Screenshots, Symbols, and Personal Thoughts: The Role of Instagram for Social Activism

Ehsan-Ul Haq
euhaq@connect.ust.hk
The Hong Kong University of Science
and Technology
HKSAR

Tristan Braud
braudt@ust.hk
The Hong Kong University of Science
and Technology
HKSAR

Yui-Pan Yau
arthur.yau@connect.ust.hk
The Hong Kong University of Science
and Technology
HKSAR

Lik-Hang Lee[*]
likhang.lee@kaist.ac.kr
Korea Advanced Institute of Science
and Technology
South Korea

Franziska B. Keller
fbkeller@ust.hk
The Hong Kong University of Science
and Technology
HKSAR

Pan Hui[†]
panhui@ust.hk
The Hong Kong University of Science
and Technology
HKSAR

ABSTRACT

In this paper, we highlight the use of Instagram for social activism, taking 2019 Hong Kong protests as a case study. Instagram focuses on image content and provides users with few features to share or repost, limiting information propagation. Nevertheless, users who are politically active offline also share their activism on Instagram. We first evaluate the effect of protests on social media activity for protesters and non-protesters over two significant protests. Protesters' exposure to protest-related posts is much higher than non-protesters, and their network activity follows the protest schedule. They are also much more active on posts related to the protest that they participate in than the other protest. We then analyze the images posted by the users. Users predominantly use symbols related to protests and share personal thoughts on its primary actors. Users primarily share content to raise their network's awareness, and the content choice is directly affected by Instagram's intrinsic interaction modalities.

CCS CONCEPTS

• **Networks** → **Social media networks**; • **Human-centered computing** → *User studies*; *Web-based interaction*; **Social content sharing**; **Social media**.

KEYWORDS

Social Activism, Online Social Network, Instagram, Protest

1 INTRODUCTION

Societies around the world have embraced social media as one of the primary communication channels [42]. In particular, Online Social Networks (OSNs) such as Facebook and Twitter have been used for "social activism"; the most recent example touted as a successful mobilisation through social media is the 'Arab Spring' [48]. Research on OSNs modalities and their role in social uprising and activism is an important area for cross-disciplinary research with a direct and long-lasting impact on society. Such impact directly falls

under the UN Sustainability Goals [1] to create opportunity for communication that can lead towards an inclusive society [49]. OSNs have proven at times their role in promoting activism in society by offering interaction methods and a platform to express support in sensitive issues such as Black Lives Matter (BLM) or Women March in the US [11].

Studies on traditional OSNs only examine online mobilization and provide no systematic evidence that users who retweet, like, or otherwise engage with online content also participate in offline political activism. Furthermore, they usually focus on Facebook and Twitter and ignore platforms that have recently become more popular, particularly in young demographics. In this paper, we examine the role of Instagram - one of the most commonly used social media platforms for the Hong Kong youth [39] - in offline protest participation. In addition, we analyze Instagram features that are utilized for social activism. To the best of our knowledge, this is the first work examining Instagram's role in protest participation and linking it to related offline behaviour. The results from this study show the role of Instagram in social activism and characterize the way users share the content on this platform.

This paper looks at whether Instagram is a viable tool for social movement, despite its significantly limited communication affordance(s). We first characterize Instagram modalities that can affect social activism on the platform. We then consider two early major demonstrations from the 2019 Hong Kong protests as our case study (See Appendix for Hong Kong Protest background and timeline). We divide our analysis into two parts: First, we use social activism and exposure score combined with the network-based approach to show that social activism on Instagram also correlates with offline protests and physical presence in protests. Second, we analyze the image content to observe the focus of discourse and find any unique tactics related to the Instagram interaction features. We observe that politically interested users manage to leverage the Instagram platform to increase their network's exposure to the ongoing social movement - and that this impact extends beyond online behavior. In order to demonstrate the link between what users post online, what they see in their feed, and their offline political engagement, we collect and study Instagram data generated between June 1 and June 20 2019, but focus on two significant protests occurring on

---
[*]Lik-Hang Lee has second affiliation with University of Oulu, Finland
[†]Pan Hui has second affiliation with University of Helsinki, Finland

---
[1]https://sdgs.un.org/

June 9 and June 16. After identifying users present at the protests in our dataset, we use the hashtags and captions of posts as linguistic parameters to distinguish between protest-related and non-protest related posts. We construct an (online) social activism score based on the volume of protest-related posts in a user's feed. This score allows us to evaluate a user's online political engagement. We show that protest participants have a higher social activism score overall but are particularly active on the day of protest. We formally quantify the user's interaction with their network and correlate it to the user's online political engagement and real-life participation in social activism. We finally analyze the content of 1,000 images randomly selected in our dataset to evaluate the common themes. We find that politically active users circumvent Instagram's limited interaction modalities by posting highly symbolic content and text or slogans in image form to spread awareness within their direct network.

2 LITERATURE REVIEW

As mentioned in the introduction, there is vast literature on the role of social media and protest mobilization. More importantly, how social media communication modalities such as video streaming, retweeting, or the use of hashtags helps in social activism [1, 21, 23, 46]. Ahmed et al., for example, highlight the role social media played during protests in Bangladesh in 2018 [1], and find that people used Facebook Live streaming for communication and as a reliable resource. A study on Egypt's Tahrir Square protests in 2011 also provides evidence for the effectiveness of communication via Facebook and links this to offline participation in protests [45]. And a study on protest participation and the Russian social media platform VK [2] shows a positive impact of social media on protest participation because it reduces the costs of collective action [12]. In the United States, Peng et al. analyze the online activism movement for Black Live Matters and show that online user behaviour is closely related to offline events: They find that discussions started shortly after instances of police violence persist longer than communications during normal times [33]. And for virtual marches, Li et al. show that identification with the rest of the protesters is a significant factor to join such protests [26]. However, social media's capacity to help mobilize or sustain a movement is not unlimited: A study [20] on a group of people with age ranging between 22 to 29 shows that social media merely provide a communication platform for already politically motivated users, but does not draw less motivated people into such discussions.

Another kind of social movement is leaderless protest, and Hong Kong protests are classified as same [21]. Other online communication methods such as messaging apps are also effective during social movements. Hirsch and Henry discuss the use of text message tools during the swarm-style protests and find such tools effective for diffusing and updating information [17]. Other research shows the effectiveness of technology during the protests more generally [47]. Finally, several papers have used social media to predict user' participation and future unrest in societies [29, 36, 37].

