Temporal Dynamics of User Engagement on Instagram: A Comparative Analysis of Album, Photo, and Video Interactions

Elin Thorgren
Linköping University
Sweden

Alireza Mohammadinodooshan
Linköping University
Sweden

Niklas Carlsson
Linköping University
Sweden

ABSTRACT

Despite Instagram being an integral part of many people's lives, it is relatively less studied than many other platforms (e.g., Twitter and Facebook). Furthermore, despite offering diverse content formats for user expression and interaction, prior works have not studied the temporal dynamics of user engagement across albums, photos, and videos. To address this gap, we present a pioneering temporal comparative analysis that unveils nuanced patterns in user interactions across content types. Our analysis sheds light on interaction longevity and disparities among album, photo, and video engagement. Additionally, it offers empirical comparisons through statistical tests, examines contributing factors such as post and uploader characteristics, and analyzes content composition's impact on user engagement. The findings reveal distinct temporal engagement patterns. Despite initial spikes in interactions post-upload, albums exhibit somewhat more sustained interest, while photos and videos have shorter engagement lifespans. Moreover, a consistent trend between shallow (likes) and deep (comments) interactions persists across content types. Notably, concise content, characterized by shorter descriptions and minimal hashtags/mentions, consistently drives higher engagement, emphasizing its relevance across all content formats. These insights deepen comprehension of temporal nuances in user engagement on Instagram, offering valuable guidance for content creators and marketers to tailor strategies that evoke immediate and sustained user interest.

CCS CONCEPTS

• **Information systems → World Wide Web**; **Information systems applications**.

KEYWORDS

Instagram, User engagement, Temporal dynamics, Interactions, Photo, Video, Album

ACM Reference Format:
Elin Thorgren, Alireza Mohammadinodooshan, and Niklas Carlsson. 2024. Temporal Dynamics of User Engagement on Instagram: A Comparative Analysis of Album, Photo, and Video Interactions. In *ACM Web Science Conference (Websci '24), May 21–24, 2024, Stuttgart, Germany*. ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3614419.3644029

Websci '24, May 21–24, 2024, Stuttgart, Germany

ACM ISBN 979-8-4007-0334-8/24/05.
https://doi.org/10.1145/3614419.3644029

1 INTRODUCTION

Social media platforms like Instagram have become an integral part of modern communication and are today an important aspect of many people's lives. Through the use of various content formats like albums, photos, and videos, it offers its users several unique avenues for expression, engagement, and interaction. Understanding how user engagement and interactions unfold over time across these content types is important for content creators, marketers, and platform developers seeking to optimize engagement strategies and user experiences.

However, despite Instagram's popularity, its offering of diverse content formats, and prior works (e.g., [16]) having explored engagement dynamics, the temporal dynamics of user engagement across these formats remains underexplored. To address this gap, we present a pioneering study of the temporal user engagement patterns on Instagram, placing particular focus on the comparative analysis of interactions observed across albums, photos, and videos. By characterizing and scrutinizing the temporal dynamics of these content types on Instagram, including the images visible to the users as well as other meta-information about each post, we uncover new nuanced insights into user behaviors and preferences within the realm of visual content consumption on Instagram.

Contributions: The contributions are threefold. First and most importantly, we present a novel temporal comparative analysis that rigorously examines the temporal behavior of user interactions with albums, photos, and videos. Our in-depth comparative analysis of their engagement patterns over time provides insight into interaction longevity and highlights intriguing disparities in the longevity of user engagement across content types. Second, we complement our temporal comparisons with head-to-head comparisons of the overall per-post distributions, where we present a mix of statistical tests to support our empirical observations.

Third, we study the impact of contributing factors such as post characteristics, uploader characteristics, and a media-based analysis that considers the content composition. For example, by investigating the influence of content composition factors like description length, hashtags, and mentions on user engagement, the study unveils compelling correlations between content characteristics and interaction levels across all post types. We are also one of the first to present a view-based analysis of the video posts on Instagram. Finally, we look at some differences observed in the interaction with the major uploader categories.

Example observations: Our analysis uncovers several distinct temporal patterns in user engagement. For example, despite an initial surge in interactions across all post types shortly after upload, the interaction rates quickly decrease across the different media types, with albums being the format with the most sustained interest compared to the relatively shorter lifespans of engagement seen with photos and videos. Notably, the analysis highlighted the enduring disparity between shallow interactions (likes) and deeper interactions (comments) across all post types, underscoring a consistent user trend.

Furthermore, our investigation into factors that drive engagement revealed the significant impact of content composition. Concise descriptions, with fewer words and limited use of hashtags and mentions, consistently correlated with higher interaction levels across all post types. This finding underscores the ability of focused and succinct content to attract greater user engagement, regardless of the content format.

Our observations strengthen the understanding of temporal nuances in user engagement, highlight user preferences, and provide valuable insights for content creators and marketers to tailor their strategies, emphasizing formats that not only prompt immediate engagement but also sustain user interest over extended periods.

Context of related work (Sec. 7): While extensive research has explored temporal engagement dynamics in photo and video-sharing platforms [4, 25, 37], investigations specific to Instagram's context remain limited. Notably, Vassio et al. [35, 36] provided insights into engagement patterns among Italian influencers' posts, differing from our approach in several key ways. Our study diverges by broadening the temporal analysis to encompass and compare the dynamics seen for different content types, unveiling distinct engagement patterns, and analyzing a diverse set of influencers. In examining post, uploader, and media factors impacting engagement, our work differs notably from studies like Bakhshi et al. [3], which highlighted the impact of faces in photos, and Mazloom et al. [24], focused on brand-related media files. Our research uniquely analyzes various post types and quantifies the engagement dynamics among them while also capturing a broader set of influencer/poster categories than prior works.

Outline: Sec. 2 presents the datasets collected and analyzed in this paper. Sec. 3 then presents a post-based comparative analysis in which we compare the engagement across albums, photos, and videos. The following sections then present a view-based analysis (Sec. 4), a media-based analysis (Sec. 5), and an analysis based on our uploader categorization (Sec. 6). Finally, we present related works (Sec. 7) and our conclusions (Sec. 8).

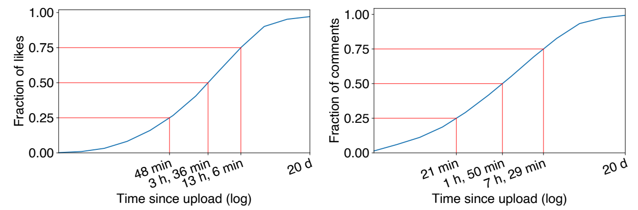
2 DATASET OVERVIEW & TEST METHOD

2.1 Datasets

For the analysis presented in this paper, we collected two datasets.

Statistical Data: The initial dataset comprises data from the top 1,000 Instagram accounts, determined by follower count, along with almost all their posts over a year, spanning from Dec. 15, 2021, to Dec. 14, 2022. For account selection, we picked the top-1K accounts from the top lists of influencers and brand accounts by starngage.com [1] for which we could find accounts that existed during the above data collection period and that had uploaded at least one post during that timeline. For this dataset, we used CrowdTangle (a public insights tool owned and operated by Facebook) [33] to collect the following information:

- **User data**: Information about each user account, including name, username, and the number of followers at the time it was last collected by CrowdTangle.



(a) Likes (b) Comments

Figure 1: Cumulative fraction of interactions (d: days, h: hours, min: minutes).

