

The Computational Memorability of Iconic Images

Lisa Saleh, Nanne van Noord*

Multimedia Analytics Lab, University of Amsterdam

Abstract

The perception of historic events is frequently shaped by specific images that have been ascribed an iconic status. These images are widely reproduced and recognised and can therefore be considered memorable. A question that arises given such images is whether the memorability of iconic images is intrinsic or whether it is shaped. In this work we analyse the memorability of iconic images by means of computational techniques that are specifically designed to measure the intrinsic memorability of images. To judge whether iconic images are inherently more memorable we establish two baselines based on datasets of diverse imagery and of newspaper imagery. Our findings show that iconic images are not more memorable than modern day newspaper imagery or when compared to a diverse set of everyday images. In fact, by and large many of the iconic images analysed score on the low end of the memorability spectrum. Additionally, we explore the variation in memorability of reproductions of iconic images and find that certain images have been edited resulting in higher memorability scores, but that the images by and large are reproduced with memorability close to the original.

Keywords

Memorability, Iconicity, Computer Vision

1. Introduction

The need to capture historic events in a visual frame has been around long enough to have captured events from centuries ago [18]. From ancient greek wall paintings of the battle of Marathon (490 BC) to the Black Death (14th century) in elaborate oil paintings [24, 6]. Historic events are often remembered through the visual imagery it is captured in [18]. Due to technological development in recent history, it has become conventional to capture events with photography. A unique share of historic photographs displayed in media are considered iconic. To illustrate, Tank Man, the image of a man in front of the tanks in a street of Tienanmen Square in Beijing, is an image which many people immediately can recall. Iconic photographs can be defined as “photographs that are widely recalled and recognized by individuals across social groups and generations, that they connect with specific historical events, and that have emotional significance for them and their national community” [7]. Iconic photographs are thus inherently connected with historical events and may influence the views of people on this historic event. For example, the photograph of Alan Kurdi during the Syrian refugee crisis in 2015 caused such an emotional response with most individuals that would influence the international politics surrounding the Syrian refugee crisis [1]. Thus iconic images may even influence

CHR 2022: Computational Humanities Research Conference, December 12 – 14, 2022, Antwerp, Belgium

*Corresponding author.

✉ lisa.saleh@student.uva.nl (L. Saleh); n.j.e.vannoord@uva.nl (N. van Noord)

🆔 0000-0002-5145-3603 (N. van Noord)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

the course of history. Due to this influence iconic images have been widely researched with the aim to understand what makes them unique.

Different aspects of iconic images have been researched. One often noted criteria for iconicity is the importance of the captured historic event. Due to the democratisation of photography, images of recent historic events are abundant. Thus an image that merely displays a significant historic event is not sufficient to reach iconic status. For an image to become iconic it must be widely spread in the media [11]. Therefore it is useful to know what images the media tends to publish and which characteristics they carry. For example, one study on media portrayal of photographs of hurricane Katrina uncovered that photographs with some visual themes are more often published than others [8]. Besides visual themes, there are plenty of other aspects that can influence if an image can be iconic. To recall the definition of an iconic photograph, the photograph must be widely recalled and recognized. An image that is widely recalled and recognized needs either to be really memorable or to be displayed very often. A more memorable image is more likely to be recalled. Thus high memorability of an image may contribute to its iconic status. As there has been no research on the intrinsic memorability of iconic images, this will be the focus of this paper. Specifically, we investigate whether there are general patterns concerning memorability of iconic images that are salient even when considering the unique circumstances by which these images have become iconic.

Memorability of images is a complex characteristic to predict; Using only the surface features of the image for predicting memorability is not comprehensive. Automating the prediction of memorability with Convolutional Neural Networks (CNN) has been a successful solution to this problem [16]. It is proven that features such as object size and brightness do affect image memorability [10]. However, a CNN like MemNet outperforms these surface features in predicting memorability [16]. When considering image memorability what a method like MemNet measures is the intrinsic memorability of an image, which can be interpreted as the innate probability of an image to be remembered. It does not guarantee that an image will be remembered, and it is certainly not the only factor, but all other factors being equal an image with a higher intrinsic memorability is more likely to be remembered.

Besides the advantages of automatizing the process by using a CNN, the usage of a prediction CNN has another advantage: As iconic photographs are images that are widely spread it is likely that a great share of people has already seen said image. Therefore it would be difficult to test the memorability of these images with a memory task. It would be a challenge to guarantee that these people have never seen these images before. Thus the challenge of testing the memorability of iconic photographs context-independent can be solved by using the above-mentioned MemNet. Additionally, the iconicity of some images is determined by the fame of the subject. MemNet is not trained to recognize celebrities or other symbols. In this research, the focus is on what image aspects the memorability influenced, without analysis of the context of the images. This makes the usage of MemNet a great fit. In this paper, iconic photographs are tested on their memorability by the use of the CNN MemNet to explore to what extent the memorability of iconic photographs can be measured.

To investigate this we formulate the following two research questions:

RQ 1: How does the memorability of iconic images compare to a baseline?

RQ 2: To what extent does the memorability of iconic images differ across variations?

