Amsterdam University Press

COMPUTATIONAL COMMUNICATION RESEARCH 4.1 (2022) 68-97
HTTPS://DOI.ORG/10.5117/CCR2022.2.002.LU

ARTICLE

The Pervasive Presence of Chinese Government Content on Douyin Trending Videos

Yingdan Lu
Stanford University

Jennifer Pan
Stanford University

Abstract
As audiences have moved to digital media, so too have governments around the world. While previous research has focused on how authoritarian regimes employ strategies such as the use of fabricated accounts and content to boost their reach, this paper reveals two different tactics the Chinese government uses on Douyin, the Chinese version of the video-sharing platform TikTok, to compete for audience attention. We use a multi-modal approach that combines analysis of video, text, and meta-data to examine a novel dataset of Douyin videos. We find that a large share of trending videos are produced by accounts affiliated with the Chinese government. These videos contain visual characteristics designed to maximize attention such as high levels of brightness and entropy and very short duration, and are more visually similar to content produced by celebrities and ordinary users than to content from non-official media accounts. We also find that the majority of videos produced by regime-affiliated accounts do not fit traditional definitions of propaganda but rather contain stories and topics unrelated to any aspect of the government, the Chinese Communist Party, policies, or politics.

Keywords: China, political communication, Video, Douyin, Tiktok, Computer Vision

Introduction

As audiences have moved to digital and social media platforms, governments, and in particular authoritarian regimes, have also increased their

© Yingdan Lu & Jennifer Pan

presence on these platforms. The goal is to reach audiences. However, in order to capture audiences on increasingly popular platforms dominated by entertainment such as Instagram and TikTok, authoritarian regimes have found themselves in competition with celebrities, entertainers, and influencers.

While autocrats can and do censor politically relevant information, it is infeasible for autocrats to remove information in order to compete for attention. Authoritarian regimes can and do censor social media platforms by shutting them down or blocking access to them, but this would remove the platform completely rather than enable regimes to compete for audience attention on them (Roberts, 2018). Moreover, vast volumes of social media content make it difficult and costly for authoritarian regimes to selectively remove content in order to capture attention. Instead, a growing body of literature in political science and communication has revealed two primary ways in which autocratic regimes compete for attention. The first is by using bots, trolls, computational propaganda, and other inauthentic and often deceptive means to fabricate the appearance of support and interest (Bjola, 2017; Keller et al., 2020; King et al., 2017; Marwick & Lewis, 2017; Sanovich, 2017; Woolley & Howard, 2017). The second is by using content distribution strategies such as clickbait headlines to increase clicks, views, and likes (Lu & Pan, 2020). In this paper, we aim to contribute to and advance this body of research by examining what information control strategies authoritarian governments are using to compete for audience attention on digital media platforms. We do this by collecting and analyzing videos from Douyin (抖音), a video-sharing social media platform similar to TikTok and owned by the same parent company.

Douyin and TikTok began as mobile apps that allowed users to make and share short videos. Unlike Facebook and Twitter, content only appears as videos that users scroll up and down to see more of, not as a mix of text, images, and videos. Douyin videos are tall (filling the entire mobile screen) rather than square or landscape as videos are on Youtube, Snapchat, and Instagram's stories (for a comparison of Facebook, Twitter, YouTube, and Douyin homepages, see Figure A1 in the Appendix). While Douyin, like other social media platforms, encourages users to engage with other users, having a network—of friends or followers —is not essential to the Douyin experience. Douyin fills users' feeds before they follow a single person, and even after following, users' feeds do not prioritize or only contain content from those they follow.

We analyze content from a video-sharing platform because despite the rapid growth of these platforms—Douyin was founded in September 2016 and had over 600 million daily active users as of January 2021[1]—and despite increasing government presence on them (Confessore & Wakabayashi, 2017), few academic studies have focused on such platforms. Research in communication suggests that the characteristics and multi-modalities of video may be particularly conducive for persuasion. The multiple modalities of video can more comprehensively stimulate human cognitive systems (Marmolin, 1992; Sundar, 2000). Because of the visual modality, video can portray reality more directly than text, enabling faster information processing and deeper encoding in memory, and inducing, as a result, more powerful emotional responses and behaviors (Barry, 1997; Graber, 1990; Iyer & Oldmeadow, 2006; Messaris & Abraham, 2001; Newhagen & Reeves, 1992; Powell et al., 2015; Stenberg, 2006; Zillmann et al., 2001). Compared to text and other single-modality content, visual information has been found to be more effective in attracting attention (Dahmen, 2012; Pieters & Wedel, 2004), generating salience (Powell et al., 2019), heightening credibility (Hameleers et al., 2020), improving trust (Hoogeveen, 1997; Sundar, 2008; Sundar & Limperos, 2013), and increasing audience reach and engagement (Goel et al., 2016; Li & Xie, 2020; Pancer & Poole, 2016).