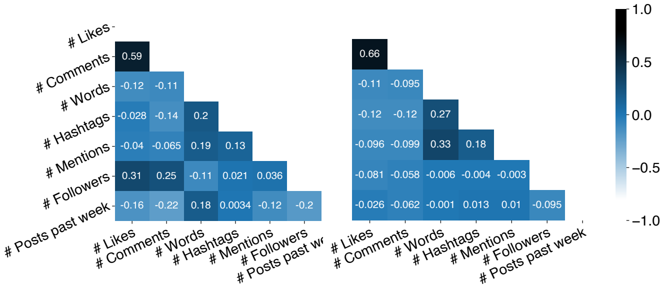
- **Post data**: Information about each post, including the type of post (album, photo, or video), description, date of upload, and the total number of likes, comments, and video views at the time it was last collected by CrowdTangle. It is worth noting here that CrowdTangle, for reporting the video views, counts the number of times a video has been watched for at least 3 seconds. This is similar to the way Facebook deals with this metric.
- **Temporal data**: For each post, we also obtained (up-to) 75 temporal snapshots of the number of likes and comments accumulated since the upload. These snapshots are gathered by CrowdTangle over a 20-day period after upload, with the snapshots being distributed over this time period to mimic a logarithmic curve with inter-snapshot times increasing (logarithmically) from 15 minutes to 24 hours. Here, it is also important to note that the last snapshot provided by CrowdTangle technically does not have a maximum threshold (only a minimum threshold of 20 days). For this reason and to ensure fair comparisons across posts that otherwise may have had different amounts of time to garner attention, we selected to use only the temporal data for the initial 20-day period. Therefore, when studying the temporal dynamics, we in the following refer to the 20-day statistics as the total number of interactions achieved.

While our data contains all posts available via CrowdTangle, we note that CrowdTangle only provides access to public posts, and many posts, therefore, may be missed. Yet, our dataset contains 450K posts, with the individual top-users having posted between 1 to 25K posts per user over the studied one-year period.

Media Content: We gathered nearly all media files accompanying the aforementioned posts from the top-100 users, comprising a total of 45,000 media files, with each user contributing between 1 and 4,600 files. For a fair comparison between posts containing photos, videos, and albums, like in [12], we always extracted and analyzed the first image (or frame) that the user sees of each post. While user engagement could be influenced by subsequent images in albums or frames in videos, this constraint enables uniform image analysis when comparing content across post types.

2.2 High-level Characteristics

Temporal Dynamics of Interactions: Figure 1 shows the average cumulative fraction of likes and comments for all users over the first 20 days after each post was uploaded. The red lines highlight the average age when a post has acquired 25%, 50%, and 75% of



(a) Aggregated calculation **(b) User-based calculation**

Figure 2: Correlation matrices for the statistical data.

Table 1: The most complete/important variables in dataset.

Name	Description	Characteristics	Scale
Likes	Number of total likes after 20-days	User engagement	Log
Comments	Number of total comments after 20-days	User engagement	Log
Views	Number of video views after 20-days	User engagement	Log
Words	Number of words in the posts' description	Post	Log
Hashtags (#)	Number of hashtags in the posts' description	Post	Linear
Mentions (@)	Number of mentions in the posts' description	Post	Linear
Followers	Number of followers for the user at upload	Uploader	Log
Posts past week	Number of posts 7-days prior to upload	Uploader	Log

their total interactions (during the 20-day period). As observed, Instagram posts are short-lived, with 75% of the likes being received during the first 13.1 hours after upload and 75% of the comments being received during the first 7.5 hours.

Statistical Correlations: Prior to examining interaction differences across post types, we conduct an initial analysis to identify correlations among underlying factors and assess redundancy or shared explanatory information among variable groups. Figure 2 presents correlation matrices for the most important variables extracted from the collected data together with the respective pairwise Pearson's correlation coefficients. To account for differences in the relative distributions, some variables were log-transformed. (See Table 1 for a summary and description of each variable.) In the sub-figures we distinguish between correlations calculated over all (as an aggregate) or the average correlation seen across the users.

As observed, the correlation among the variables is moderate to weak, with the strongest correlation being between likes and comments, followed by the various correlations between these user interactions and the uploader characteristics (followers + number of posts). The weak and moderate correlation of this second type suggests that there may be other compounding factors that also impact user engagement levels. As expected, the relatively stronger correlations between user engagement metrics (likes + comments) and uploader characteristics (followers + number of posts) observed for the aggregate model are mostly canceled out for the per-user correlations. In contrast, the post-related characteristics (words + hashtags + mentions) see bigger correlations among themselves with the per-user correlations than for the aggregate set.

View-based Analysis: For videos, we also have access to the view count, which was found to be highly correlated with the number of likes and comments. For example, with the aggregate model, these correlations were 0.92 (with likes) and 0.83 (comments), respectively. For the uploader characteristics, the correlations were again weaker: 0.42 with followers and -0.24 with the number of posts

Table 2: Summary of the number of users, posts, and engagement for each category of users.

Category	Number of users	Number of posts				Engagement per user		
		Album	Photo	Video	Total	Likes	Comments	Views
Musicians	252	16,664	15,126	8,769	40,559	94,695,705	687,304	47,822,288
Others	249	24,847	36,268	7,731	68,846	75,405,771	832,286	45,528,958
Actors	236	15,240	9,672	4,158	29,070	57,138,957	364,897	27,983,632
Brands	172	90,828	157,820	47,780	296,428	224,458,563	2,234,431	238,912,556
Athletes	91	4,366	7,787	1,641	13,794	100,104,093	641,083	36,529,477
Total	**1,000**	**151,945**	**226,673**	**70,079**	**448,697**	**551,803,089**	**4,760,001**	**396,776,911**

in the past week. While part of the higher correlations with the likes can be explained by the video posts, in general, seeing somewhat higher correlations between the user engagement metrics (e.g., the correlation between likes and comments for videos is 0.82), the high correlations suggest that likes can be a good proxy for the relative number of views obtained for different posts.

Categorical Breakdown: To better understand the dataset, we also manually categorized the users into five different categories. In particular, each user was categorized based on their foremost profession except for accounts that belonged to, for example, brands, organizations, sports clubs, and magazines. These types of users had their own category: brands, thus a majority of them were brands. The rest of the users were categorized into one of the following categories: actors, musicians, athletes, or others. Others include professions such as influencer, comedian, and politician. Table 2 summarizes the number of users and posts in each category, as well as the (average) total engagement per user within the 20-day period (i.e., total across the category divided by the number of users in the category).