2. Related work

2.1. Iconic Images

There has been little to no computational research on iconic images, but in the following we will give a brief overview of other research directions. As iconic images are not a bound set, different definitions and what criteria they should meet have been proposed [7, 17, 2, 11]. There are differences in the criteria brought up, but the following six proposed by Perlmutter [22] are used most often: "(1) significance of the reported event; (2) capacity to represent the event as a whole; (3) celebrity of the image promoted by the media; (4) prominence of display of the image; (5) frequent repetition of the image across media outlets; (6) ability to generate a primordial theme in society such as good versus evil." These criteria are not met to the same extent for every iconic image; it is not a blueprint for iconic images. For example, both the image of the Falling Man of the Twin Towers and the hijacked plane crashing into the Twin Towers represent the September 11 attacks. The image of the Falling Man has a less capacity to represent the event as a whole than the image of the airplane, as the attack itself is displayed more comprehensively in the latter image. But even though the Falling Man image does not meet the second criteria to the full extent, it is still an image to be considered iconic.

Additionally, other studies have added criteria to be more comprehensive. For example, in some studies more stress has been placed on the symbolic meaning that people have of the image [7]. An image can become a symbol with a meaning beyond the historical image it is attached to. The Guerrillero Heroico image of Che Guevara is used as a symbol for revolution beyond in Cuba [20]. Even though this is not significant for every iconic image, it is still a unique aspect for some. Furthermore, the criteria from Perlmutter [22] put the focus on the image and less on the reception of the image. One aspect of the reception that is unmentioned in these six criteria is that an iconic image also should be widely recalled and part of the collective memory of a certain with a group's identity [7]. This is another criterion often added to complement the criteria from Perlmutter. That an image must be widely recalled is important to add because this can be a way of researching what image have an iconic status with qualitative research as done by for example Hoeven [12].

But as other factors of the reception, like the symbolic meaning or being a part of the collective memory, of an iconic image are not fully encoded in the image itself and are subjective, it is yet challenging to study computational [19]. It is yet difficult to study the semantics of images computationally, as this often is can not be conducted from the visual data or metadata. Therefore the study of the perception of iconic images is often limited to interviews or large-scale surveys [19]. Hence the study of iconic images in computer vision is yet to be immersed, and potentially the computational study of memorability can open this door.

2.2. Memorability

Due to technological advancements capturing and sharing have become more convenient. Mass media and its consumption lead people to be exposed to countless images every day. People are exposed to many images daily and some of those are better remembered than others. Which image people remember and which they forget is related to the memorability of the image. Different academic fields have an interest in image memorability. For example,

research in cognitive sciences answered how certain activities in the brain correlate to the memorability of images. Additionally, image memorability has been a subject of research in the computer vision field in recent years [16, 10, 9, 14, 4]. Taking the next step from computing how memorable images are with MemNet, to analyzing what image attributes leads to memorable image and what connection memorability has with other qualities like the emotion it portrays. In line with researching what attributes influence memorability, GANalyze was created [10]. With GANalyze images are tweaked to increase memorability. Some attributes emerged to often be increased when optimizing the memorability, such as redness, brightness and object size [10]. Thus images that have higher redness, brightness or a bigger object size are more likely to be memorable. Even though these attributes might be simple and intuitive ways to predict memorability, the prediction of memorability is complex and can not solely be explained by these attributes [16].

When diving further into the correlation between memorability and some image aspects, collected across different studies, the following observations were made. In some studies, it appeared that a number of object categories, like people, animals and vehicles were relatively more memorable [21]. Images with humans are especially memorable when the face is visible and has eye contact with the camera [13]. In scenic images, indoor scenes appeared to be relatively more memorable [13]. Additionally, it appeared that spatial layout is highly correlated with memorability; Images where the object is bigger and centred in the image are more likely to be memorable [21]. This is in line with research which showed that images with a high aesthetic level are also more likely to be memorable [15]; the spatial layout of an image is strongly connected to its aesthetics. Moreover, it also appeared that images in which humans portray negative emotions like disgust and fear, tend to be recalled better [16]. Finally, only the most memorable images have a correlation with popularity [16]. Even though mapping for these above attributes and many more, there is still about 25% of the variance in memorability that is unaccounted for [23]. Thus looking for the attributes in images that are correlated with memorability will give some intuitive sense of memorability, but will not be comprehensive.

2.3. MemNet & LaMem

In this paper, memorability will be measured computationally through a score given by MemNet. MemNet is a CNN trained on an annotated dataset called LaMem. LaMem is a large-scale dataset constructed of multiple already existing datasets that were annotated in a perception study with human subjects who performed a memorability task. The datasets used are highly diverse, which leads to LaMem being a diverse dataset totalling 60.000 images [16]. It contains images of humans, animals, landscapes, and even art, which also includes abstract images. This diversity of image types is displayed in Figure 1.

For the LaMem annotation task a stream of images was shown to participants and after a varying number of distractors the participants had to indicate if they had seen the image before. Thus when an image was remembered often the image would be scored as more memorable. After performing this task on a group of participants a memorability score for each image could be calculated. This score is the annotation for each image in the LaMem dataset. This makes LaMem thereby a sufficient foundation for a CNN to compute a memorability score (0 – 1 scalar) for different types of images [16]. LaMem was shown to have an inter-annotator



Figure 1: Sample images from LaMem [16] with diversity in image subject. The images are arranged by their memorability score, decreasing from left to right.

rank correlation of 0.68, with MemNet achieving a remarkably high rank correlation of 0.64 [16, p. 6]. MemNet has since been further validated in subsequent studies and is, therefore, a reliable measure of intrinsic image memorability [23, 21, 14, 4].

