Understanding Image Memorability: Neural Correlates, Behavioral Characteristics, and Predictive Models

Amir Shokri
Electrical & Computer Engineering Department
Semnan University
Semnan, Iran
amirsh.nll@gmail.com

***Abstract*—This paper presents a comprehensive review on the evolving research landscape of image memorability, aiming to synthesize existing knowledge and identify gaps in understanding the roles of emotional content, aesthetics, and visual salience. Through a systematic examination of recent studies and the critical analysis of the LaMem dataset, this review explores advancements in computational models for predicting image memorability, highlighting the shift towards classification-based approaches and the incorporation of semantic features. The findings reveal a complex interplay of factors influencing memorability and challenge the efficacy of traditional regression-based models, suggesting that the novel classification approach not only benchmarks but also, in some aspects, surpasses human consistency. By offering insights into the cognitive processes behind image memorability and proposing a novel computational model, this review contributes to a deeper understanding of the subject and outlines directions for future research, emphasizing the need for a holistic approach to studying image memorability.**

I. Introduction

To what extent can an image be remembered? Certain images are undeniably more stirring than others, particularly those featuring familiar content, such as photographs of friends or family, places visited, and events attended. Studies show that, however, some images are inherently more memorable than others without such recognizable mental content [2]. The second case, according to recent studies, is measurable memorability [9], very closely related to semantic features [4] and to some extent even predictable [9]. The advertising industry stands to gain significantly from the prediction of image memorability, utilizing the memorability score to quantitatively assess the efficacy of prototype advertising designs across different consumer products. Additionally, there is a potential impact on the system's capacity to enhance object recognition or scene comprehension through predictive memorability.

The innate attributes of memorable images, however, remain not entirely comprehended. Content and spatial features such as "enclosed spatial structure" and "people with visible faces" positively affect memorability scores, while aesthetic beauty and unusualness of an image are negatively associated with memorability scores [4].

In today's dynamic environment, media platforms like media advertising, social networks, recommendation systems, and information retrieval require substantial computational resources to manage the expanding volume of data. Consequently, the capability to comprehend content within these media systems plays a pivotal role in optimizing their processing. Various notions such as aesthetics and visual salience [1], emotion [3], attractiveness [3], social popularity [3], and memorability [6] may interfere with content comprehension. This study concentrates on the memorability of images, an emerging and underexplored concept within the realms of computer vision and multimedia.

While image memorability has become a topic of interest in the field of computer vision more recently, the concept of visual memory has been studied extensively within psychology for numerous decades. Research has demonstrated that individuals can recall a vast array of images after a single viewing, even when later presented with a multitude of similar images [12]. However, the likelihood of recalling any given image can vary based on factors such as the context in which the user views the image and the image's specific visual attributes [6]. Notably, despite the subjective nature of memorability, there tends to be a consensus among people regarding the memorability of certain images [12]. These insights provide a valuable foundation for further research aimed at understanding the elements that make an image particularly memorable to the average person.

Researchers have investigated the external and internal features in computer vision that make a photo memorable [15]. and makes an image memorable [6], showing that color, object statistics (number of objects, simple image features derived from pixel statistics, reporting average pixel coverage in current object classes) are not strongly correlated with memorability. Meanwhile, scene has a high correlation and object semantics with memorability. Memorability is not connected to other subjective concepts like attractiveness and aesthetics. Importantly, the experiments in [6] show sufficient human adaptation during annotation. have confirmed picture memory and thus prove the possibility of predicting the memorability of pictures. Several small datasets have been publicly released to support research in this area, such as scene categories [14, 15], face image datasets [17], affective effect on image memorability [18], and visualization images [18]. The first large-scale image memory dataset (LaMem) in particular, thanks to some research at MIT, contains approximately 60,000 crowdsourced annotated images along with a memory prediction model (MemNet) to benchmark published work. has been [16].