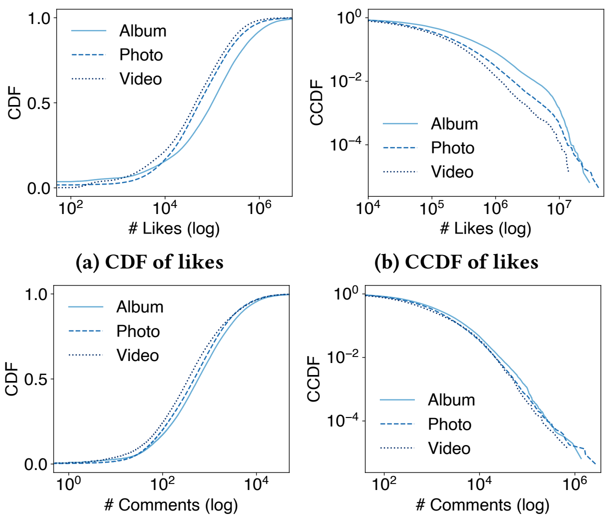
2.3 Statistical Test Methodology

When comparing two or more distributions, we used a series of statistical tests. Here, we briefly summarize these tests. First, the Kruskal-Wallis test [20] was used to determine the statistical significance between two or more sample distributions. This is a non-parametric method, also called one-way analysis of variance (ANOVA) on ranks. When significant differences between the sample classes were observed with the Kruskal-Wallis test, we then: (1) performed pairwise tests using Dunn's test [8] to identify specific pairs of classes that exhibited differences, (2) used the Mann-Whitney U test [23] to confirm significant differences between pairs of sample medians and (3) used the bootstrap method [9] with 10K iterations and a 99% confidence interval to determine the statistical significance between pairs of sample means. Except for bootstrapping, we applied a p-value threshold of 0.005 in our tests.

3 POST-BASED ANALYSIS: ALBUMS VS. PHOTOS VS. VIDEOS

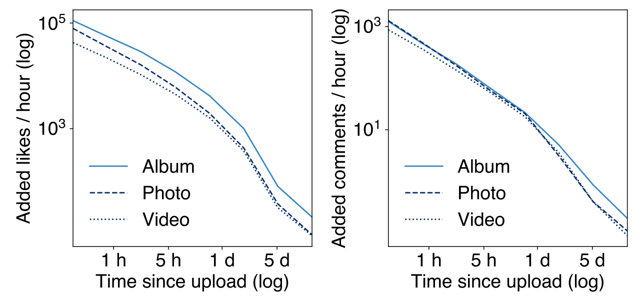
3.1 Lifetime Interactions per Type

The number of likes and comments are often used to measure the level of shallow and deep interactions, respectively [2, 19]. To capture the distribution of user interactions across various post types, Figure 3 presents the cumulative distribution function (CDF) and complementary cumulative distribution function (CCDF) for the total number of likes and comments with respect to each post type. To capture the big variations in variation levels, logarithmic x-axis are used. As indicated by the separation of the curves, regardless



(a) CDF of likes (b) CCDF of likes
(c) CDF of comments (d) CCDF of comments

Figure 3: Distributions of total interactions



(a) Likes (b) Comments

Figure 4: Temporal reduction in interaction intensity for different post types.

of interaction metric, the albums receive the most interactions and the videos receive the least interactions. Furthermore, as illustrated by the relatively straight-line shape of the CCDFs, all post types follow a heavy-tailed distribution, with a limited number of users and posts receiving most of the interactions. Here, photos seem to have a heavier tail compared to albums, suggesting that the photo category includes some additional outliers with many interactions.

Statistical tests confirmed significant differences in the engagement with the three different types of posts. For example, after observing significant differences between the distributions using the Kruskal-Wallis test (p-values < 10^{-147}), pairwise Dunn's tests (distribution comparisons) and pairwise Mann-Whitney U tests (median value comparisons) were all found significant with p-values < 0.005, and so were pairwise Bootstrapping with 10K iterations and a 99% confidence interval (mean value comparisons).

3.2 Temporal Type Comparisons

To capture the temporal differences in the interaction rates with different post types, we next studied the rate that interactions where accumulated as a function of the time since a post was first uploaded. Here, temporal data for each post was first divided into eight logarithmically increasing time buckets. The average interaction rate for each bucket was then computed by tallying interactions from the last instance in the previous bucket to the last instance in the current one, and then dividing this by the time span. Figure 4 summarizes these results.

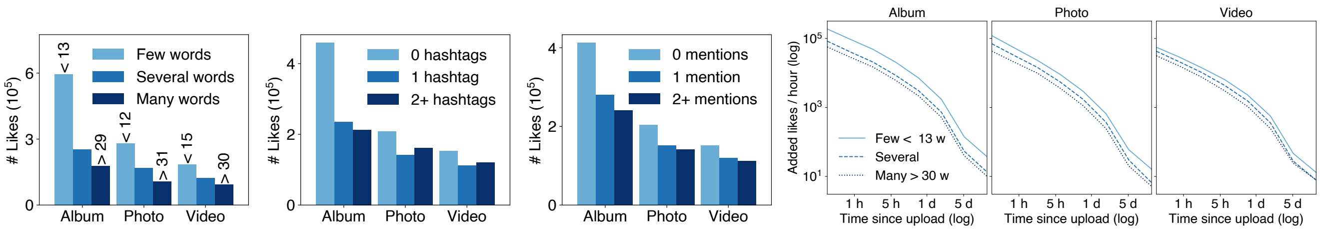
From the figure, we make several observations. First, for both interaction types, users follow a similar pattern across the three post types. For example, in all cases, the interaction rates quickly decrease as the posts age, resulting in a drop by close to a factor of 100 over the first 24 hours. Second, the relative differences in the number of interactions that each type obtains over the lifetime appear to hold true at all stages of the posts' lifetime. For example, albums receive the most interactions at all stages of the posts' lifetime, while the videos receive the least interactions (on average) throughout the 20-day time period. Third, we observe relatively smaller differences between the types when considering the comments (deeper interactions) compared to the likes (shallow interactions). Fourth, we observe the biggest drop in interaction rates for the photos, suggesting that photos might fall off the users' radar more quickly than albums and videos. We next look closer at how post characteristics and uploader characteristics impact the observed interaction rate differences.

3.3 Impact of Post Characteristics

By evaluating how various key characteristics of posts influence their potential success across the three different post types under consideration, we have found that the posts that have short descriptions (e.g., with few words, no hashtags, or no mentions) are more successful than those with longer or more complex descriptions. These findings were consistent irrespective of which post type was considered and whether considering shallow or deep interactions. To substantiate this claim and highlight some subtle differences, we next look closer at the impact of three key characteristics and how they impacted the relative success rate of posts.

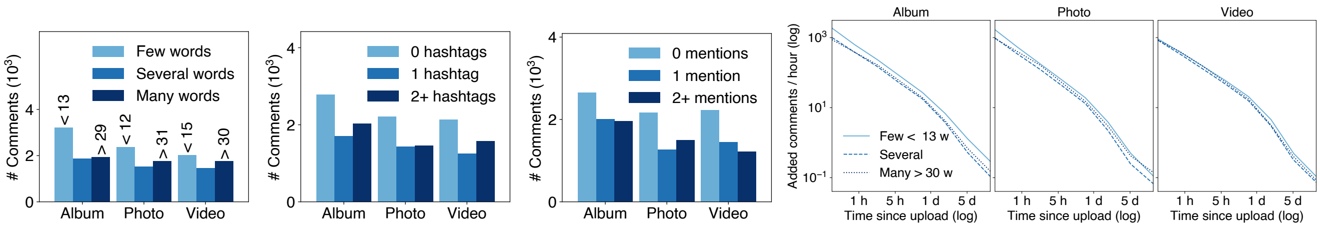
Word count: Consider first the number of words used to describe the posts. For this analysis, we do not distinguish between regular words, numbers, hashtags, mentions, and emojis; they are each counted as one word. Figures 5a and 6a show the average values and Figures 5d and 6d show the average hourly rate that likes and comments are added as a function of time since upload for three classes of posts: (1) few words (< 13 words), (2) several words (13 to 20 words), and (3) many words (> 20 words). These classes were picked to most closely have the classes be split around the 33^{th} and 67^{th} percentile of the overall words per post distribution. From the figure, we clearly see that posts containing few words are most successful, regardless of which type of post it is (i.e., album, photo, or video) or type of interactions (i.e., likes or comments). The main difference is instead for the least successful posts, where we can see that the posts with many words in their description receive the fewest likes, whereas the posts with an intermediate word count receive the fewest comments. This may be due to some uploaders' long descriptions actually encouraging users to comment.