3. Memorability of Iconic Images

The following step-by-step procedure, shown in Figure 2 represents our proposed method that can be used for measuring a memorability across a set of images.

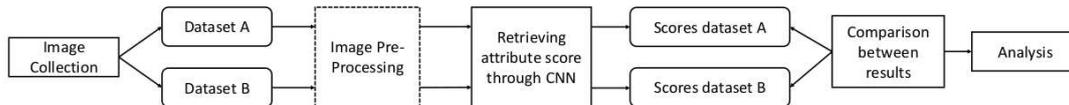


Figure 2: Flowchart overview of the memorability comparison procedure for two datasets.

The procedure we propose involves a comparative process between two datasets (A and B). Here, dataset A represents the collection to be studied, whereas dataset B functions as a baseline. The general idea of this approach is to not only measure the memorability scores but also judge whether the scores differ from a meaningful baseline. Each dataset is pre-processed by transforming the images to a suitable input resolution for the CNN (typically a low resolution like 256×256) and normalizing them. Once the dataset are pre-processed the images will be passed through the pre-trained CNN to obtain a per-image memorability score.

The method for comparing between the memorability scores of datasets A and B depends on the sizes of the respective datasets. In the case of small datasets, the analysis will be mostly done by qualitatively looking at the images to draw conclusions based on the specific images. When dealing with large-scale datasets, the comparison can be based on a statistical analysis. Whilst our baseline datasets are large-scale, the dataset of iconic images used is too small scale to reliably perform statistical analysis.

Further analysis beyond the comparative analysis will consist of sampling images which are

either random or statically interesting, according to their attribute score. Sampling random images will give an indicative perspective on how the data is constructed. This method of analysis is based on Distant Viewing [3], a method for studying large visual corpora, where the main take way is to perform the computational analysis whilst also viewing the corpora. The latter is import because the intuitive semantics that may get lost in a computational analysis will thereby also be studied. As the semantics are important for iconic images, this method suits this paper well. Conclusions drawn from these analyses can then be linked to prior work.

3.1. Comparing the Memorability of Iconic Images

For our dataset of iconic images we use a selection of 26 images that were initially in a large-scale survey on iconic images [12]. Smits and Ros [25] further used this set and collected variations of the images in the form of online circulations. These 26 images are a part of a global visual memory: 'a limited set of images that people all over the world have seen and remembered' [12]. Each image and its scores are depicted in Figure 3. These images are in order from highest to lowest memorability score. When comparing the scores with the images we make the following observations. Firstly, it is notable is that the three images with the highest memorability score are similar. The images of Che Guevara, Sharbat Gula, and the Migrant Mother are all portraits where the face of the subject fills much of the image. Their expression is visible to the viewer as all three are looking in the (rough) direction of the camera. That these portraits have a high memorability score is in line with the prior work; Images with bigger object size, a face that has eye contact with the camera and humans as their main object are more likely to have a higher memorability score. Which makes it in the lines of expectation that images like the Holocaust survivors, Raising a flag over the Reichstag and Tank Man have a lower memorability score; These images have a small object size and either has no human faces or a high amount of different faces. On the whole we see a diverse range of memorability scores across the dataset.

3.2. Iconic Images compared to a Baseline

To place the memorability scores in context, we compare to two different baselines. The first baseline used is the LaMem dataset; This dataset consists of images with different subjects and themes, and functions as a diverse baseline representing a 'general image collection' [16]. This is an annotated dataset for memorability, but the annotations were not used; Images of the LaMem dataset were inserted in MemNet to create a memorability score generated with the same circumstances as the dataset it is compared with. The second baseline used was the GoodNews dataset, which consists of all kinds of images used in the New York Times from 1818 until 2019 [5]. These images vary from news photography to sports photography, to images from the cooking appendix and so on. The range of types of images is displayed in Figure 4. These images are mainly images that were made by professional photographers and selected by the editors of the paper. Thus these images often have some journalistic value and meet the aesthetic standards of the paper. Therefore the GoodNews dataset is used as a baseline to represent images depicted in media. As iconic images are widely published in the media, the GoodNews dataset will help us determine whether iconic images are particularly memorable

	Sharbat Gula 0.9120		Che Guevera 0.8894		Migrant Mother 0.8767
	Survivor of Hutu death camp 0.8290		The Falling Man 0.8190		Man on the Moon 0.8168
	Abu Ghraib prisoner 0.8098		Hindenburg Disaster 0.7940		Alan Kurdi 0.7897
	Mao Zedong 0.7767		Mohandas Ghandi 0.7725		Hijacked Airplane 0.7697
	Times Square Kiss 0.7610		Napalm Girl 0.7588		Viet Cong Execution 0.7533
	Burning Monk 0.7476		Vulture and the girl 0.7462		Situation Room 0.7447
	Assassination of Inejiro Anasuma 0.7248		Raising the flag on Iwo Jima 0.7173		Kent State Shootings 0.7151
	Spanish Soldier 0.6982		Coup in Chile 0.6981		Tank Man 0.6939
	Raising a Flag over the Reichstag 0.6796		Holocaust survivors 0.6449		

Figure 3: Iconic images sorted on their memorability score. Each image is captioned with the name it is known as and its memorability score. The photographer and year of origin can be found in Table 1.

when compared to other media images.