Several studies have focused on the role of social media in Hong Kong protests [2, 21, 35]. The usage of messaging applications in Hong Kong protests is also attributed to presumed data security and

privacy [2]. In contrast, Instagram, as a medium for social activism is getting attention [14, 23, 28]. However, there are no studies on Instagram for social activism or protest participation based on data on interactions between users. Cornet et al. analyze graphic content from Instagram posts related to three social movements but only use this to identify five themes [10]. Another recent paper discusses the influence of an Instagram user on other users. However, without taking network connections into account, [40]. Finally, Trevisan et al. show that while likes are the preferred way of user interactions on Instagram, political accounts may get more comments than non-political accounts [42].

In this study, we focus on the interaction between protesters and non-protesters, and images with protest-related hashtags. To the best of our knowledge, this is one of the first studies of social activism on Instagram, and the first study analysing the Hong Kong protests of 2019 through a social media perspective.

3 MOTIVATION: INSTAGRAM AS A SOCIAL ACTIVISM PLATFORM?

Instagram is generally associated with young demographics across the world[19], making it an important platform to study for social activism. Additionally, Instagram is one of the most popular social networks among the 18-30-year-olds, which represent more than 60% of the people participating in the Hong Kong protests, but remains an understudied platform [24]. However, unlike other social media platforms, Instagram's features do not lend themselves easily to political and social activism. Earlier work on social media activism on other platforms have focused on multi-layered activity and message cascades such as retweet and mention networks [15]. However, Instagram lacks a direct proxy to such modalities. Instagram's interaction features also limit the opportunities to propagate content in the way retweets or shares do on Twitter or Facebook, respectively. This difference in information transmission has a visible effect on the type of content posted by users, such as sharing of screenshots as a workaround for the lack of sharing features [10]. As a result, Instagram significantly limits studying the propagation of information cascades, hindering its adequacy for social activism. Some characteristics of the platform that limit the spread and outreach of messages are as follows:

Image content at the core: Instagram primarily focuses on image content. The primary view of the interface only displays images, with an additional interaction cost to access the image caption or the comments. As such, the textual content is secondary to the image. Users may circumvent this issue by posting screenshots of text or adding the caption manually to the picture. This limitation makes it harder to share slogans or engage in more complex discussions of political and social issues.

(Re)Sharing: By design, Instagram offers limited post sharing capabilities. Users can share posts through time-limited stories or direct messages to other users. Sharing a post is one of the most used features in different platforms to show support or contradict other users' opinions. For instance, on Twitter, users can retweet (a twitter terminology to share content, with an option of adding their message to it or as it is) the content from other users. Retweet networks, based on sharing of content, have been widely studied and used in Computational Politics research [16], particularly for

polarization and stance detection problems. Sharing content on social media is also linked with higher visibility in terms of engagement as compared to just liking or commenting on it; sharing a post will make it appears as a post in the feeds of the user's network [5]. The absence of post sharing features significantly limits the propagation of messages on Instagram and the capacity to study a social movement on the platform.

Virality Measure: In social media communications, user-engagement is mainly studied as a combination of popularity (no. of likes), virality (no. of shares), and commitment (no. of comments) [6]. The empirical analysis for these interaction measures has been widely used in research studies focused on citizens' engagement in social media during social events, organizational, and governmental communications [5, 13]. However, on Instagram, there is no apparent measure of virality embedded in the posts compared to other social media. The previous work reports that people tend to comment less than like a post on Instagram. Such limitations and nuances in the design and users' practices make it challenging for researchers to analyze citizen engagement effectively.

Exposure Time: Instagram stories are short-lived by default, with seven-second visibility. As such, an Instagram user can only see the stories from another user in their network for seven seconds. This story will then move to the end of the list. This design choice can be credited with many positive factors, such as a round-robin and chronological display of stories that give users a fair exposure to the entirety of their network. However, the short visibility duration may also lead to lower exposure to specific content. To see an Instagram story for a second time, the viewer has to scroll to the end of the list or go directly to the original poster's profile, significantly increasing the interaction cost. Due to these factors, it is not easy to precisely quantify the impact of stories on the exposure.

Second Degree Interactions and Trending Topics: Unlike other networks such as Twitter, a user's timeline on Instagram only contains the posts directly shared by the user's network. Instagram users can not see the activities of the users that they follow, e.g. users can not see what other users liked or followed. Similarly, Instagram does not accumulate and show trending topics for the user. Such limitation can also curb the exposure to the users and may eventually affect social activism.

This paper shows that even though exposure to large scale social activities is limited on Instagram, users' network asserts enough exposure to correlate with their offline protest participation. We show how users divert a tool primarily designed for sharing photos into a social activism instrument. We also aim to highlight new methods to study social movements on a social network that prevents traditional techniques such as post sharing networks.

4 DATA COLLECTION

To understand why some individuals join a protest while others stay away, we compare protesters with non-protesters. Some explanations for the different behaviour may be relatively trivial: for instance, some non-protesters may not have heard about the protest, may not be politically interested, or are out of the country when the demonstration occurs. To focus on the explanations, we use the approach suggested by Larson et al. [22] to identify protesters and a group of non-protesters that could have attended, but chose

not to do so. As we consider two protests on June 9 and June 16 2019, we focus our study on content generated between June 1 and June 20. We collect data in a three-step process. First, we manually curate a list of hashtags related to the protest in Hong Kong in both Cantonese and English. The list of hashtags is accessible online. [3] We then recover the set of users who have posted at least one post with one of the hashtags, filtering out users who do not have any posts from the period under examination (1st to June 20) among their 50 most recent posts. We wrote our crawler to search for the posts for a given hashtag on Instagram. We extract the users based on these search results for all hashtags. Finally, we collect the 50 most recent posts of these users aware of and interested in the protest issue. We performed this data collection at the end of June 2019 to ensure posts related to the studied protests in our user set.

Larson et al. [22] use the complete 2-step neighbourhood of Twitter users in their approach. However, Instagram only allows for two types of public interaction: like and comment. Unlike Twitter, it does not offer a way to re-post another account's post as a new post (analogous to retweet). [4] We, therefore, only consider the users' one-step neighbourhood and the content shared there. Besides, we only focus on the accounts a given user follows, not her followers, as users only see posts from the former in their feed. Based on these steps, we can identify a set of users aware and interested in the issues at the heart of the protest and the other Instagram users whose posts are likely to show up on their feed, i.e. their network or the accounts they follow.