Statistical tests (using 0.005 level) confirmed that the differences were statistically significant for all pairwise comparisons of the



(a) Word count (b) Hashtags (c) Mentions (d) Temporal impact of word count (w: words)

Figure 5: Differences in shallow interaction level (number of likes) for posts with different post characteristics.



(a) Word count (b) Hashtags (c) Mentions (d) Temporal impact of word count (w: words)

Figure 6: Differences in deep interaction level (number of comments) for posts with different post characteristics.

distributions and medians. However, for the means (using bootstrapping), we did not observe significant differences in the number of comments when comparing album posts with several vs. many words in their descriptions. Otherwise, the observed differences were significant also here.

Hashtags and mentions: Figures 5b and 6b present the average impact that hashtags have on interactions, and Figures 5c and 6c show the corresponding results for mentions. In each figure, we show the average interactions for albums, photos, and videos, categorized by the presence of hashtags and mentions in the post descriptions. As observed, posts without any hashtags or mentions tend to gather the highest interactions. Our statistical tests validated the significance of observed differences between posts with and without mentions at the 0.005 level. Also, for hashtags, the Kruskal-Wallis test showed statistical significance (with the highest p-value being $1.97 \cdot 10^{-57}$ for likes and $1.73 \cdot 10^{-104}$ for comments). The Dunn's test and Mann-Whitney test indicated no significant difference in likes between videos with and without hashtags (p-values of 0.73 and 0.52, respectively).

Discussion: While hashtags and mentions are commonly used as a means to reach new users and help uploaders attract a broader audience, these results show that the most successful posts among the most successful users seldom use these techniques. This is in contrast to what was observed during the early years of these services [10, 17]. One reason may be that these users already have a big following and do not gain much from adding visible meta information. Another reason may be that some users tag other users directly in the media file themselves. Here, we did not try to determine the underlying factor but simply note that the behavior of this user group has changed while remaining highly successful in attracting engagement.

3.4 Impact of Uploader Characteristics

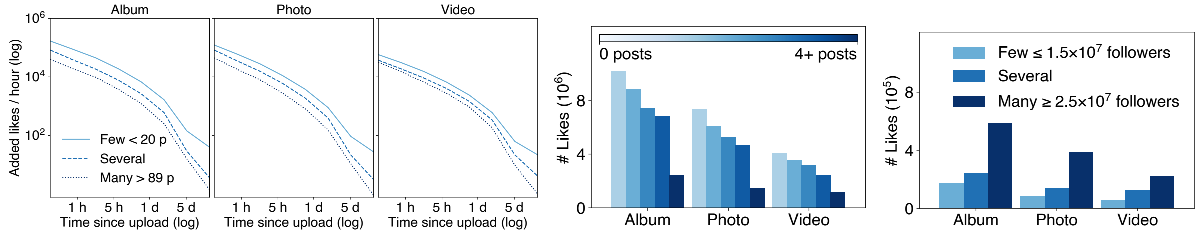
Uploader characteristics play a significant role in shaping users' interactions on social media platforms. Here, we focus on the influence of three key aspects: post frequency, the uploader's network size, and the impact of prior uploads on post interactions.

Frequency of Posts: Figures 7a and 8a show the temporal impact of the number of prior posts uploaded during the 7 days leading up to each post. We again define the categories to contain approximately the same number of posts. Notably, we observe that users who upload fewer than 20 posts the week prior to the post (i.e., less than 3 posts/day on average) tend to receive more interactions on their posts compared to users with higher posting frequencies.

Looking at the uploaders with many posts (>89 posts/week), we observe some differences depending on if we consider likes (shallow) or comments (deep) interactions. For example, these posts receive the least number of additional likes, suggesting that the likes are distributed among their numerous uploads but receive more (early) comments than the category of posters with an intermediate number of weekly posts. These trends apply to all post types.

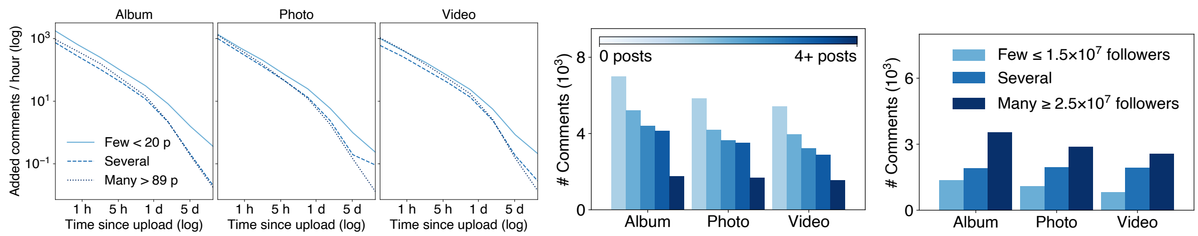
Statistical tests have shown significant differences between all pairwise comparisons of medians and distributions. However, bootstrapping indicated no significant difference in comments for photo posts with several and many previously uploaded posts. This result may be attributed to the observed trend among photos in Figure 8a, which does not apply to other post types.

The higher interaction rates with posts that have seen few prior posts in the prior 7 days also hold true when going down to very few weekly posts. This is illustrated in Figures 7b and 8b, where we show the average impact as seen for the posters with fewer uploaded posts within the 7 days prior to the post. In general, these results suggest a negative correlation between previous uploads and the interaction levels that a post achieves.



(a) Temporal impact of uploaded posts 7 days prior **(b) Average based on prior posts** **(c) Average based on # followers**

Figure 7: Differences in shallow interaction level (i.e., number of likes) for posts with different uploader characteristics.



(a) Temporal impact of uploaded posts 7 days prior **(b) Average based on prior post** **(c) Average based on # followers**

Figure 8: Differences in deep interaction level (i.e., number of comments) for posts with different uploader characteristics.

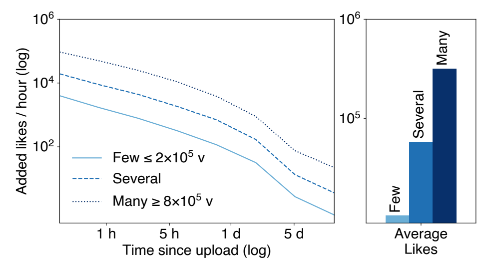
Number of Followers: Figures 7c and 8c illustrate the average interactions that albums, photos, and videos receive based on the number of followers the uploader has at the time of upload. Like past comparisons, we split the set of posts into three size-based categories containing approximately the same number of posts. Perhaps not unexpectedly, we observe a clear correlation between the number of followers and the number of interactions, with posters with more followers generally receiving more interactions. This highlights the value of a larger social network. The results are supported by statistical tests, which indicate a significant relationship between the size of the number of followers and both types of interactions (likes and comments) in terms of median, mean, and distribution (p-values < 10^{-62}).