Figure 4: Samples from the GoodNews dataset, which contains diverse images from The New York Times [5].

The distribution of the memorability scores of the LaMem and GoodNews dataset is visualised in Figure 5. Additionally, the scores of the iconic images are plotted with a scatter plot at the bottom of the Figure. From the distributions we can observe that the GoodNews images are generally more memorable than the LaMem images. Additionally, the LaMem dataset has a wider distribution than the GoodNews dataset. There are fewer images in the GoodNews dataset with a very low memorability score, which might reflect that all images in GoodNews have been approved by an editor of the paper. It is very likely an editor would select a picture with at least some traits that correlate with memorability. For example, images that are selected by an editor often need to have either a clear main object, thus big object size or some aesthetic value. Taking into account scatter plot of the iconic images we observe that iconic images are generally (slightly) less memorably than the images depicted in the media.

To further clarify how the iconic images compare to both datasets we plot the distribution of the iconic images according to the quartiles of both LaMem and GoodNews in Figure 6. We can observe that most iconic images fall into the lower quartiles for both datasets, scoring slightly higher when compared to LaMem. As the LaMem dataset is representative of all kinds of images, it can be concluded that these iconic images are slightly less memorable images. When comparing the iconic images to the GoodNews dataset, it stands out that most of the iconic images have a memorability score that overlaps with the first quartile of the GoodNews dataset. The distribution of the iconic images in the quartiles of the GoodNews dataset is significantly shifted to the less memorable side. Thus iconic images are less memorable than images depicted in the media.

4. Memorability across Variations

Until now we have analysed the most canonical versions of the 26 iconic images. However, it is known that edits have been made in prominently published versions of iconic images. To explore whether these edits have been made (implicitly or explicitly) to boost the memorability of iconic images we further analyse the circulations collected by Smits and Ros [25]. For each of the 26 iconic images they used the Google Cloud Vision API to retrieve online circulations, thereby creating a dataset of 900k images. Per image the number of variations differs from the Che Guevara portrait having over 100k variations retrieved to the image of Mao Zedong and

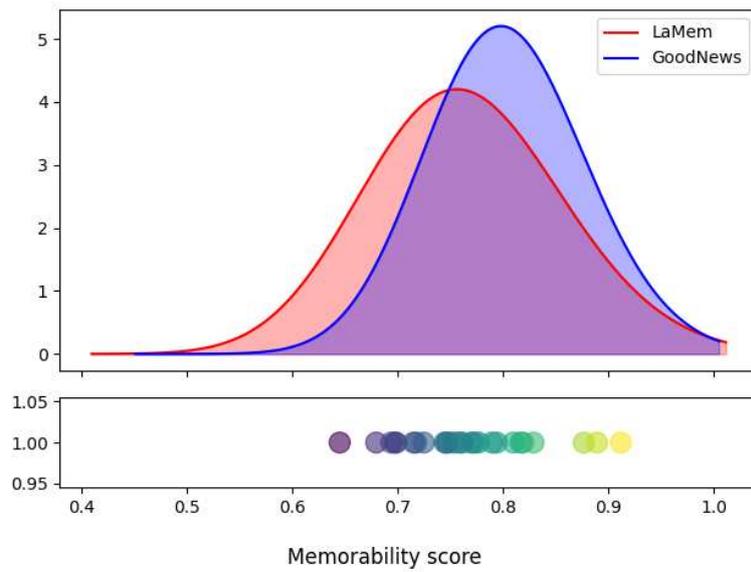


Figure 5: The upper graph is the distribution plot of the LaMem and GoodNews dataset, where the distribution is the density for the memorability scores. The graph below is a scatter plot of the iconic images and how they compare the datasets.

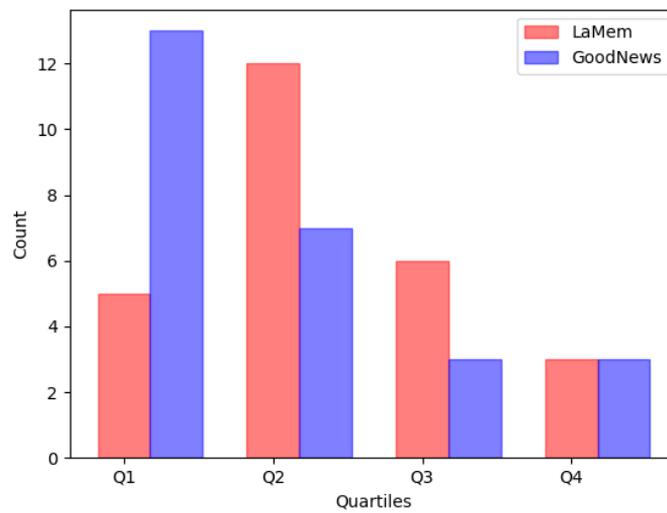


Figure 6: Distribution of the iconic images across the LaMem and GoodNews datasets. The y-axis represents how many iconic images have the memorability score for that quartile.

the founding of the PRC, having less than 3k variations. Figure 7 shows how many images were retrieved for each of the images.