Both demonstrations involved hundreds of thousands of participants in a confined area (see Figure 4 in Appendix). We use dates as mentioned in timeline Table 5 (Appendix), the time during that day, and the location of the demonstrations to classify users as protest participants if their protest-related post is marked with the protest location at that time. We collect the location data with the posts if the user has tagged her location. We separate them from users who have posted during the same time frame but from a location elsewhere in Hong Kong and without any protest-related hashtags. This method helps avoid false positives, such as posting protest-related content outside the protest region. We call the latter users non-protesters. Table 1 provides an overview of the accounts collected. Note that the number of accounts is relatively small in our study. Like Larson et al. [22], we follow an approach that does not attempt to analyze a large number of users themselves but instead scours a large number of users to identify the users that allow us to make a meaningful comparison. In this case, three groups that are similar except for the outcome of interest: the participation in the first or the second demonstrations and non-protesters. We also note that our dataset of protesters is small as compared to social activism studies on other platforms such as Twitter. However, Instagram does not provide any official API to collect the users' timeline or network information. In addition, we also limit our hit rate to the

---

[3]https://docs.google.com/spreadsheets/d/1sNMhW3APf8WsILfKI4UL6uOLvu9hGj-HpLQ27rWaTaw
[4]Instagram has stories where a user can share a post from another Instagram user as their own story. Such stories are not visible in the timeline of other users unless they explicitly click on the given user's stories icon or the storyboard. Any further communication on these stories is in the form of private messages between users. The rest of the network cannot see these discussions, so we do not collect them as part of the direct exposure discussed below.

Table 1: Number of users and their network size. Collected Network shows the number of users present in our dataset

	Protesters	Network	Network (Collected)
P_a	49	27987	21580
P_b	102	57557	40848
$P_a \cap P_b$	15	5530	5530
Non Protesters	42	19984	10433

Instagram servers. These limitations result in slower data collection and network sampling, and hence a smaller protester sample size.

Research Ethics and Data Privacy: We understand that the data used in this research deals with sensitive personal information. To deal with this issue, we have immediately anonymized the data in 2019 during the data preprocessing and the raw data was deleted after that.

5 METHODOLOGY

We analyze the language and the first-degree connections of the protesters and non-protesters. We use the hashtags the text in the captions of the posts, as hashtags are one of the methods used to get the theme of the post [4, 8, 44]. Instagram users use hashtags to highlight what the posts are about, similarly to Twitter.

We divide the dataset into three different sets of users and their networks. As discussed in Appendix, Hong Kong Protests are a series of protests, but we only consider the two largest ones on June 9 and 16. From our dataset of identified protesters, we separate the protesters and their network into two groups, based on whether they attended the first and second protests. We name the set of protesters participating in the first protest as P_a and their network as P_{an}. For the second set of protesters and their network, we use the terms P_b and P_{bn}, respectively. Meanwhile, for non-protesters and their network, we use the terms NP and NP_n. The number of such users is given in Table 1. In our dataset, 15 users belong to both P_a and P_b.

5.1 Hypothesis Formulation

Our hypotheses are based on the concept of exposure, that is, the potential posts that users can see from their network [32]. Exposure to certain kinds of posts on other social media platforms is thought to affect users' decisions and actions, such as opposing contrarian views and enforcing users' own opinions in online communities [31]. But [31] also shows that political interests and participation leads to selective exposure.

Online social networks use ranking algorithms to determine their users' feed. Such algorithms are usually a black box to the public knowledge domain and tend to incorporate the network structure, i.e. when scrolling through their feed, users see posts by those they follow [41, 43]. Suppose the latter accounts share many protest-related posts. In this case, the user is exposed to protest-related information and, therefore, likely to be affected by such exposure. However, how likely they are to do so may also depend on how their network talks about the protest. We thus suggest the following hypotheses. First, we test if online activism (measured

as a user's fraction of posts that contain protest-related hashtags) even translates into offline protest participation (i.e. posting at a protest location):

H1. *The social activism score among protesters will be higher than the score among non-protesters from the day of the first protest to the end of our study.*

Protests tend to be contagious, and social media is often used to mobilize protest participants. We would, therefore, expect that protest participants are more exposed to protest-related posts through their network. We measure this **protest exposure** using the ratio of protest-related posts to non-protest related posts posted by the accounts that a user is following:

H2. *Through their network, protesters will have greater exposure to protest-related posts than non-protesters.*

Users may use social media to estimate how popular a cause is and look for additional clues, such as the number of likes. Instagram, just like Twitter, does not have any "down-vote" function. If users want to express their displeasure with a post, they either have to leave a comment or decide not to interact with the post. Therefore, a high ratio of comments to likes is often considered a sign of a controversial or unpopular post on Twitter. However, we argue that protest-related posts inherently trigger discussions, increasing the number of both positive and negative comments. Similarly, protesters might use comments on a post for coordination among themselves. We thus analyze the numbers of likes and comments separately - instead of calculating a comments-to-likes ratio - to evaluate the activity around protest-related posts.

H3. *Protesters will have higher activity around protest-related posts in their network than non-protesters, and the protest appraisal (number of likes) will be higher for protest-related posts.*

For H1, we use the hashtags and the related keywords in the captions to measure the score. For H2, we categorise the potentially visible posts to the users based on their networks.

5.2 User Interactions

For user interaction analysis, we look at the two interaction methods provided by Instagram: Like and Comment. For each set of users, we look at the number of comments and likes received by the protest-related posts and compare with the same for non-protest related posts. We do it separately for the seed users P_a, P_b, NP and their networks.

5.3 Social Activism

To quantify social activism on social media, we look at the posts by each user in the sets of protesters and non-protesters P_a, P_b, NP. In this case, we consider social activism as users posting anything with a protest-related hashtag (according to our hashtag list) on their Instagram account. Hence, we define the social activism score for each user as the ratio of protest-related posts (P_p) to the rest of the posts (P_{np}). For any user i, the social activism score will be
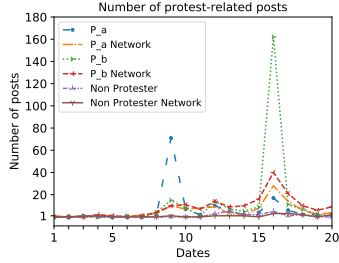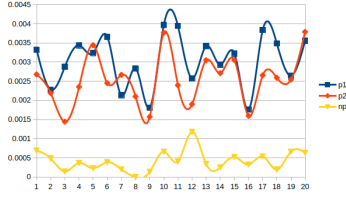
$$a_i = \frac{\sum p_p}{\sum p_{np}} \qquad (1)$$

For each set of users, the social activism score is taken as the average individual activism score of all users (n) within a specific user set.

Table 2: Social activism score based on posts hashtags. Protesters have higher score than non-protesters.