4 VIEW-BASED ANALYSIS (VIDEO ONLY)

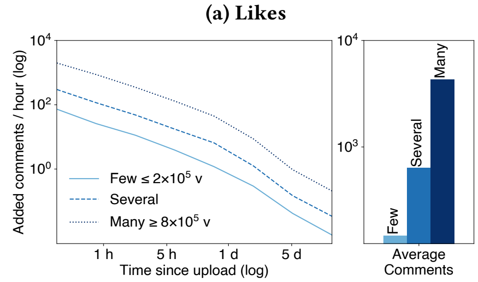
In this section, we look closer at the relationship between video views and other user interactions.

Views Impact on Likes and Comments: Figure 9 shows the temporal and average impact that the number of views has on the number of likes and comments that a video post obtains. Comparing the view-based categorization of posts (with the three categories defined to contain a similar number of posts), there is a noticeable connection between views and other user interactions. For example, a higher number of views correlates with an increase in interactions. This is not surprising and helps motivate why the number of likes and comments may serve as a good proxy for users watching a video (or post in general), even when the number of views is not known (as is the case for photos and albums). Having said that, in the following, we take advantage of the video view counts to study the impact of both post and upload characteristics.

Impact of Post Characteristics on Views: Figure 10 illustrates the impact that the key post characteristics word count, number of hashtags, and number of mentions have on the average number of views. For simplicity, we use the same category limits as in Figures 5 and 6. Similar to what was observed for likes (and to some extent comments), the posts with concise descriptions, containing fewer words, tend to receive the highest number of views. Furthermore, posts without hashtags and mentions in their descriptions attracted the highest user engagement. Statistical tests affirmed these differences, with significant p-values noted: $3.14 \cdot 10^{-66}$ for words, $6.81 \cdot 10^{-33}$ for hashtags, and $3.06 \cdot 10^{-220}$ for mentions.

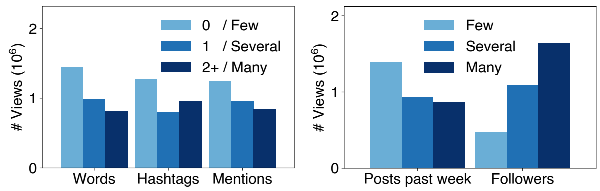


(a) Likes



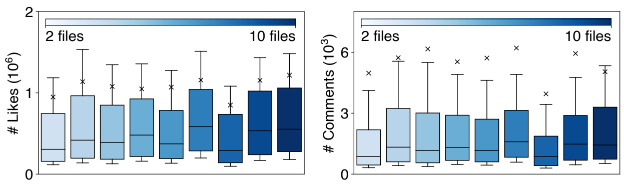
(b) Comments

Figure 9: Temporal interaction rates and average number of likes and comments for posts containing videos that achieved different number of views (v: views).



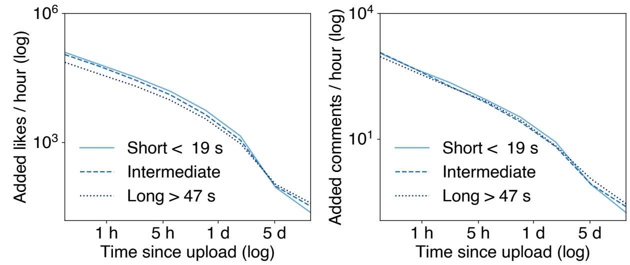
(a) Post characteristics **(b) Uploader characteristics**

Figure 10: Impact of the post- and uploader characteristics on the average number of video views.



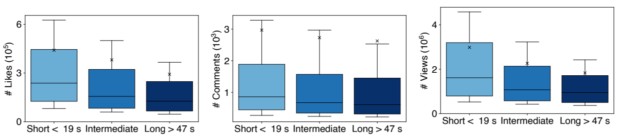
(a) Likes **(b) Comments**

Figure 11: Interactions with posts containing albums with different number of media files.



(a) Likes **(b) Comments**

Figure 12: Temporal interaction rates with posts containing videos of different duration (s: seconds).



(a) Likes **(b) Comments** **(c) Views**

Figure 13: Engagement with videos with different durations.

Impact of Uploader Characteristics on Views: Figure 10b summarizes the impact of uploader characteristics on the average number of views, categorized by follower count and the number of posts uploaded in the previous 7 days, utilizing consistent limits for comparable categories as previously employed.

We again observe a positive correlation between the number of followers that an uploader has and the number of views that their videos receive and a negative correlation between the number of views and the number of posts uploaded in the previous 7 days. These trends are consistent with what we observed for likes but not always comments (for which the "many" category of the number of prior uploads sometimes outperforms the "several" category). Finally, our statistical tests supported the findings in Figure 10b, indicating significant differences (with the highest p-value being $1.55 \cdot 10^{-52}$ for posts in the past week and $3.81 \cdot 10^{-235}$ for followers).

5 MEDIA-BASED ANALYSIS

We next performed a media-based analysis of all files included in posts by the top-100 most followed users. Focusing on the top-100 users allowed us to thoroughly collect and analyze all media files posted by these prominent accounts across the entire year. We ensured consistency by replicating all experiments conducted on the top-1000 dataset, confirming that the subset reveals identical trends and conclusions in both our post-based (Sec. 4) and view-based (Sec. 5) analyses. Notably, the top-100 users generally garnered more interactions and used briefer descriptions compared to their counterparts in the top-1000 dataset. Subsequently, our analysis proceeds to utilize extracted features that either were media-type specific or that captured content-specific characteristics of the individual media within the posts.

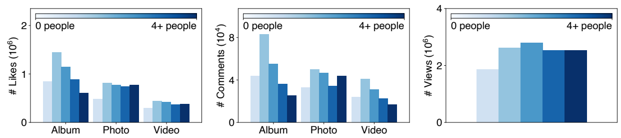
5.1 Features Specific to Media Type

Consider first two features specific to albums and videos.

Albums: A distinguishing feature of an album is the number of files in each album. Figure 11 shows the average engagement with posts containing albums with different numbers of media files (ranging from 2 to 10 files). Here, we used boxes to show the 25-75th percentiles, whiskers for the 20-80th percentiles, lines for the medians, and markers to show the averages. The graphs illustrate that albums with 2 media files generally attract the least interactions and that albums with more files tend to receive more likes (Figure 11a) and comments (Figure 11b), although the trend is much less pronounced for comments. Statistical tests revealed significant differences in medians, means, and distribution between albums containing 2 and 10 media files (p-values $< 2.11 \cdot 10^{-44}$), suggesting that popular users favor larger albums. These differences imply potential advantages for larger albums in reaching more users, possibly through the Instagram algorithm or shares.

Videos: A distinguishing feature of videos is their play duration. Figure 12 illustrates engagement patterns across videos of varying durations, segmented into comparable sets. While most clear for likes, we see that short videos initially garner both the most likes and the most comments, followed by intermediate-duration videos. However, longer videos gain more likes (and comments) at a later stage, reflecting sustained interest. This temporal pattern aligns with user behavior: shorter videos attract immediate engagement, while longer ones entice viewers over time. Despite this, shorter videos consistently boast the highest average likes, comments, and views, as depicted in Figure 13.

Statistical tests confirm significant differences in interactions across video durations, except for comments between intermediate and long videos. (We note that also these exceptions would have been considered significant at the 0.05 level, instead of our chosen 0.005 level, as they all scored p-value in the range 0.01-0.03.) These findings suggest that while longer videos acquire more interactions later on, shorter videos generally perform better overall. These results imply a preference among popular users for shorter videos or suggest that shorter content has wider potential engagement.