This study involves creating a computational model using the LaMem dataset to estimate the memorability of images. The results show that this new model surpasses the performance of the previously advanced MemNet model and even exhibits greater consistency than human evaluations within the LaMem dataset. Two primary aspects distinguish this model from earlier ones. First, it approaches the task of prediction not as a regression issue, as was done in prior works, but rather as a classification problem, estimating the likelihood of images falling into various memorability classes. The emphasis is on gathering images that are likely to be memorable. Second, the model doesn't just depend on features obtained from a convolutional neural network (CNN) trained on other tasks or on simple hand-crafted image features. Instead, it also incorporates additional semantic features that are associated with the image captions to refine its predictive performance.

The remainder of the study is structured as follows. Section 2 outlines the suggested computational model for predicting image memorability. Section 3 contains empirical results, encompassing our exploration of the correlation between intriguing and memorable concepts, as well as the generality of the model. The study concludes in Section 4.

II. The Behavioral Characteristics of Image Memorability

A. *The Multifaceted Connection Between Image Content and Memorability*

The study of how certain aspects of images affect their memorability reveals a complex interplay between the content of the image and its likelihood to be remembered. Prior to the use of memorability scores, research delved into a range of image attributes to understand what makes some stand out in our memory. These studies examined both basic aspects, like the difference in memorability between images presented in three-dimensional format versus two-dimensional, or in full color versus grayscale, and more complex traits such as the uniqueness or rarity of the subject matter, particularly with human faces. Additionally, it has been noted that the capacity to remember an image is not just a matter of visual perception; conceptual uniqueness also contributes significantly to whether an image sticks in our memory. Simultaneously, this preceding research highlighted the intricate nature of visual memory storage, underscoring the reliance on specific details for recollection, such as recognizing whether recognition of a particular image has occurred before [22]. Additionally, it emphasized the ability to remember substantial details about the arrangements and contexts in which objects are perceived [2]. Despite this, our capacity for recalling random patterns is notably limited unless they exhibit object-like characteristics [23], indicating that visual memory is not solely driven by visual details. These findings have been synthesized under the concept of "meaningfulness," suggesting that memorability is enhanced when images contain recognizable content, emphasizing that such meaningful images are more effectively remembered compared to those lacking such content [24].

Utilizing these fundamental findings, the process of assigning memorability scores to individual images involves assessing the relative significance of established factors, identifying the proportion of variations, and uncovering novel memorability influencers for the entire image. The amalgamation of all these factors aims to comprehensively elucidate the overall memorability of an image. Examples of the primary factors influencing variability in image memorability include the observation that, unlike landscape images, they exhibit a lower mean memorability (mean memorability score = 0.61). Images featuring people tend to be notably memorable on average (mean memorability score = 0.82) [7]. Furthermore, images portraying unconventional content, like a hand-shaped chair [25], generally have high memorability (mean memorability score = 0.83). On average, stimuli eliciting fear and amusement are more memorable, whereas images evoking satisfaction and awe tend to be less memorable [5]. As previously mentioned, memorability demonstrates a weak association with image attractiveness and subjective aesthetic judgments [5]. Similarly, while low-level image features such as basic image features and color contribute to the diversity of memorability, their impact is relatively modest [7, 26].

B. *Image memorability for recognition memory vs recollection*

A broader exploration of image memorability involves the investigation of "recognition memory," wherein participants are tasked with determining the familiarity of images (Box 1). An alternative yet complementary memory task, extensively studied in the context of word lists [27], involves presenting subjects with images and subsequently challenging them to recall what they saw without any cues, making this memory task more challenging than recognition memory. Is this the most challenging memory task? A recent study addressed the challenge of quantifying the variability of recall for picture recall by asking subjects to draw and then view their pictures [28]. They found no relationship between changes in picture recall quantified for recall versus recognition memory [28]. These findings align with the notion that the alterations in memorable pictures for these two memory tasks might be different.