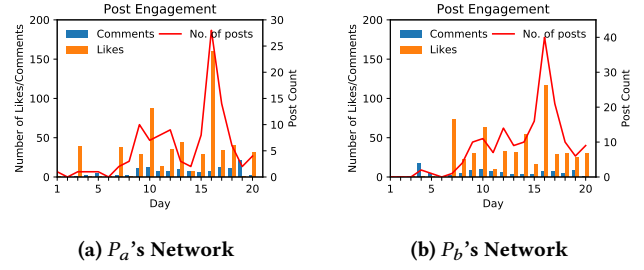| Group	Score	Standard Error
P_a	0.75	0.1
P_b	0.59	0.09
NP	0.05	0.02

Figure 1: Number of posts for protesters, non-protesters, and their networks. Protesters display a peak in their activity on the day they participated in a protest. Non protesters do not display a significant change in activity. The dates are given as the day in the month of June 2019.

Figure 2: Exposure comparison for both Protester groups and non-protesters. Protesters see more posts related to protests than protesters. The dates are given as the day in the month of June 2019. Y-axis shows average exposure.

$$Score_s = \frac{\sum a_i}{n}, a_i \in s_i \qquad (2)$$

Equation 1 gives the score to individual users and Equation 2 shows the score for complete set of users. Table 2 shows the results for this analysis in our dataset along with the standard error in measure. The result shows that P_a and P_b have higher activism scores than non-protesters and that the participants of the first protest - the avant-garde - have a higher score than those of the second protest.

5.4 Exposure

This section analyzes the type of posts the user in our main sets are potentially exposed to. Given our hypotheses, we expect such exposure to be higher for protesters than non-protesters. To do so, we look at the exposure from two different perspectives. Firstly, we take the standard definition of exposure as a contact between any

(a) P_a's Network

(b) P_b's Network

Figure 3: Average number of likes and comments for posts in P_a and P_b's network. Day wise analysis of posts for both set of protesters' network. The dates are given as the day in the month of June 2019.

two entities for a specific period of time (from initial time t_0 to final time t_f), and can be written as follows:

$$E = \int_{t_0}^{t_f} C(t)d(t) \qquad (3)$$

Where C is the function in time t that shows the relationship of interaction between subjects in a specific time interval. In our case, we primarily look at the posts in the per day unit. In our case, the function C would be the number of politically related posts out of the total number of posts that a user is exposed to. Hence, a simplified equation to measure the exposure of political content for a given user i on a particular day will be

$$e_i = \frac{\sum_j p_p}{\sum_j (p_p + p_{np})}, j \in N(i) \qquad (4)$$

Where p_p shows the political posts, p_{np} shows non political posts from any user j in the network $N(i)$ of user i. We show the results for this method in Figure 2 as an average of all users in the same network on a given day. In our second method for exposure, we also consider the interaction measures (likes and comments count) on the posts that users are likely to see. Segev et al. [40] suggest that, on Instagram, likes are more engaging than comments. However, a higher number of comments and likes can help increase the outreach of the post, and we have already reasoned that comments might be positive in this context as well. We, therefore, define both the average number of likes and comments as political engagement exposure but look at them separately in Figure 3.

6 FINDINGS

Protesters have higher social activity score. Table 2 shows the political scores for both the protesters set and the non-protesters set. During these 20 days, protesters have posted more content about protests. Protesters from the second demonstration display, on average, a slightly lower social activity score than protesters from the first demonstration. This phenomenon is likely because the second demonstration, reported to have more than 2 million people [5], must have drawn a more diverse population and mobilized not just the most politically active segment of Hong Kong society.

We also display the top 100 hashtags for protesters in Figure 6.a (Appendix) and 6.b, and for non-protesters in Figure 6.c. We represent in green the hashtags related to protests, and in red the hashtags not directly related to protests. The hashtags' size is directly proportional to their frequency in our dataset. Protesters feature a much higher number of protest-related hashtags in their posts than non-protesters. We notice an interesting linguistic feature: Non-protesters posting protest-related hashtags tend to use hashtags in English while protesters mix both English and Cantonese language hashtags. This phenomenon might be related to the non-protester set may encompass more international profiles than the protesters' sets. Indeed, even the hashtag "香港" (Hong Kong) does not appear in the top 10 of the hashtags on non-protesters' posts, while it is predominant in protesters' posts.

Protesters have higher exposure than non-protesters, and their network's political posts are more engaging than non-protesters' networks. We evaluate the exposure to protest-related posts of protesters and non-protesters by analysing the daywise post activity for the protestors and non-protesters and their respective networks. We first display the post-activity for all users in Figure 1. We notice two peaks around the protest days (June 9 and June 16) where the protesters have most of the protest-related posts. In terms of exposure, both protesters set are potentially exposed to a higher number of protest posts compared to non-protesters. We thus see a classic clustering effect: individuals with high social activism are following individuals with high social activism. What is less clear is how this comes about: are they attending the protests because of what appeared on their feed that day, or is it just the case that their whole network, including them, is more prone to protest, both online and offline?

Figure 2 would instead suggest that mobilisation to these particular protests was not the result of the online activism by peers. Here, we focus on relative exposure, i.e. the fraction of posts appearing in the feed that contains protest-related hashtags. As expected, non-protesters have less exposure to protest-related posts than protesters. To further strengthen our argument, we perform a 2-way ANOVA test showing that the three samples are significantly different from each other ($p < 0.001$). However, protest participants do not appear to be exposed to a more substantial fraction of protest-related posts in the days leading up to the protests and certainly not during the days of the actual protest (June 9 and 16), which have some of the lowest exposure scores. Instead, we notice upticks in their exposure on the days after the protest. This phenomenon could indicate that the protesters active on Instagram are ahead of the curve - they are not followers but leaders of the protest.

The non-protesters, on the other hand, were most exposed on June 12, after their exposure increased somewhat in the wake of the first protest. Although June 12 does not correspond to any of the two protests examined, another demonstration occurred on this day, and protesters attempted to prevent the bill's second reading by the legislative council. Non-protesters thus appear to be exposed with some delay to protest-related information. They may also receive more exposure to the other demonstrations on weekdays instead of the weekend.

Protesters have a higher activity around protest-related posts in their network. We display the top 100 hashtags clouds collected in posts of the network of protesters in Figures 6.d and

6.e, and the network of non-protesters in Figure 6.f (Appendix). Interestingly, our original intuition that protesters would be more exposed to posts related to protests is not immediately apparent. It appears that the network activity translates into likes and comments more than posts. However, as our study takes place in the first 20 days of June, only the most involved users may post about protests. We expect the networks to catch up later in July and August.