(a) Likes (b) Comments (c) Video views

Figure 14: Average engagement with posts with different number of visible people in the media.

5.2 Content Features

We next look closer at the content of the media files themselves, specifically the images presented to the users. To allow comparisons between the different forms of posts, we let the first image of an album represent an album, and the first frame of a video represent a video, and then we use machine learning techniques to extract features from these images that help classify each media file.

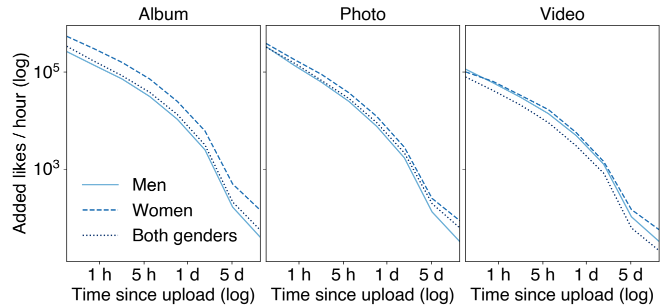
Number of people in image/frame: To detect people in images, the library Detectron2 [39] was used. Detectron2, created by Facebook, is an open-source framework that provides a large number of pre-trained state-of-the-art object detection and segmentation models. This analysis implemented their Faster R-CNN model with a Feature Pyramid Network (FPN) [21] backbone built on ResNet 101 [13]. The model and backbone were chosen for their good box average precision and fast predictions compared to other models [15, 28]. The model was implemented with default settings and can detect objects such as people, cars, cats, and dogs. However, for this analysis, only people were detected.

Figure 14 shows the average engagement based on the number of identified people in media files. Albums with a single identified person draw the highest interactions, while those featuring multiple individuals receive fewer. Interestingly, albums with people generally attract more interactions compared to those without. Statistical tests validated these trends in medians and distributions, although the average number of comments did not significantly differ between albums without people and those with two individuals.

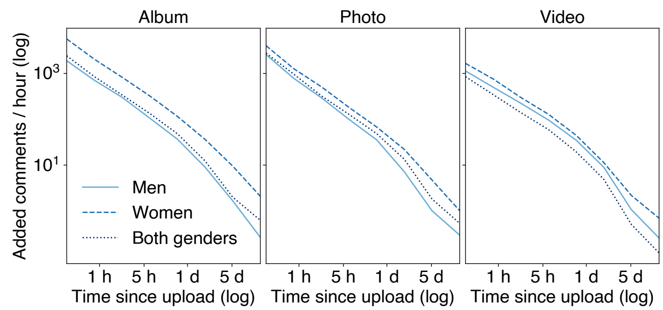
Similar trends are noticed for videos in Figure 14b but not for likes (Figure 14a) or views (Figure 14c). While comments in videos show a negative correlation, statistical tests did not support this finding across all comparisons. For the other cases (photo likes, video likes, photo comments, and video views), the presence of people in media files generally increases received interactions compared to posts without people, including with larger numbers of detected individuals.

Gender: To extract facial attributes in images, the library DeepFace [31] was used. DeepFace is an open-source framework that provides state-of-the-art facial recognition models and a facial attribute analysis module for the attributes age, gender, emotion, and race [32]. This analysis only analyzed the facial attributes of age, gender, and emotion. For the facial attribute analysis module, we employed a re-implementation of RetinaFace [7], chosen for its superior performance among various face detection models [32]." Facial attribute models, excluding emotion, were constructed using VGG-Face [26] due to its exceptional performance [14, 38]. Here, the output layers varied based on the predictable facial features.

Figure 15 shows the temporal interaction patterns based on detected genders using the RetinaFace model (minimum 86% accuracy). We note that files with exclusively women present receive

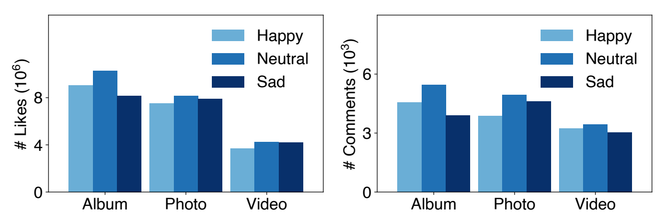


(a) Likes



(b) Comments

Figure 15: Temporal interaction rates with posts containing different genders.



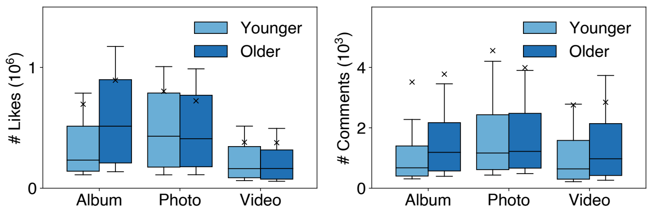
(a) Likes (b) Comments

Figure 16: Average engagement with posts containing people showing different emotions.

the highest interactions across all post types, typically followed by files where both genders appear. However, video files with only men present sometimes initially gain more engagement, particularly in Figure 15a (likes), but this engagement decreases over time. Similarly, statistical tests reveal significant differences among gender categories for albums and photos (with a significance threshold of 0.005) but not always for videos (e.g., p-values of 0.06 with Kruskal-Wallis test, 0.15 with Dunn's test, and $8.44 \cdot 10^{-4}$ with Mann-Whitney).

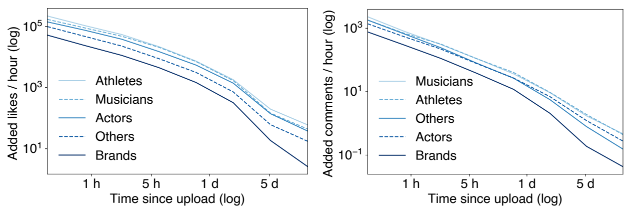
Emotions: For the emotion model, we used a CNN to detect one of seven emotions: angry, disgust, fear, happy, sad, surprise, or neutral. However, only happy, neutral, and sad were analyzed, thus they consisted of the majority of predicted emotions.

Figure 16 shows the average impact of predicted emotions-—happy, neutral, and sad-—where each emotion is detected with an accuracy of at least 80%. We note that neutral facial expressions receive the highest likes and comments across all media types, followed by



(a) Likes **(b) Comments**

Figure 17: Engagement with posts containing people of different age categories: young vs. old.



(a) Likes **(b) Comments**

Figure 18: Temporal interaction rates with posts uploaded by different categories of posters.

happy expressions in albums and sad expressions in photos. However, the observed differences between emotions generally had no to little statistical support, indicating that the type of facial expressions in media files is not a significant contributor to whether a post acquired many interactions when studied at an aggregate level (although we expect that it may play a role for individual posts).

Age: Figure 17 shows the average impact of predicted ages, categorized into younger (0-34) and older (35+) groups based on detected ages between 0 and 100 years old. We note that older individuals tend to receive more interactions in albums and videos, while younger individuals attract more interactions in photos. Whiskers representing the 20^{th} and 80^{th} percentiles support this trend, particularly in albums and videos, indicating greater popularity among older individuals. For albums, this bias was statistically supported for all tests except the means of comments. However, statistical tests did not show significant differences in medians of likes for photos (0.08) and videos (0.19). For videos, distributions of likes (0.04) and means of comments also lacked statistical significance.