Due to how the variations were collected, in iterations based on previous retrieval results, the

Table 1

Reproduction of table from Smits and Ros [25] with the amount of circulations per image.

known as	photographer	year	historical event	circulations
Migrant mother	Dorothea Lange	1936	Great Depression	41697
Falling Soldier	Robert Capa	1936	Spanish Civil War	18194
The Hindenburg Disaster	Sam Shere	1937	Zeppelin	36683
Times Square Kiss	Alfred Eisenstaedt	1945	V-Day	65164
Raising the Flag on Iwo Jima	Joe Rosenthal	1945	Pacific War	63249
Holocaust survivors	Lee Miller	1945	Holocaust	18343
Raising a Flag over the Reichstag	Yevgeny Khaldei	1945	World War II	90344
Gandhi and the Spinning Wheel	Margaret Bourke-White	1946	Mohandas Gandhi	10893
The Founding of the PRC	Hou Bo	1949	Mao Zedong	2865
Assassination of Inejiro Asanuma	Yasushi Nagao	1960	post-war Japan	3921
Guerillero heroico	Alberto Korda	1960	Che Guevara	108288
The Burning Monk	Malcom Browne	1963	Vietnam War	18122
Saigon Execution	Eddie Adams	1968	Vietnam War	18305
A Man on the Moon	Neil Armstrong	1969	Space Race	186921
Kent State Shootings	John Filo	1970	Kent State	7320
Accidental Napalm (Napalm girl)	Nick Ut	1972	Vietnam War	38619
Allende's Last Stand	Luis Orlando	1973	South-American Coups	6997
Afghan Girl	Steve McCurry	1984	Afghan War	47892
Tank Man	Jeff Widener	1989	Tiananmen Square Protest	63182
The vulture and the little girl	Kevin Carter	1993	Sudan famine	30121
Survivor of Hutu death camp	James Nachtwey	1994	Rwandan genocide	3395
The Falling Man	Richard Drew	2001	9/11	11681
Hijacked airplane	unknown	2001	9/11	6938
Abu Ghraib prisoner	Sergeant Ivan Frederick	2003	Iraq War	3601
The Situation Room	Pete Souza	2011	War on Terrorism	20102
Alan Kurdi	Nilüfer Demir	2015	Refugee crisis	24432
total				947269

differences to the original get progressively larger the deeper we get into the list of variations. The variations of the iconic image vary in different crops, different shades of colors and so on. There are also variations where other images and text are added to the iconic image, this can differ from a logo of a broadcaster to a book cover where the original image also appears. Furthermore, internet memes, collages and other types of photo-shopped images appear in the dataset. Additionally, the API is not infallible, there are some images in the dataset where the original image does not appear at all. Roughly speaking we observe that images which have been circulated less there are also fewer edits. To control for some of this variation, and to limit the scale of the comparison, we only use the first 10k variations for each iconic image.

The memorability scores for the variations of each iconic image are visualised in Figure 7. Images to the right on the x-axis should generally have less resemblance with the original image. When viewing the distributions of scores for the variations some patterns can be found. We recognise four groups: (1) distributions with relatively little spread across the variations, (2) distributions that fan outward towards the end of the plot, (3) distributions that have a large spread from start to finish, and (4) distributions where we can recognise clear clusters.

Examples of the first group, with little spread, at the Migrant Mother, the Situation Room, and the assassination of Anasuma. Within this group we can recognise two subcategories, as they