We then quantify the interactions with the content that users are likely exposed to. Figure 3a and 3b shows the average number of likes and comments for the P_a and P_b respectively. The results are consistent with the finding of previous studies that likes are more engaging than comments. While the number of likes to protest-related posts increases as the protest day comes closer for both protests and both sets of protesters, the number of comments does not change significantly. Moreover, as indicated in previous plots, the number of posts significantly increased during and leading up to the protests. After the first protest day (June 9), there is some decrease in the number of posts until June 12, when a group of people demonstrated in front of the Legislative Council of Hong Kong. After another slight decrease, the number of posts gradually increases during the two days preceding the second protest (June 16). We also notice a higher engagement for the protest-related posts on protests days.

Finally, we quantify the interactions for the potentially exposed content to the users. Figure 3 displays the comment and likes for both sets of protesters over the studied period. These pictures confirm that the networks of protesters follow their posting trends regarding protests. We, however, notice two interesting trends. The network of protesters tends to react to the first protest one day after, on June 10. However, for the second protest, the network seems to respond to posts as they are posted. On both dates, the like-to-comment ratio is at its highest. The protest on June 9 started the political discourse and made people interact with the related posts. For the second protest, the discussion continued as 2 million people are reported to have participated, and perhaps has attracted more people to engage in political discourse. Another interesting point is the lower reaction to posts related to June 12 protest. Despite the number of protests-related posts being as high as during the first protest, the number of comments is halved. However, such correlations are hard to analyse based on small scale studies. We want to extend this work to more extensive longitudinal studies.

7 IMAGE CONTENT ANALYSIS

Until this point, we primarily focused on text content and the network of users as it allows us to draw purely analytical conclusions. However, Instagram is an image sharing platform. We show a few examples of visual content associated with protest hashtags in Figure 5 (Appendix). Users primarily post visual content related to protests, whether a call to protests, memes, pictures of places and events, or art. It is interesting to note the presence of screenshots from other websites, circumventing the limited sharing features of Instagram and protests hashtags applied to more consensual images such as food pictures. Image analysis can give us a better understanding of the type of content people post during the protest. It also helps us qualitatively link the content with the interaction features available on Instagram. We follow a three-step approach:

Table 3: Top manual tags and top 3 co-ocurring tags in the protest related images.

Tag	Count	Co-ocurring
police	193	helmet, mask, gun
mask	141	helmet, police, umbrella
drawing	102	police, protester, mask
helmet	96	mask, police, shield
unrelated	89	food, meme, portrait
protest	83	mask, protester, umbrella
umbrella	73	mask, protest, street
street	66	protest, umbrella, mask
screenshot	64	twitter, police, article
protester	53	mask, helmet, umbrella

Table 4: Major Theme of the protest related images

Theme	Details
Campaign	Images focusing on the written text, banners, and screenshots
Activities	Images taken during the protests, mainly focusing on the rallies and protest locations
Actors	Images related to political figures, instituate such as Police, and journalists
Incidents	Images related to incidents of gun shots, tear gas, and arrests
Symbols	Images containing the particular images that have widely been associated with protests such as helmets, umbrella, and masks

automated object detection, manual tagging, and eventually grouping the content into five thematic areas.

Step 1: To analyse the content of the images posted by users during these movements, we download 38,090 images using Instagram search with protest-related hashtags. Using this method, we only download the images without any personal information on the users. Following a similar approach as [18], we use the Microsoft Azure Cognitive API for computer vision to analyse 1000 randomly picked images from the dataset. This API can provide a detailed description of the pictures. We extract the following features from the images as provided by the Azure API[6]: Tags, Categories and Description of the images. Table 6 shows the results of this analysis. Azure categorises many pictures as text or people. Together with the tags and random inspection of images, we observe that most of the images contain protest slogans, flags, infographics, or are screenshots of Twitter feeds, conversations and websites. However, the Azure API remains too generic for our analysis, as the tags do not add much useful information. For instance, the tag 'Nintendo' in the 'text menu' category tends to be associated with most images featuring a cartoon character.

Step 2: To compensate for the generic results of Azure's API, we manually annotate the images. A previous work by [10] on social movements on Instagram manually identifies five categories in the images posted for three different social movements, but only considers 30 images out of the 8,415 image dataset, a number much smaller than our observation set. While the previous work highlights only high-level categories, we attempt to describe the graphic content finely. We manually annotate 1,000 images from the previous step to identify the images' main attributes, objects, and categories. We try to describe the content of each image in a

maximum of ten tags. For images featuring protesters and police officers, we strive to describe their attire, including masks, helmets, or umbrellas. In total, we describe our set of 1,000 images in 735 tags. Table 3 shows the top tags in pictures, as well as the three most common co-occurring tags. By using human annotation, we can significantly refine the results of the computer vision API.

Step 3: Once we have additional tags given to the images, we further categorize the content on thematic areas. We use the Affinity Diagram method to identify the thematic area that image is used to highlight [9, 10], with the five main categories (Table 4).

The results show that people post content related to events and the context during the protests. The two most common objects identified are police and mask. Almost 20% of the pictures feature a police officer, while almost 15% feature an individual (police or protester) wearing a mask. These numbers can be explained by the fact that protesters developed a strong anti-police sentiment, and derogatory comments accompany many images depicting the police. Police officers often appear wearing the full anti-riot uniform. In the early stages of the protest, masks were used to protect protesters' identities and to protect them from tear gas. Later on, masks became a prominent symbol as the government banned wearing masks in public. The umbrella has been a familiar symbol since the Hong Kong Umbrella movement in 2014. Protesters have also used umbrellas as barricades and protection measures during clashes against the police. Overall, when protesters are the image's subject, they tend to be represented with a mask, helmet and umbrella. However, when the protest itself is the subject, protesters primarily wear a mask and, eventually, an umbrella. It appears that the subject of portraits is often protesters who are prepared for clashes with the police. At the same time, pictures featuring the protest tend to focus more on people participating in the rallies.

We notice that almost 1% of the images are unrelated. These images noticeably do not directly relate to the protests yet contain protest-related hashtags. When looking more detail at the co-occurring tags, we observe that most are pictures of food and drinks or memes. We find that such photos might refer to the *'Yellow' camp shops* representing the supporters of democracy in Hong Kong and local activists, primarily through financial support. The online activist disseminates such products (food in this case) to incite consumers to support the *'Yellow' camp shops* to sustain such financial assistance. We also confirm the findings of [10]; Instagram users tend to share screenshots from other websites, as Instagram does not have the sharing features of other social media. Finally, 1.4% of the pictures describe the protests in a negative tone. This finding is exciting as we failed to locate hashtags negatively depicting protests in our data collection process. It appears that users posting such images accompany them with the commonly used set of protest-related hashtags instead of creating their new hashtags.