C. Image Memorability and other cognitive phenomena

Memorability of images is associated with various discernible factors, and one such factor is the visual salience of cognitive phenomena. This pertains to the phenomenon where specific regions in images emerge and capture our attention. By examining the consistency of fixation patterns during free viewing across subjects, the visual salience of an image can be gauged [29]. Numerous studies have established a correlation between measures of visual salience and image memorability [5, 26, 30, 31, 32]. Like recall, the saliency of images increases when they feature one or more objects [26, 29], especially those presented predominantly in close-up and within a cluttered context [5]. However, in instances where images contain multiple fixation points and various objects, the correlation between memorability and saliency significantly diminishes [26], indicating a distinction between saliency and memorability. Similarly, differences in memorability exist among identical face images concerning parts, shape, fixation patterns, and image features [8, 17]. The research indicates that while there is a connection between what captures our visual attention and what we tend to remember, these concepts are not the same. In line with this, a recent investigation into how memorability is linked with various cognitive processes — such as directing attention instinctively through cues and searching, or intentionally through mental control and focusing on certain depths, as well as the effect of priming — found that these processes do not fully explain the variations in memory strength. This underscores the unique quality of memorability as separate from these other mental phenomena.

D. Image memorability depends on image set context

Memorability scores for images exhibit strong reproducibility when observed within the context of a randomly chosen set of other images. The reliability of a cross-subject ratings for image memorability remains consistent even when similar images from the same category are presented sequentially within a series of selected images, such as embedding a photo of a lighthouse within a sequence of other lighthouse images. but they take on new values [31]. It is possible to predict the memorability scores of the image compared to the random criterion by the degree of distinctiveness of an image from other images in the new image set with contextual changes in magnitude and sign. scene classification)[31]. Images exhibiting the most comparable activation patterns to the rest of the set in a categorization context experience a significant decrease in memorability, while images that stand out the most may show a relative increase in memorability. These findings imply that a comprehensive understanding of image memorability necessitates not only a delineation of individual image characteristics but also an appreciation of the context in which those images are situated.

III. The Neural Correlates of Image Memorability

A. Image memorability is reflected in the magnitude of the response to novel images

The neural underpinnings of image memorability can be seen in the amplitude of responses to new images. When exploring the cognitive processes that enhance an image's memorability, it appears that changing how images are presented in a new context can affect their memorability and the likelihood of their being remembered later. Alternatively, changes in memorability may become apparent when images are recognized as familiar, implying that the concept of memorability could be tied to the processes of memory signaling or storage. Current evidence seems to support the first hypothesis more, without completely ruling out the latter. Early studies using functional magnetic resonance imaging (fMRI) identified areas of the brain associated with the memorability of images [14, 15]. These studies suggested that memorability could be determined within certain categories, like faces or scenes, by measuring the blood oxygen level-dependent (BOLD) signal increases in the higher visual cortical areas as subjects looked at new images. Similar findings were reported when using electroencephalogram (EEG) readings, where images that elicited stronger N170 brainwave activity upon first sight were more likely to be remembered [24]. This was further supported by studies using magnetoencephalography (MEG), where the identity of pictures was more accurately inferred from brain responses to less memorable images, even when these images were ultimately not remembered, such as in rapid serial visual presentation (RSVP) scenarios [16].

B. Conceptualizing image memorability and object identity representations

Understanding the distinctions in visual representations between highly memorable images and those less so can be pivotal. A particular study examined this by monitoring neural activity in the inferior temporal cortex (ITC) of monkeys performing a visual memory task, similar to the one outlined in Box 1 [6]. The study found a significant link between the intensity of neural responses within the ITC and the memorability ratings of new images, with a Pearson correlation of 0.62. Notably, images that were more memorable initiated neural responses that were 20% greater than those less memorable, implying a visual appeal that made them stand out. Historical research into the ITC's response to natural scenes didn't fully acknowledge the range of neural responses, despite thorough investigations into visual representations in these areas [3, 4, 37]. This oversight might be due to the subtle, yet quantifiable, effects of variations in neural response strength, termed "magnitude coding" in the ITC [38, 39]. Such variations are thought to be counterbalanced by neural mechanisms ensuring stable overall neural firing rates, such as segmentation normalization [41] and homeostatic plasticity [40]. However, memorability research contradicts this by showing that the differences in neural response strength are not only significant (up to 20%) but also correlate closely with a behavioral aspect: the efficacy with which images are retained in memory.