6 CATEGORICAL ANALYSIS

Finally, we examine the user engagement with different uploader categories: actors, athletes, brands, musicians, and others.

High-level Temporal Analysis: Figure 18 shows the temporal interaction rate associated with each uploader category. We make several observations. First, athletes receive the highest number of likes, followed by musicians, while musicians garner more comments than athletes. Second, brands consistently receive the fewest interactions across all categories. They also show a sharper decline in interactions over time compared to other categories, potentially due to a higher volume of uploaded posts.

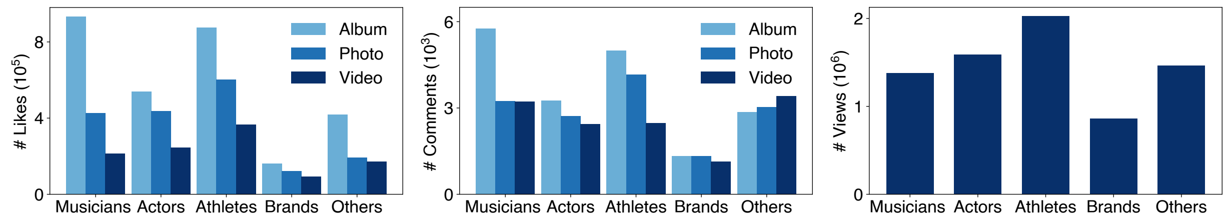
Post-type Comparison: We next consider how user engagement varies across post types within each category. These results are summarized in Figure 19. Looking at likes and comments, it is clear that albums generally receive the most attention across all categories, except for comments for the “others" category, where models and influencers dominate video comments. Surprisingly, the “others" category receives fewer views despite its substantial comment count. This skewness may be due to the Instagram algorithm internals or multiple comments posted by a selected few users. Second, except for albums, athletes, followed by actors, received the most likes and video views, with musicians excelling in album likes and comments. Third, comparing the relative ordering of likes for videos (athletes, actors, musicians) and the order of comments for the same categories (musicians, actors, athletes), we note that it appears that people tend to engage deepest with musicians' videos and have the most frequent but shallow interactions (likes and views) with the athletes' videos. It can also be argued that the large number of likes to the musicians' albums can be seen as a type of deeper interaction. Finally, we note that all pairwise tests (Dunn and Mann-Whitney U) within a category showed significant differences in the amount of likes except between photos and videos of the “others" category. For comments, the least significant difference (and only that did not show significance for at least one of the pairwise tests) was again between photos vs. videos, but this time for musicians.

7 RELATED WORK

This study aligns with prior research exploring user engagement dynamics on social media platforms, particularly Instagram. Previous works like those by Jang et al. [17] and Huang et al. [16] have delved into Instagram engagement. Notably, Vassio et al.'s work [35, 36] stands out for its temporal analysis of engagement using data from Crowd Tangle. While similar in focus, their study examines Italian influencers' engagement dynamics, differing from our broader content analysis across various influencer categories, and they do not compare the dynamics across content types. Other temporal dynamics studies in the literature concentrate on platforms like Facebook [25, 30], Twitter [11, 27, 29], or YouTube [5], diverging from our Instagram-centered investigation encompassing different content types.

Huang et al.'s study [16] examined posts' and uploaders' characteristics on Instagram, exploring the optimal publish time for increased interactions. They discovered that posts uploaded after noon received more interactions, regardless of the day. However, their focus differs from our work as they did not analyze the temporal dynamics of user engagement over time, setting their study apart from our temporal-centric analysis.

Part of our work also relates to works studying how image and video features influence post popularity. Here, Bakhshi et al.'s pioneering work [3] revealed that photos with faces on Instagram garnered over 30% more interactions, although faces' attributes showed no significant impact. Jang et al.[18] focused on teen and adult engagement, observing higher interactions among teens despite fewer uploads compared to adults. Mazloom et al.[24] explored brand media files, noting higher engagement in images combining people and the brand's product. Lindell et al. [22] found pose orientation affected engagement, with left cheek poses receiving more likes. Our work stands distinct by examining varied post types,



(a) Likes **(b) Comments** **(c) Video views**

Figure 19: Average engagement with each post type for each category.

assessing and statistically quantifying differences in the engagement dynamics among them, and by considering diverse uploader categories.

Other related works have trained models for predicting the popularity of current or future posts. Examples of such models included [12, 40, 41] or excluded [6] the impact of media content. For example, Zhang et al. [40] proposed a model to predict the popularity of posts based on only the image and description of each photo post. Using the model, they found that the unpopular images contained posters and food. However, descriptions were seen as a more reliable estimate for their prediction model. Gayberi and Oguducu [12] created a model to predict the popularity of posts based on posts' and uploaders' characteristics. Additionally, the media characteristics were accounted for by extracting features such as people, vehicles, and whether the environment was indoor or outdoor. Similarly, Tricomi et al. [34] trained predictive models on image and caption embeddings, metadata, and engagement metrics. By analyzing the feature importance and correlation, the study showed that likes on Instagram are mainly driven by images, while comments are primarily stimulated by captions. We note that both our goal and methodology is substantially different than those building predictive models.

8 CONCLUSIONS

This paper has presented a temporal analysis emphasizing the comparisons between interactions seen with albums, photos, and videos, capturing the intriguing dynamics and preferences observed among different content types on Instagram. Our analysis revealed several statistically significant patterns and trends in user engagement across these post types.

Despite the initial burst of engagement seen across all post types shortly after upload, albums maintained relatively higher interaction rates over an extended period compared to photos and videos. This extended engagement with albums suggests a sustained user interest and continued appreciation for this content format beyond the immediate upload period.

Further investigation into the factors influencing engagement highlighted the impact of content composition. For example, shorter descriptions with fewer words and minimal use of hashtags and mentions were associated with higher interaction levels across all post types. This suggests that concise, focused content tends to attract more attention and engagement from users, irrespective of the content format.

Moreover, the analysis highlighted the relative differences in user engagement depth between shallow interactions (likes) and deeper interactions (comments). Across all post types, the disparity between likes and comments remained consistent, signifying a trend where users tend to engage more frequently with shallow interactions, and this disparity persisted over time.

Finally, our categorical analysis, revealing distinctive engagement patterns among various categories, such as musicians and brands, underscores the broader significance of tailoring content strategies based on specific user segments.

One limitation of this study is that we focus on the top 1K followed Instagram accounts. As one line of future work, the extensibility of the reported insights to the broader population of Instagram accounts can be studied. Despite this limitation, the observed temporal nuances highlight the importance of understanding not only the volume of interactions but also their duration and sustainability over time. Content creators and marketers can leverage these insights to tailor their content strategies, emphasizing the creation of content formats that not only generate immediate engagement but also sustain user interest over an extended period.

REFERENCES

[1] [n. d.]. Starngage. https://starngage.com/plus/en-us. Accessed: 2023-03-31.

[2] Kholoud Khalil Aldous, Jisun An, and Bernard J. Jansen. 2019. View, Like, Comment, Post: Analyzing User Engagement by Topic at 4 Levels across 5 Social Media Platforms for 53 News Organizations. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 13. 47–57.

[3] A. Bakhshi, D. Shamma, and E. Gilbert. 2014. Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proc. CHI.* 965–974.

[4] Y. Borghol, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti. 2012. The untold story of the clones: Content-agnostic factors that impact YouTube video popularity. In *Proc. ACM KDD.* 1186–1194.