We note that most content is posted under the symbols and actors category by analysing the most common thematic areas. Based on our qualitative analysis, the symbols and actors category usually contain the users' created messages and expressions of support with other protesters. One particular aspect of such messages being in high numbers can be associated with Instagram interactions. As such, campaign-related messages containing information about the rallies, etc., would likely be re-shared more to spread

more information about the demonstrations. However, as afore-mentioned, Instagram lacks this feature and work around tactics like screenshots and sharing is less common as it is not intuitive in the platform. On the contrary, the users' own messages that users can easily post are more common in terms of symbols and actors.

8 CONCLUSION AND DISCUSSION

In this paper, we have analyzed protest mobilization on a platform that is often thought to be for selfies and beauty [11] than a political: Instagram. Although presenting limited functionalities for social activism, we showed that Instagram was highly relevant to the Hong Kong Anti-extradition Bill protests. An effect of online social media activism translates onto offline activism. Users present at the demonstrations were more active online than non-protesters (H1). They had more interactions and exposure to protest-related content on Instagram than those who did not participate (H2). We showed that users and their networks correlate in terms of political engagement. However, protest participants' offline activity may temporally precede online activity in their network. Protesters and non-protesters differ in the content and the likes and comments on their Instagram feed (H3). Any exposure to protest-related information appears to lag for non-protesters. Regarding the published image content, we showed that pictures of the police were the most frequent, reinforcing the anti-police sentiment. We also showed that activists adapt to the characteristics of this highly visual platform by focusing on symbols, sharing screenshots of other media, and embedding text in the images to improve the media richness.

Our results show the use of Instagram in a social movement, and these results can operationalize into the differences in users' behavior and the use of the platform. Looking from the perspectives of the Instagram interaction features, **we show that even though there is the limited functionality of *sharing* a content. Instagram's *follow network* is instrumental in enhancing social activism**. It is evident through our results of exposure of protest-related content for the protesters. In terms of interaction design, quantifying the user effort to see the content from the network can optimize the effects of network exposure. The feed-ranking algorithm and the paid promotions can affect such exposure based on users' profiles, networks, and overall activities on the platform. In addition, it becomes even more important when such ranking manipulation is coupled with Instagram's interface. For instance, when a user scrolls through the Instagram feeds on a particular mobile device, the whole screen space is usually covered by one post. How much effort the user has to put (scrolling down the content) can affect the effectiveness of the network exposure.

In terms of images, the primary content on the platform, **we confirm the results of a previous study about screenshot usage**, and the assumption that it is related to the lack of sharing option. In addition, our content analysis shows a more specific usage of screenshots. **Screenshots are used to share the campaign-related generic messages but not the offline activities**. These messages mostly contain written slogans and banners related to protests. However, screenshots are not typical for showing offline activities that highlight the protesters or raising the voice to support protesters. Nonetheless, screenshots do not represent most of the content. A large sample of the image content consists of

helmets, masks, umbrellas, and guns. Use of *umbrella* goes back to the 2014 Umbrella Movement in Hong Kong. This distinction shows that **Instagram users prefer to share content that is more representative of offline activities and raise the support for the protesters using symbolic language**. We also highlight the usage of images to indirectly report the stance of some entities, in this the restaurants. Instagram users tried to promote the business for those restaurants while increasing their outreach to other activists by using protest-related hashtags.

Finally, the analysis of hashtags shows the distinction between protesters and non-protesters. Protesters use the native language (Cantonese) while non-protesters have a higher percentage of English. One of the key reasons is that the protesters leverage their native language to communicate the ground operations on the protest sites among themselves [21] and also the native language contains higher emotional support [23].

In summary, consistent with the prior findings [34] on social media's role, Instagram symbols in images to call for collective action and show of support. The visuals have been strongly supported by literature to provoke the necessary psychological reactions in driving the protests [7, 27]. The language of the protesters provides a sense of unity. Instagram works as a valuable channel for disseminating images and slogans (contained in hashtags) to raise the group moral [25] and fuel up future events [38]. We put forward the idea that Instagram incorporates both centralized and decentralized characteristics under the decentralized movement depending on how the images are used (screenshots vs user-generated images). Instagram offers its users to work in networked settings and as individual voices.

Limitations. In addition to valuable insights, our work is still limited in certain aspects. One of the primary reasons for that is limited capabilities regarding large scale data collection on Instagram. Moreover, we do not study the Instagram stories in our work. Instagram is also known to promote paid promotions. We do not have any information and characterization of content as such. The network-based approach in Instagram also has a limitation. Sometimes, a user's friends might have set their profiles to private. This situation might result in lesser data for exposure analysis. There is no workaround for this situation, and it will remain as it is until Instagram changes the privacy settings.

REFERENCES

[1] Nova Ahmed, Zareen Tasnim, and Jasmine Jones. 2019. Digital Silence and Liberating Stories: During a Student-Driven Movement. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, alt02.

[2] Martin R Albrecht, Jorge Blasco, Rikke Bjerg Jensen, and Lenka Mareková. 2021. Collective Information Security in Large-Scale Urban Protests: the Case of Hong Kong. *arXiv preprint arXiv:2105.14869* (2021).

[3] Bloomberg.com. 2019. *A Brief History of the Protests*. https://www.bloomberg.com/graphics/hong-kong-protests-timeline/ Accessed:2010-09.

[4] Yarimar Bonilla and Jonathan Rosa. 2015. #Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States. *American Ethnologist* 42, 1 (2015), 4–17. https://doi.org/10.1111/amet.12112

[5] Enrique Bonsón, David Perea, and Michaela Bednárová. 2019. Twitter as a tool for citizen engagement: An empirical study of the Andalusian municipalities. *Government information quarterly* 36, 3 (2019), 480–489.

[6] Enrique Bonsón and Melinda Ratkai. 2013. A set of metrics to assess stakeholder engagement and social legitimacy on a corporate Facebook page. *Online Information Review* (2013).

[7] Felix Brünker, Fabian Deitelhoff, and Milad Mirbabaie. 2019. Collective Identity Formation on Instagram – Investigating the Social Movement Fridays for Future.

[8] Hsia-Ching Chang. 2010. A new perspective on Twitter hashtag use: Diffusion of innovation theory. *Proceedings of the American Society for Information Science and Technology* 47, 1 (2010), 1–4. https://doi.org/10.1002/meet.14504701295

[9] Juliet Corbin and Anselm Strauss. 2014. *Basics of qualitative research: Techniques and procedures for developing grounded theory.* Sage publications.

[10] Victor P. Cornet, Natalie K. Hall, Francesco Cafaro, and Erin L. Brady. 2017. How Image-Based Social Media Websites Support Social Movements. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems (CHI EA '17).* ACM, New York, NY, USA, 2473–2479.