C. How is image memorability quantified?

Image memorability quantification typically involves participants performing a visual recognition task via online platforms, like Amazon Mechanical Turk. This enables memory performance assessment across a broad set of participants, often 80 per image. During these tasks, participants view an image and report if it's a repeat from earlier in the sequence or a new one. Subsequent analysis has shown that recall scores are stable over time, with most studies examining the recall after 100 intervening trials between the initial viewing and a second viewing after four minutes. Memorability scores are derived from a participant's average ability to remember an image, taking into account the hit rate (HR) and adjusting for new image familiarity by using the false alarm rate (FAR). The correction for FAR is calculated for each image to address the variable tendency of images to seem familiar. The memorability score (MB) for any given image i is thus MB(i) = HR(i) - FAR(i). Adjustments to these scores may be made to balance time intervals between first-time and repeated image presentations. These recall scores are then normalized between 0 and 1, indicating the proportion of subjects recalling an image after first seeing it, despite being exposed to numerous images subsequently. This methodology has been employed for a wide array of images, as seen in the LaMem dataset, which has memorability data for 60,000 images. The distribution of these memorability scores is significant, often ranging from 1 to 0.2, with an average of 0.76. It's important to note that obtaining reliable memorability scores doesn't require active participation in a memory task. These scores have been shown to be stable over periods from minutes to weeks. This method has also been adapted for use in animal studies, such as with rhesus monkeys, to investigate neural correlates of image memorability. An alternative measure for assessing recall is the d' statistic, which is the difference between the HR and FAR.

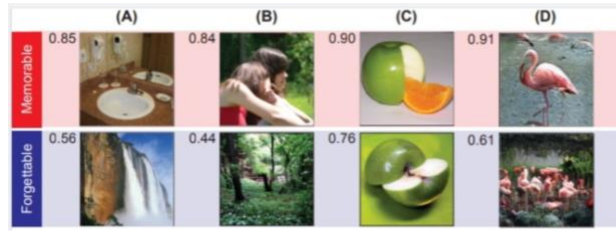

Figure 1. Quantifying image memorability.
A) Employing a visual recognition memory task for computing picture memorability scores. During each trial, participants assess whether the images are familiar or novel. Memorability scores for familiar pictures are then computed based on the subjects' average performance, adjusted for false alarms. B) Illustrating the memorability score distribution for the LaMem dataset, encompassing approximately 60 thousand images sourced from various origins [5].

D. Image memorability and Deep Artificial Neural Networks

Deep neural networks have played a pivotal role in recent advancements in comprehending image memorability, contributing in two distinct conceptual ways. Primarily, they have been used as engineering tools for memorably manipulating images and quantifying them. As an illustration, a deep neural network named MemNet was created to predict memorability scores for diverse images [5]. This tool has been employed as a substitute for image memorability scores in studies investigating the neural correlates of memorability [6]. Additionally, the GANalyze multi-component neural network was developed to generate new images and input images with content similar yet more memorable than the manipulated image [24]. Insights into the features influencing memorability are derived from this tool. Secondly, deep neural networks have served as models to understand how memorability is processed neurally in the brain. These models complement findings that deep neural networks, trained to categorize objects by associating images with object category labels, exhibit a functional organization highly akin to the form processing pathway observed in non-human primates and human brains [3, 4]. Remarkably, variations in image memorability are also evident in deep neural networks trained for object classification [6].

