural Information Processing Systems*, 19, 545–552.

[Lyu, M., Choe, K. W., Kardan, O., Kotabe, H., Henderson, J., & Berman, M. \(2020\). Overt attentional correlates of scene memorability and their relationships to scene semantics. *Journal of Vision*, 20\(9\), 1–17. <https://doi.org/10.31234/osf.io/3e8qm>](#)

Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in

natural images. *Vision Research*, 45(18), 2397–2416.

<https://doi.org/10.1016/j.visres.2005.03.019>