[11] Rachel L Einwohner and Elle Rochford. 2019. After the March: Using Instagram to perform and sustain the Women's March. In *Sociological Forum*, Vol. 34.

[12] Ruben Enikolopov, Alexey Makarin, and Maria Petrova. 2017. Social media and protest participation: Evidence from Russia. *Available at SSRN 2696236* (2017).

[13] Sheena Erete and Jennifer O Burrell. 2017. Empowered participation: How citizens use technology in local governance. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems.* 2307–2319.

[14] Korina Giaxoglou. 2018. # JeSuisCharlie? Hashtags as narrative resources in contexts of ecstatic sharing. *Discourse, context & media* 22 (2018), 13–20.

[15] Andrea Hamm, Yuya Shibuya, Stefan Ullrich, and Teresa Cerratto Cerratto Pargman. 2021. What Makes Civic Tech Initiatives To Last Over Time? Dissecting Two Global Cases. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems.* 1–17.

[16] Ehsan Ul Haq, Tristan Braud, Young D Kwon, and Pan Hui. 2020. A survey on computational politics. *IEEE Access* 8 (2020), 197379–197406.

[17] Tad Hirsch and John Henry. 2005. TXTmob: text messaging for protest swarms. In *CHI'05 extended abstracts on Human factors in computing systems.* ACM, 1455–1458.

[18] Yuheng Hu, Lydia Manikonda, and Subbarao Kambhampati. 2014. What we instagram: A first analysis of instagram photo content and user types. In *Eighth International AAAI conference on weblogs and social media.*

[19] Christoph Kastenholz. 2021. Council Post: Gen Z And The Rise Of Social Commerce. https://www.forbes.com/sites/forbesagencycouncil/2021/05/17/gen-z-and-the-rise-of-social-commerce/?sh=14dab2d5251d

[20] Avril Keating and Gabriella Melis. 2017. Social media and youth political engagement: Preaching to the converted or providing a new voice for youth? *The British Journal of Politics and International Relations* 19, 4 (2017), 877–894.

[21] Yong Ming Kow, Bonnie Nardi, and Wai Kuen Cheng. 2020. Be Water: Technologies in the Leaderless Anti-ELAB Movement in Hong Kong. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.* 1–12.

[22] Jennifer M Larson, Jonathan Nagler, Jonathan Ronen, and Joshua A Tucker. 2019. Social networks and protest participation: Evidence from 130 million Twitter users. *American Journal of Political Science* 63, 3 (2019), 690–705.

[23] Carmen Lee and Dennis Chau. 2018. Language as pride, love, and hate: Archiving emotions through multilingual Instagram hashtags. *Discourse, Context & Media* 22 (2018), 21–29.

[24] Francis L. F. Lee, Samson Yuen, Gary Tang, and Edmund W. Cheng. 2019. Hong Kong's Summer of Uprising: From Anti-Extradition to Anti-Authoritarian Protests. *China Review* 19, 4 (2019), 1–32. https://www.jstor.org/stable/26838911

[25] Yu-Hao Lee and Gary Hsieh. 2013. Does Slacktivism Hurt Activism? The Effects of Moral Balancing and Consistency in Online Activism. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13).* Association for Computing Machinery, New York, NY, USA, 811–820.

[26] Hanlin Li, Disha Bora, Sagar Salvi, and Erin Brady. 2018. Slacktivists or activists?: Identity work in the virtual disability march. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, 225.

[27] Tawseef Majeed and Ali Abushbak. 2020. The Role of Social Media in Mobilizing Crowd Protest–A Case Study of Palestinian Anger Against Israel's Military Measures at Al-Aqsa Compound on Instagram. 40 (06 2020), 511–526.

[28] David Matley. 2018. "Let's see how many of you mother fuckers unfollow me for this": The pragmatic function of the hashtag# sorrynotsorry in non-apologetic Instagram posts. *Journal of Pragmatics* 133 (2018), 66–78.

[29] Sathappan Muthiah, Bert Huang, Jaime Arredondo, David Mares, Lise Getoor, Graham Katz, and Naren Ramakrishnan. 2015. Planned protest modeling in news and social media. In *Twenty-Seventh IAAI Conference.*

[30] NewsAsia. 2019. Hong Kong protests: Key dates as peaceful rallies against extradition Bill turn to violent clashes. https://www.channelnewsasia.com/news/asia/hong-kong-protest-timeline-extradition-bill-china-11676634

[31] Yuki Ogawa, Kakuko Miyata, Hitoshi Yamamoto, and Ken'ichi Ikeda. 2013. Association Between Selective Exposure and Attitude on Twitter. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 03 (WI-IAT '13).* IEEE Computer Society, Washington, DC, USA, 122–125. https://doi.org/10.1109/WI-IAT.2013.164

[32] Sameer Patil and Apu Kapadia. 2012. Are You Exposed?: Conveying Information Exposure. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work Companion (CSCW '12).* ACM, New York, NY, USA, 191–194. https://doi.org/10.1145/2141512.2141575 event-place: Seattle, Washington, USA.

[33] Hao Peng, Ceren Budak, and Daniel M Romero. 2019. Event-Driven Analysis of Crowd Dynamics in the Black Lives Matter Online Social Movement. In *The World Wide Web Conference.* ACM, 3137–3143.

[34] Thomas Poell and José Van Dijck. 2015. *Social media and activist communication.* 527–537.

[35] Haodi Qi, Hanyu Jiang, Wende Bu, Chengzi Zhang, and Kyong Jin Shim. 2019. Tracking Political Events in Social Media: A Case Study of Hong Kong Protests. In *2019 IEEE International Conference on Big Data (Big Data).* IEEE, 6192–6194.

[36] Naren Ramakrishnan, Patrick Butler, Sathappan Muthiah, Nathan Self, Rupinder Khandpur, Parang Saraf, Wei Wang, Jose Cadena, Anil Vullikanti, Gizem Korkmaz, et al. 2014. 'Beating the news' with EMBERS: forecasting civil unrest using open source indicators. In *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 1799–1808.

[37] Suhas Ranganath, Fred Morstatter, Xia Hu, Jiliang Tang, Suhang Wang, and Huan Liu. 2016. Predicting online protest participation of social media users. In *Thirtieth AAAI Conference on Artificial Intelligence.*

[38] M. Rovisco. 2016. A new 'Europe from below'? Cosmopolitan citizenship, digital media and the indignados social movement. *Comparative European Politics* 14 (2016), 435–457.

[39] SCMP. 2018. What social media platforms dominate Hong Kong? https://www.scmp.com/news/hong-kong/community/article/2142652/what-social-media-platforms-run-hong-kong-and-what-are

[40] Noam Segev, Noam Avigdor, and Eytan Avigdor. 2018. Measuring Influence on Instagram: A Network-Oblivious Approach. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval (SIGIR '18).* ACM, New York, NY, USA, 1009–1012. event-place: Ann Arbor, MI, USA.