E. Visual memory storage and visual familiarity signaling

The processing of novelty and the storage of visual memory, primarily in the inferior temporal cortex (ITC) and the perirhinal cortex, is associated with signaling related to the familiarity of an image. Within these structures, familiarity manifests as repetition suppression, which involves adaptation-like reductions in population response following the viewing of a novel image [22–23]. Repetition suppression, potentially aligning with picture memorability, operates multiplicatively [17–18], The phenomenon becomes more evident when images that are easily remembered are shown repeatedly, eliciting stronger neural responses initially, which are then diminished with further repetitions. For example, the same rate of reduction applied to different initial values (say, a 10% decrease) will have a more significant absolute impact on larger numbers than smaller ones (for instance, reducing 100 by 10 versus reducing 10 by 1). Yet, it's not entirely clear if this pattern of response diminishment in the inferotemporal cortex (ITC) and perirhinal cortex is a full explanation for how images are remembered or forgotten. Additional research points to the role of the hippocampus in supporting the cognitive process of visual recognition memory, particularly when individuals assess the familiarity of an image that is similar to one they have seen before. This is assisted by the hippocampus through a mechanism referred to as pattern separation. Furthermore, research indicates that areas beyond the medial temporal lobe, such as the frontoparietal regions, contribute to visual memory by responding differently to images that are remembered as opposed to those that are not, highlighting a distinction in neural activity for such memories.

F. Impacts of local image in memorability

In this study [33], the researchers investigate the concept of image memorability and its impact on various fields, such as marketing, design, and photography. They aim to quantify image memorability and explore how local images influence it. To achieve this, they introduce new datasets with local images and create a visual memory game for quantifying image memorability. Using deep learning models (ResNet 50 and ResNet 101), they predict memorability scores. The findings reveal that local images significantly affect image memorability, emphasizing the role of image features and contextual meaning in image memory. The research introduces the SemMem dataset, which includes local images, highlighting the need for diverse datasets in understanding image memorability.

The study also discusses the methodology, participant demographics, and transparency in the research process. They collected data from Iranian adults, noting that contextual significance plays a crucial role in image memory. Furthermore, the researchers compare their SemMem dataset to existing ones, such as MemCat and LaMem, and demonstrate the use of ResNet architecture for predicting image memorability. Overall, this research provides valuable insights into image memorability and its practical implications for industries reliant on impactful visual content.

References

[1] S. Frintrop, E. Rome, and H. I. Christensen, (2010). "Computational visual attention systems and their cognitive foundations: A survey, "ACM Trans. Appl. Percept., vol. 7, no. 1, pp. 1–39, J.

[2] T. F. Brady, T. Konkle, G. A. Alvarez, and A. Oliva. (2008). Visual long-term memory has a massive storage capacity for object details. PNAS, 105(38):14325, 14329.

[3] U. Rimmele, L. Davachi, R. Petrov, S. Dougal, and E. A. Phelps, (2011). "Emotion enhances the subjective feeling of remembering, despite lower accuracy for contextual details", Psychology Association.

[4] P. Isola, D. Parikh, A. (2011). Torralba, and A. Oliva. Understanding the Intrinsic Memorability of Images. In NIPS 24.

[5] Khosla, A., et al. (2015) Understanding and predicting image memorability at a large scale. In International Conference on Computer Vision (ICCV).

[6] P. Isola, D. Parikh, A. Torralba, and A. (2011). Oliva. Understanding the Intrinsic Memorability of Images. In NIPS 24.

[7] Isola, P., et al. (2014) What Makes a Photograph Memorable? IEEE Trans Pattern Anal Mach Intell 36, 1469-1482.

[8] Bainbridge, W.A., et al. (2013) The intrinsic memorability of face photographs. J Exp Psychol Gen 142, 1323-1334.

[9] P. Isola, J. Xiao, A. Torralba, and A.(2011). Oliva. What makes an image memorable? In CVPR, IEEE Intl. Conf. on.

[10] Wichmann, F.A., et al. (2002) The contributions of color to recognition memory for natural scenes. J Exp Psychol Learn Mem Cogn 28, 509-520.