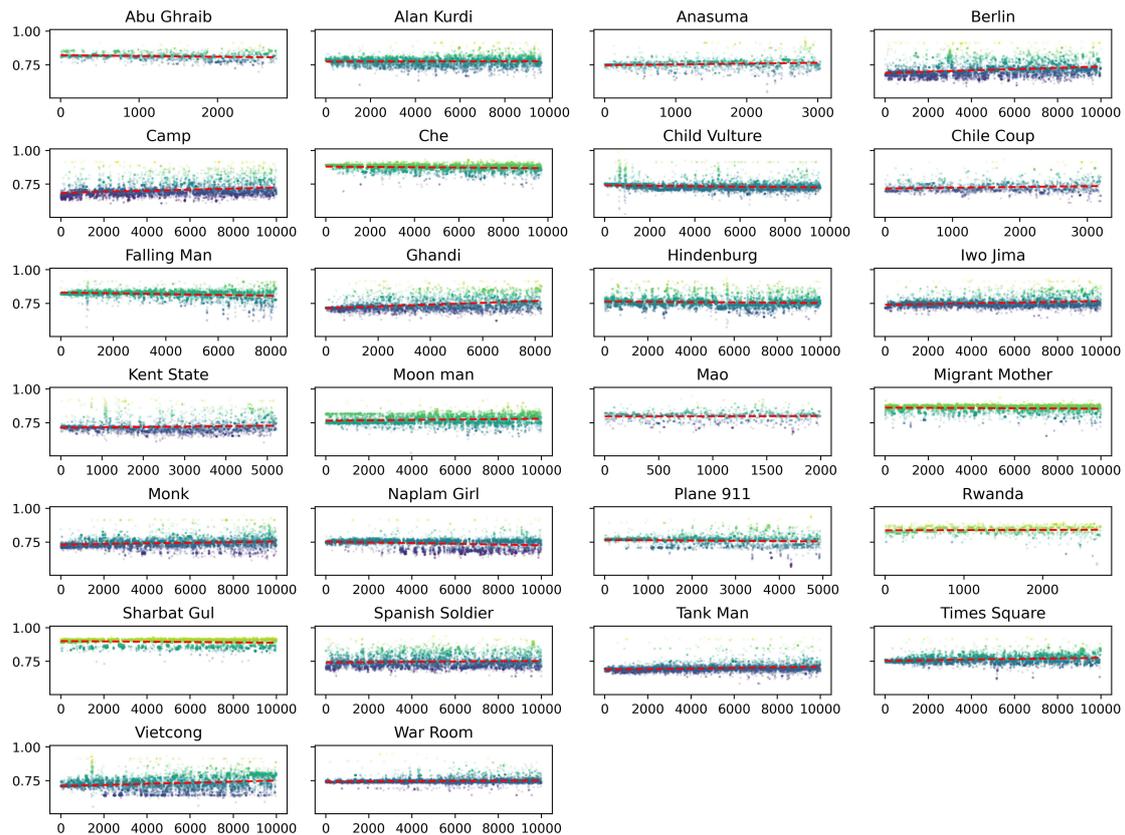


Figure 7: Grid with a scatter plot per iconic image, where each scatter plot visualises the memorability score across the variations in sequence. Higher x values are further from the original image.

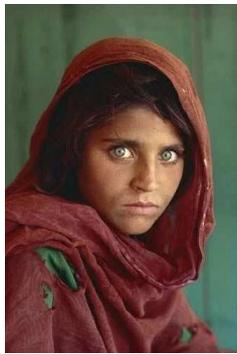
are either images that have less than 10k variations or they are images where the canonical form is most dominant. Because for the latter subcategory we have only looked at the first 10k images retrieved, it is still very likely that there are also many variations that differ more strongly circulating on the internet. But these would only be retrieved further in the dataset than the first 10k images selected. However, it is still noticeable that the first 10k variations have more resemblance with the original iconic image than for other widely circulated iconic images.

For the second group examples have most of their spread at the end of the graph, such as the Burning Monk, the Falling man and the Tank Man images. In these images, it appears that there is definitely a big share of images that have a big resemblance with the original picture but variations that differ more already appear within the first 10k variations.

In the third group there is a lot of spread from beginning to finish, these are images like Gandhi, Raising the flag over Iwo Jima and the Image of the Spanish Soldier. These images appear in a lot of different variations. This can be explained by it being popular photos for web pages, book covers and other types of editing where the context of the image gets altered heavily.

Lastly is the group where we can recognise different clusters. This group includes images

like Napalm Girl, the Hijacked Plane and the portrait of Sharbat Gula. All these images have an alternative popular variation, which leads to a cluster forming in the figure. In the following section, the Napalm Girl image will be highlighted to give an example of what these variations look like. Two variations of the Sharbat Gula image and the Hijacked plane are displayed in Figure 8 with their memorability score in the caption of each image. The variations that are highlighted in this figure are average images from the main clusters. In the Sharbat Gula variations graph, the bigger cluster on the top of the graph is the original image (Figure 8a) and the cluster that lies under most of the images are depicted in Figure 8b. In the Hijacked Plane, there is a similar pattern, the graph of the variations has two distinct clusters, one at the top which consists of images like the original image (Figure 8c and the other cluster which has a lower memorability score than the first cluster. This cluster consists of images like in Figure 8d, which has a wider crop than the original image such that the building on the left still remains in the image.



(a) The original image with a memorability score of 0.9120



(b) The variation of the original with a recent image placed alongside has a memorability score of 0.8511



(c) The published image with a memorability score of 0.7697



(d) A variation of the published image with a wider crop and a memorability score of 0.7095

Figure 8: Two variations of the Sharbat Gula and the Hijacked Plane image. Where the Figure 8a is the original portrait of Sharbat Gula and Figure 8b is a variation where a more recent portrait is placed alongside the original. The Figure 8c is the original published version of the Hijacked Plane and the wider crop variation in Figure 8d

4.1. The Memorability of the Variations of the Napalm Girl Image

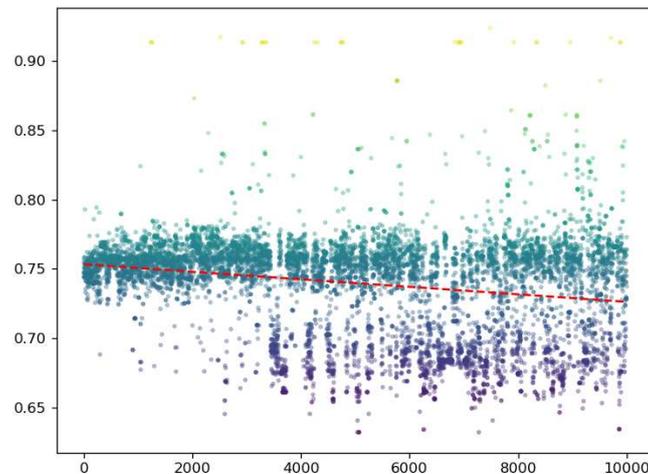


Figure 9: Scatter plot of the memorability scores on the y-axis and the sequence of variations of the Napalm Girl image on the x-axis, where the red dotted line is a trend line.