[41] Kjerstin Thorson, Kelley Cotter, Mel Medeiros, and Chankyung Pak. 2021. Algorithmic inference, political interest, and exposure to news and politics on Facebook. *Information, Communication & Society* 24, 2 (2021), 183–200.

[42] Martino Trevisan, Luca Vassio, Idilio Drago, Marco Mellia, Fabricio Murai, Flavio Figueiredo, Ana Paula Couto da Silva, and Jussara M Almeida. 2019. Towards Understanding Political Interactions on Instagram. *arXiv preprint arXiv:1904.11719* (2019).

[43] Min-Hsuan Tsai, Charu Aggarwal, and Thomas Huang. 2014. Ranking in heterogeneous social media. In *Proceedings of the 7th ACM international conference on Web search and data mining.* 613–622.

[44] Oren Tsur and Ari Rappoport. 2012. What's in a Hashtag?: Content Based Prediction of the Spread of Ideas in Microblogging Communities. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining (WSDM '12).* ACM, New York, NY, USA, 643–652. event-place: Seattle, Washington, USA.

[45] Zeynep Tufekci and Christopher Wilson. 2012. Social media and the decision to participate in political protest: Observations from Tahrir Square. *Journal of communication* 62, 2 (2012), 363–379.

[46] Ehsan ul Haq, Tristan Braud, Young D Kwon, and Pan Hui. 2020. Enemy at the Gate: Evolution of Twitter User's Polarization During National Crisis. In *2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM).* IEEE, 212–216.

[47] Volker Wulf, Konstantin Aal, Ibrahim Abu Kteish, Meryem Atam, Kai Schubert, Markus Rohde, George P Yerousis, and David Randall. 2013. Fighting against the wall: social media use by political activists in a Palestinian village. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM.

[48] Volker Wulf, Kaoru Misaki, Meryem Atam, David Randall, and Markus Rohde. 2013. 'On the ground'in Sidi Bouzid: investigating social media use during the tunisian revolution. In *Proceedings of the 2013 conference on Computer supported cooperative work.* ACM, 1409–1418.

[49] Jian Xiao and Isabel Fangyi Lu. 2021. Art as intervention: Protests on urban transformation in China and Australia. *Journal of Urban Affairs* (2021), 1–20.
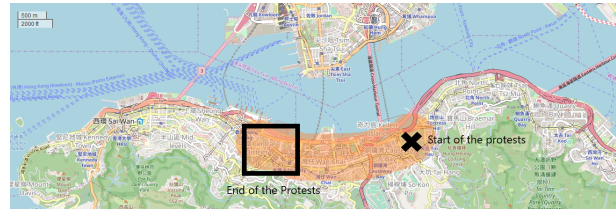
APPENDIX

Background: Hong Kong Protests

Hong Kong protests[7] started in reaction to a bill entitled "The Fugitive Offenders and Mutual Legal Assistance in Criminal Matters Legislation (Amendment) Bill 2019"[8], proposed by the Hong Kong government in February 2019. Small rallies and protests followed this announcement. Large scale protests started with the second reading of the bill in June 2019. According to the organizers, more than two million people joined the first protests. Since its introduction, the original bill was modified two times to scale down the terms and conditions. On July 9, 2019, the Chief Executive of Hong Kong, Carrie Lam, announced that the bill was "dead" and finally agreed to a complete withdrawal on September 4, 2019. However, in reaction to police actions against the protests, the opposition movement extended its demands to include more general social issues, and protests continued into 2020. Table 5 summarizes the important dates up to July 1 2019 [3, 30]. This paper focuses on the two largest demonstrations that happened on June 9 and 16, 2019.

Figure 4: Map of the protests (north of Hong Kong Island). Area of the protests represented in Orange. Both studied protests started in Victoria Park and ended between Admiralty and Central.

[5] https://en.wikipedia.org/wiki/2019_Hong_Kong_protests
[6] https://en.wikipedia.org/wiki/2019_Hong_Kong_extradition_bill

Table 5: HK Protests Timeline until July 1.

Date	Event
Early 2018	A HK resident kills his girlfriend in Taiwan and returns to Hong Kong. Due to the lack of an extradition agreement between Hong Kong and Taiwan, the HK police cannot charge or extradite him.
Feb. 12	Proposal for fugitive extradition enabling the case-by-case transfer of fugitives to jurisdictions for which HK lacks a formal extradition treaty.
Mar. 27	The proposal is scaled down in reaction to the backlash and strong criticism from the civil society.
Mar. 31	First protest against the bill, attended by between 5,200 (police estimation) and 12,000 (organizer estimation) protesters.
Apr. 3	First reading of the bill at the HK Legislative Council (LegCo).
Apr. 28	Between 22,800 (police estimate) and 130,000 (organizer estimate) protesters march against the bill.
May 11	Opposition and pro-government lawmakers clash at HK LegCo.
May 31	HK government brings more changes to the bill.
Jun. 9	1.03 million individuals join the first major protest according to the organizers (270,000 according to the police).
Jun. 12	Activists gather in front of the LegCo building to prevent the second reading of the bill.
Jun. 16	Largest demonstration in Hong Kong to date, with between 338,000 (police) and 2 million (organizers) participants
Jun. 21	Protests around Police Headquarters
Jul. 1	Protesters storm the LegCo building

Table 6: Top categories, detected by MS Azure API, in the protest related images.

Category	Object Tags
others	text , screenshot , person , cartoon , design
outdoor	text , person , outdoor , clothing , screenshot
text sign	text , design , screenshot , poster , cartoon
text mag	text , person , screenshot , poster , human face
text	text , screenshot , design , poster , font
text menu	text , screenshot , design , nintendo , abstract
abstract	text , screenshot , person , outdoor , clothing
people crowd	person , outdoor , clothing , man , people
abstract nonphoto	cartoon , drawing , text , sketch , nintendo
people group	person , clothing , text , man , human face

(a) (b) (c) (d) (e)

Figure 5: Typical image content associated with protest hashtags on Instagram. (a) pictures of events, (b) screenshots of articles and social media posts, (c) food pictures, (d) portraits of actors, (e) art created by users.

(a) Hashtag cloud for P_a (b) Hashtag cloud for P_b (c) Hashtag cloud for Non Protesters

(d) Hashtag cloud for P_a network (e) Hashtag cloud for P_b network (f) Hashtag cloud for NP network

Figure 6: Hashtag clouds for protesters, non protesters and their networks