[5] Y. Borghol, S. Mitra, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti. 2011. Characterizing and modelling popularity of user-generated videos. *Performance Evaluation* 68, 11 (2011), 1037–1055.

[6] S. Carta, AS. Podda, DR. Recupero, R. Saia, and G. Usai. 2020. Popularity prediction of instagram posts. *Information* 11, 9 (2020), 453.

[7] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou. 2020. Retinaface: Single-shot multi-level face localisation in the wild. In *Proc. IEEE/CVF CVPR.* 5203–5212.

[8] Olive Jean Dunn. 1964. Multiple comparisons using rank sums. *Technometrics* 6, 3 (1964), 241–252.

[9] B. Efron. 1992. *Bootstrap methods: another look at the jackknife.* Springer.

[10] E. Ferrara, R. Interdonato, and A. Tagarelli. 2014. Online popularity and topical interests through the lens of instagram. In *Proc. ACM Conference on Hypertext and social media.* 24–34.

[11] K. Garimella and R. West. 2021. Evolution of Retweet Rates in Twitter User Careers: Analysis and Model. In *Proc. AAAI ICWSM*, Vol. 15. 1064–1068.

[12] M. Gayberi and S.G. Oguducu. 2019. Popularity prediction of posts in social networks based on user, post and image features. In *Proc. International Conference on Management of Digital EcoSystems.* 9–15.

[13] K. He, X. Zhang, S. Ren, and J. Sun. 2016. Deep residual learning for image recognition. In *Proc. IEEE CVPR.* 770–778.

[14] G. Huang, M. Mattar, T. Berg, and E. Learned-Miller. 2008. Labeled faces in the wild: A database forstudying face recognition in unconstrained environments. In *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition.*

[15] J. Huang, V. Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, et al. 2017. Speed/accuracy trade-offs for modern convolutional object detectors. In *Proc.*

IEEE CVPR. 7310–7311.
[16] J. Huang, C. Wang, M. Su, Q. Dai, and MZA. Bhuiyan. 2018. Inspecting influences on likes and comments of photos in instagram. In *Proc. IEEE SmartWorld.* 938–945.
[17] JY. Jang, K. Han, and D. Lee. 2015. No reciprocity in" liking" photos: analyzing like activities in instagram. In *Proc. ACM conference on hypertext and social media.* 273–282.
[18] JY. Jang, K. Han, D. Lee, H. Jia, and P. Shih. 2016. Teens engage more with fewer photos: temporal and comparative analysis on behaviors in instagram. In *Proc. ACM Conference on hypertext and social media.* 71–81.
[19] Cheonsoo Kim and Sung-Un Yang. 2017. Like, comment, and share on Facebook: How each behavior differs from the other. *Public Relations Review* 43, 2 (2017), 441–449.
[20] W.H. Kruskal and W.A. Wallis. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association* 47, 260 (1952), 583–621.
[21] TY. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie. 2017. Feature pyramid networks for object detection. In *Proc. IEEE CVPR.* 2117–2125.
[22] Annukka K. Lindell. 2019. Left cheek poses garner more likes: the effect of pose orientation on Instagram engagement. *Laterality* 24, 5 (2019), 600–613. https://doi.org/10.1080/1357650X.2018.1556278 PMID: 30526363.
[23] H.B. Mann and D.R. Whitney. 1947. On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics* (1947), 50–60.
[24] M. Mazloom, R. Rietveld, S. Rudinac, M. Worring, and W. Van Dolen. 2016. Multimodal popularity prediction of brand-related social media posts. In *Proc. ACM Multimedia.* 197–201.
[25] A. Mohammadinodooshan and N. Carlsson. 2023. Effects of Political Bias and Reliability on Temporal User Engagement with News Articles Shared on Facebook. In *Proc. PAM.* Springer, 160–187.
[26] O. Parkhi, A. Vedaldi, and A. Zisserman. 2015. Deep face recognition. *British Machine Vision Association* (2015).
[27] Jürgen Pfeffer, Daniel Matter, and Anahit Sargsyan. 2023. The Half-Life of a Tweet. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 1163–1167.
[28] V. Pham, C. Pham, and T. Dang. 2020. Road damage detection and classification with detectron2 and faster R-CNN. In *Proc. IEEE Big Data.* 5592–5601.
[29] Arthi Ramachandran, Lucy Wang, and Augustin Chaintreau. 2018. Dynamics and Prediction of Clicks on News from Twitter. In *Proceedings of the 29th on Hypertext and Social Media* (Baltimore, MD, USA) *(HT '18)*. Association for Computing Machinery, New York, NY, USA, 210–214. https://doi.org/10.1145/3209542.3209568
[30] F. Sabate, J. Berbegal-Mirabent, A. Cañabate, and P. Lebherz. 2014. Factors influencing popularity of branded content in Facebook fan pages. *European management journal* 32, 6 (2014), 1001–1011.
[31] S. Serengil. [n. d.]. DeepFace: A Lightweight Face Recognition and Facial Attribute Analysis (Age, Gender, Emotion and Race) Library for Python. https://github.com/serengil/deepface. Accessed 2023-05-22.
[32] S. Serengil and A. Ozpinar. 2021. Hyperextended lightface: A facial attribute analysis framework. In *Proc. International Conference on Engineering and Emerging Technologies.* 1–4.
[33] CrowdTangle Team. [n. d.]. CrowdTangle. https://www.crowdtangle.com/. Accessed: 2023-03-01.
[34] Pier Paolo Tricomi, Marco Chilese, Mauro Conti, and Ahmad-Reza Sadeghi. 2023. Follow Us and Become Famous! Insights and Guidelines From Instagram Engagement Mechanisms. In *Proceedings of the 15th ACM Web Science Conference 2023* (Austin, TX, USA) *(WebSci '23)*. Association for Computing Machinery, New York, NY, USA, 346–356. https://doi.org/10.1145/3578503.3583623
[35] L. Vassio, M. Garetto, C. Chiasserini, and E. Leonardi. 2021. Temporal dynamics of posts and user engagement of influencers on facebook and instagram. In *Proc. IEEE/ACM ASONAM.* 129–133.
[36] Luca Vassio, Michele Garetto, Emilio Leonardi, and Carla Fabiana Chiasserini. 2022. Mining and modelling temporal dynamics of followers' engagement on online social networks. *Social Network Analysis and Mining* 12, 1 (2022), 96.
[37] K. Wang, P. Wang, X. Chen, Q. Huang, Z. Mao, and Y. Zhang. 2020. A feature generalization framework for social media popularity prediction. In *Proc. ACM Multimedia.* 4570–4574.
[38] Lior Wolf, Tal Hassner, and Itay Maoz. 2011. Face recognition in unconstrained videos with matched background similarity. In *Proc. IEEE CVPR.* 529–534.
[39] Y. Wu, A. Kirillov, F. Massa, WY. Lo, and R. Girshick. 2019. Detectron2. https://github.com/facebookresearch/detectron2.
[40] Z. Zhang, T. Chen, Z. Zhou, J. Li, and J. Luo. 2018. How to become Instagram famous: Post popularity prediction with dual-attention. In *Proc. IEEE Big Data.* 2383–2392.
[41] A. Zohourian, H. Sajedi, and A. Yavary. 2018. Popularity prediction of images and videos on Instagram. In *Proc. International Conference on Web Research.* 111–117.