A number of different crops of the Napalm girl image circulate online. The scatter plot which displays the memorability scores and their variations is depicted in Figure 9. The outliers on the upper side of this plot, images with a score above 0.85, are either images heavily photoshopped or images that got in this dataset but in which the Napalm Girl image does not appear. One of the aspects that stands out from this graph is the big cluster of images which are under the trend line and appear from about 4.000 on the x-axis. When sampling these images it appears that these are mainly the Napalm Girl image with a tighter crop than the original image. This is displayed in Figure 10. Where the original image is displayed in Figure 10a and the tighter crop in Figure 10b. The part of the image that is only visible in the wider crop consists of a big part of the sky and a photographer on the right. The tighter crop is more focused on the children, which are the main subject of the picture, and thereby take up a larger portion of the image. The tighter crop having a higher memorability score fits the notion that images with a bigger object size are more memorable and hints that memorability might be an (implicit) criteria for edits done by photo-editors.

Another frequent variation was the image in Figure 10c. This image has a significantly higher score than the original picture and was the only variation with a score in this range that still fully depicted the original image. The original image was edited with a red-hued filter over the image. That this red-hued image has such a higher memorability score is expected as the redness of an image has been demonstrated to positively affect memorability.

5. Discussion

The findings in this paper are based on the result generated by MemNet. Even though MemNet is a validated method of predicting memorability, the actual memorability and the computed



(a) The original full photograph with a memorability of 0.6767



(b) Widely published tighter crop with a memorability of 0.7652



(c) A variation with a red filter and a memorability of 0.8200

Figure 10: Three different variations of Napalm Girl found in online circulation.

memorability score could still differ. In conducted research, it became apparent that image memorability and some qualities were correlated, like aesthetics and certain emotions. But even though they are correlated, MemNet is not trained on these qualities; Thus, for example, images that do not have the more common features that positively influence the memorability, like a big object size, but are very memorable because they portray a strong negative emotion that also influences memorability, possibly do not get a high memorability score from MemNet. This highlights a mismatch between intrinsic memorability and actually being remembered. A clear example of this mismatch is the Napalm Girl image, this image has a memorability score below the average of LaMem. But Napalm Girl displays a scene of terror, that evokes strong negative emotions like sadness and anger. Those strong emotions make this image more memorable and highlight a limitation of (computational) intrinsic memorability methods. Moreover, frequent exposure of less memorable images might also lead to increased remembrance, which might also play a role for this image.

A possible limitation of this work is that the dataset was selected by Dutch researchers. Despite being selected with the aim to represent international iconic images, the images are predominantly known in the Western World. This could influence the results, but this might

also interact with MemNet. As MemNet is trained on LaMem which consists mainly of images from the Western World, this bias is matched - whilst it should not influence the analysis itself it does limit to what extent we can generalize about the results. Additionally, the dataset of the iconic images is mostly from the 20th century when color photography was not as common. Most of the 26 images are in black and white. When experimenting with the differences in memorability score for the same image in greyscale as in color, it affected the memorability score; While the LaMem dataset averaged a memorability score of 0.7645 this score dropped to 0.7456 when all images were converted to greyscale. From this we can observe that colour plays a role, but not to the extent that it changes our conclusions. Even with taking these points into account, the observations we made align with existing theories on memorability.

6. Conclusion

In answering the research question: How does the memorability of iconic images compare to a baseline? It appeared that of the iconic images, the portraits where the face was a big part of the image had a higher memorability score. This confirmed the research on memorability where a correlation between higher memorability and big object size, and a face as an image object was established. Additionally, the iconic images with small object sizes and small faces had a lower memorability score. The memorability scores for these iconic images align with theories on memorability. Furthermore, it appeared that images that are depicted in the media are generally more memorable than all other images. Few images are depicted in media that have a relatively low memorability score since all images depicted in media are selected by editors. Iconic images are generally slightly less memorable than other images and are on the lower side of the memorability of images in the media.

To answer the second research question: To what extent does the memorability of iconic images differ across variations? When comparing the spread of the memorability of the variations of the iconic images, there were certain patterns to be found. Some had clear clusters, which were other popular variations of the original image. Looking at examples of different variations and memorability scores they show that altering an image can influence the memorability score. In these examples, it appeared that variations of the images with a red hue or tighter crop were more memorable. This is in line with previous work.

On the whole we can conclude that computational measures of memorability do not fully capture the memorability of iconic images, as many iconic images are remembered much better than what their memorability score would imply. While the reasons for this may be manifold we expect that frequent exposure and strong emotional content play an important role.

References

- [1] R. Adler-Nissen, K. E. Andersen, and L. Hansen. "Images, emotions, and international politics: the death of Alan Kurdi". In: *Review of International Studies* 46.1 (2020), pp. 75–95. DOI: 10.1017/s0260210519000317.