[11] Bartlett, J.C., et al. (1984) Typicality and familiarity of faces. Mem Cognit 12, 219-228.

[12] S. Lazebnik, C. Schmid, and J. (2006). Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In CVPR, IEEE Intl. Conf. on, volume 2, pages 2169 – 2178.

[13] Vokey, J.R. and Read, J.D. (1992) Familiarity, memorability, and the effect of typicality on the recognition of faces. Mem Cognit 20, 291-302.

[14] Z. Bylinskii, P. Isola, C. Bainbridge, A. (2015). Torralbaa, and A. Oliva, "Intrinsic and extrinsic effects on image memorability", Vision.

[15] P. Isola, J. Xiao, D. Parikh, A. Torralba, and A. Oliva, (2014). "What makes a photograph memorable?, " Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 36, no. 7, pp. 1469–1482.

[16] Aditya Khosla, Akhil S. Raju, Antonio Torralba, and Aude Oliva, (2015). "Understanding and predicting image memorability at a large scale", in International Conference on Computer Vision (ICCV).

[17] W. A. Bainbridge, P. Isola, and A. Oliva, (2013). "The intrinsic memorability of face photographs., "Journal of experimental psychology. General, pp. 1323–34.

[18] A. B. Michelle, A. A. Vo, Z. Bylinskii, P. Isola, S. Sunkavalli, A. Oliva, and H. Pfister, (2013). "What makes a visualization memorable?, "IEEE Transactions on Visualization & Computer Graphics, vol. 19, no. 12, pp. 2306–2315.

[19] Valsecchi, M. and Gegenfurtner, K.R. (2012) On the contribution of binocular disparity to the long-term memory for natural scenes. PLoS One 7, e49947.

[20] Huebner, G.M. and Gegenfurtner, K.R. (2012) Conceptual and visual features contribute to visual memory for natural images. PLoS One 7, e37575.

[21] Konkle, T., et al. (2010) Conceptual distinctiveness supports detailed visual long-term memory for real-world objects. J Exp Psychol Gen 139, 558-578.

[22] Vogt, S. and Magnussen, S. (2007) Long-term memory for 400 pictures on a common theme. Exp Psychol 54, 298-303.

[23] Wiseman, S. and Neisser, U. (1974) Perceptual organization as a determinant of visual recognition memory. Am J Psychol 87, 675-681.

[24] Brady, T.F., et al. (2019) The Role of Meaning in Visual Memory: Face-Selective Brain Activity Predicts Memory for Ambiguous Face Stimuli. J Neurosci 39, 1100-1108.

[25] Saleh, B., et al. (2013) Object-Centric Anomaly Detection by Attribute-Based Reasoning. In Conference on Computer Vision and Pattern Recognition (CVPR).

[26] Dubey, R., et al. (2015) What makes an object memorable? In IEEE International Conference on Computer Vision (ICCV).

[27] Yonelinas, A.P. (2002) The nature of recollection and familiarity: A review of 30 years of research. J Mem Lang 46, 441-517.

[28] Bainbridge, W.A., et al. (2019) Drawings of real-world scenes during free recall reveal detailed object and spatial information in memory. Nat Commun 10, 5.

[29] Judd, T., et al. (2009) Learning to predict where humans look. In IEEE 12th International Conference on Computer Vision (ICCV).

[30] Mancas, M. and Le Meur, O. (2013) Memorability of natural scenes: The role of attention. In IEEE International Conference on Image Processing.

[31] Bylinskii, Z., et al. (2015) Intrinsic and extrinsic effects on image memorability. Vision Res 116, 165-178.

[32] Celikkale, B., et al. (2013) Visual Attention-Driven Spatial Pooling for Image Memorability. In 2013 IEEE Conference on Computer Vision and Pattern Recognition Workshops.

[33] A.shokri, F.yaghmaee, (2020) Quantifying Image Memorability in Adults, M.Sc. Thesis in Artificial Intelligence and robotics, Faculty of Electrical & Computer Engineering, Semnan University.