- [2] K. Andén-Papadopoulos. “The Abu Ghraib torture photographs: News frames, visual culture, and the power of images”. In: *Journalism* 9.1 (2008), pp. 5–30. DOI: 10.1177/1464884907084337.
- [3] T. Arnold and L. Tilton. “Distant viewing: analyzing large visual corpora”. In: *Digital Scholarship in the Humanities* 34.Supplement_1 (2019), pp. i3–i16. DOI: 10.1093/llc/fqz013.
- [4] Y. Baveye, R. Cohendet, M. Perreira Da Silva, and P. Le Callet. “Deep Learning for Image Memorability Prediction: The Emotional Bias”. In: *Proceedings of the 24th ACM International Conference on Multimedia*. Mm ’16. Amsterdam, The Netherlands: Association for Computing Machinery, 2016, pp. 491–495. DOI: 10.1145/2964284.2967269. URL: <https://doi.org/10.1145/2964284.2967269>.
- [5] A. F. Biten, L. Gómez, M. Rusiñol, and D. Karatzas. “Good News, Everyone! Context driven entity-aware captioning for news images”. In: *CoRR* abs/1904.01475 (2019). arXiv: 1904.01475.
- [6] P. Bruegel the Elder. *Triumph of Death*. 1563. URL: <https://www.museodelprado.es/en/the-collection/art-work/the-triumph-of-death/d3d82b0b-9bf2-4082-ab04-66ed53196ccc>.
- [7] A. A. Cohen, S. Boudana, and P. Frosh. “You Must Remember This: Iconic News Photographs and Collective Memory”. In: *Journal of Communication* 68.3 (2018), pp. 453–479. DOI: 10.1093/joc/jqy017.
- [8] N. S. Dahmen and A. Miller. “Redefining iconicity: A five-year study of visual themes of Hurricane Katrina”. In: *Visual Communication Quarterly* 19.1 (2012), pp. 4–19.
- [9] R. Dubey, J. Peterson, A. Khosla, M.-H. Yang, and B. Ghanem. “What Makes an Object Memorable?” In: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015, pp. 1089–1097. DOI: 10.1109/iccv.2015.130.
- [10] L. Goetschalckx, A. Andonian, A. Oliva, and P. Isola. “GANalyze: Toward Visual Definitions of Cognitive Image Properties”. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 2019.
- [11] R. Hariman and J. L. Lucaites. *No caption needed: Iconic photographs, public culture, and liberal democracy*. University of Chicago Press, 2007.
- [12] R. van der Hoeven. “The Global Visual Memory: A Study of the Recognition and Interpretation of Iconic and Historical Photographs”. PhD thesis. Universiteit Utrecht, 2019.
- [13] P. Isola, D. Parikh, A. Torralba, and A. Oliva. “Understanding the Intrinsic Memorability of Images”. In: *Advances in Neural Information Processing Systems*. Ed. by J. Shawe-Taylor, R. Zemel, P. Bartlett, F. Pereira, and K. Weinberger. Vol. 24. Curran Associates, Inc., 2011.
- [14] P. Jing, Y. Su, L. Nie, and H. Gu. “Predicting Image Memorability Through Adaptive Transfer Learning From External Sources”. In: *IEEE Transactions on Multimedia* 19.5 (2017), pp. 1050–1062. DOI: 10.1109/tmm.2016.2644866.
- [15] P. Jing, Y. Su, L. Nie, and H. Gu. “Predicting Image Memorability Through Adaptive Transfer Learning From External Sources”. In: *IEEE Transactions on Multimedia* 19.5 (2017), pp. 1050–1062. DOI: 10.1109/tmm.2016.2644866.

- [16] A. Khosla, A. S. Raju, A. Torralba, and A. Oliva. “Understanding and Predicting Image Memorability at a Large Scale”. In: *2015 IEEE International Conference on Computer Vision (ICCV)*. 2015, pp. 2390–2398. doi: 10.1109/iccv.2015.275.
- [17] M. Mortensen. “Constructing, confirming, and contesting icons: the Alan Kurdi imagery appropriated by #humanitywashedashore, Ai Weiwei, and Charlie Hebdo”. In: *Media, Culture & Society* 39.8 (2017), pp. 1142–1161. doi: 10.1177/0163443717725572.
- [18] M. Moss. *Toward the visualization of history: the past as image*. Lexington Books, 2008.
- [19] N. van Noord. “A survey of computational methods for iconic image analysis”. In: *Digital Scholarship in the Humanities* (2022). doi: 10.1093/llc/fqac003.
- [20] M. Parker. “The Retrospective Iconicity of ‘Guerrillero Heroico’”. In: *Salford Postgraduate Annual Research Conference*. 2009, p. 292.
- [21] S. Perera, A. Tal, and L. Zelnik-Manor. “Is Image Memorability Prediction Solved?” In: *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. Los Alamitos, CA, USA: IEEE Computer Society, 2019, pp. 800–808. doi: 10.1109/cvprw.2019.00108.
- [22] D. D. Perlmutter. *Photojournalism and foreign policy : icons of outrage in international crises*. Praeger series in political communication. Westport, CO [etc: Praeger, 1998.
- [23] N. C. Rust and V. Mehrpour. “Understanding Image Memorability”. In: *Trends in Cognitive Sciences* 24.7 (2020), pp. 557–568. doi: <https://doi.org/10.1016/j.tics.2020.04.001>.
- [24] *Sarcophagus with the Battle of Marathon*. 490 Bc. URL: <https://www.livius.org/pictures/italy/brescia-brixia/marathon-relief/>.
- [25] T. Smits and R. Ros. “Quantifying Iconicity in 940K Online Circulations of 26 Iconic Photographs”. In: *Computational Humanities Research*. 2020.