One of the reasons for the limited number of studies in this area is that there are no established norms for how video should be analyzed as data for quantitative social science research.[2] While research on the effects of video in various forms is well established,[3] the analysis of large-scale image datasets to answer social science questions is a recent phenomenon (Joo & Steinert-Threlkeld, 2018; Mebane et al., 2018; Peng, 2018; Peng & Jemmott, 2018; Reeves et al., 2019; Sobolev et al., 2020; Williams et al., 2020; Xi et al., 2020; Zhang & Pan, 2019). Social science research that quantitatively analyzes video data at scale is even less common (Boussalis & Coan, 2020; Nyhuis et al., 2021).

Thus, a contribution of this paper is the analysis of video data and the multi-modal approach we take, which builds on the recognition that video data rarely exists in isolation. Video data, and image data more broadly, are often accompanied by meta-data and text data (Steinert-Threlkeld, 2019; Zhang & Pan, 2019). Rather than discard textual data, the multi-modal approach we use analyzes video data in conjunction with meta-data and textual data to maximize efficiency and accuracy.

We organize our study around the following research questions:

- **RQ1**: What share of trending videos on Douyin comes from regime-affiliated accounts?
- **RQ2a**: Do videos from regime-affiliated accounts exhibit visual features shown to maximize attention?

- **RQ2b**: Do videos from regime-affiliated accounts exhibit levels of attention maximizing features more similar to videos from celebrities and ordinary users than other media accounts?
- **RQ3**: What share of content from regime-affiliated accounts is propaganda?

These questions allow us to describe the presence and strategies of the Chinese regime on Douyin and contribute to research on digital media and authoritarian politics.

To answer these questions, we collected a novel dataset of 50,813 videos from the Trending Videos page of Douyin between March and June of 2020. We use the meta-data associated with each video to identify the type of account producing the content, differentiating between *regime-affiliated accounts*, which include accounts run by the Chinese Communist Party (CCP), the Chinese government, and official state media outlets, and accounts of celebrities, ordinary users, and non-official media outlets. We sample frames from each video to analyze visual features, such as brightness, warm color dominance, and face presence, known to elicit audience attention. We use the textual descriptions associated with trending topics to classify videos into different content categories, focusing on differentiating between videos containing propaganda, which we define broadly as all content related to the Chinese regime and denigrating outgroups, and other content unrelated to the CCP, the Chinese government, government policies, or leaders and officials, such as entertainment and human interest stories. Finally, we take advantage of the fact that topics trend on the Douyin not only because of organic virality to show that regime-affiliated accounts are likely producing non-propaganda content intentionally.

Data

Douyin users represent nearly 60% of China's 989 million internet users (CNNIC, 2021). Importantly, Douyin's user base is diverse, including people from wealthy, urban centers like Shanghai and those in impoverished counties in central and western China.[4] Douyin is owned by the Chinese technology company ByteDance, which also owns TikTok. While Douyin and TikTok are similar in their design and functionality, the two platforms are not identical in user interface or features. For example, for content discovery through the "Discover" page, Douyin users see recommended video hashtags and lists of trending buzzwords in textual form, while TikTok users see trending hashtags with thumbnails of videos below each hashtag (see Figure A2 in Appendix). Importantly, Douyin and Tiktok differ in terms of political governance. Douyin is available in mainland China while Tiktok is blocked. This means Douyin is bound by China's internet regulations and must adhere to the information control directives of the Chinese government.[5] The Chinese government has emphasized the need to use Douyin and other short-form video sharing platforms to communicate directly with the public (Chen et al., 2020). Starting from 2018, the Cyberspace Administration of China has worked with Douyin to set up accounts for official media and local governments.[6] As of the end of 2018, over 250,000 short videos had been created by CCP or government accounts and over 150,000 short videos had been created by official accounts on Douyin, resulting in 6.9 billions likes.[7] As of June 2020, the Chinese government reported that 25,313 regime-affiliated accounts from different administrative regions and levels had been established on Douyin (CNNIC, 2020).

Similar to other popular Chinese social media platforms like Weibo, Douyin not only allows users to browse content recommended by the platform algorithm but also allows users to easily see trending discussions. When a user opens the Douyin app, the homepage automatically plays recommended videos (Panel A of Figure 1). Users can scroll down the homepage for more videos or click on the search button on the top-right corner of the homepage (circled in red in Panel A of Figure 1). Clicking on the search button will redirect users to a "Discover" page (Panel B of Figure 1), which shows the Douyin Trending topics list (outlined in red in Panel B). On the Discover page, users can scroll down to see a full list of 50 hashtag-like topics, updated in near real time, that Douyin features as Trending topics, or click the "See the complete Trending page" button at the bottom to go to the Trending page (Panel C of Figure 1) where all 50 topics are listed. Users can see the number of videos Douyin has associated with each topic and the popularity of the topic. When users click on a topic, a video associated with that topic will play automatically (example shown in Panel D of Figure 1). Similar to the homepage, users could scroll down to see all videos in the topic. Once all videos in a topic have been viewed, users are redirected to another Trending topic. In Panel C of Figure 1, the entertainment topic "People sent Li Huanying's photos to Jia Ling (网友把李焕英照片送给贾玲)" is outlined in red. Clicking on this topic automatically plays a video (Panel D) created by *Hubei Daily*, the official newspaper in Hubei province.

Between March 18, 2020, and June 17, 2020, we collected all topics listed on the Douyin Trending page on an hourly basis. Then, retrospectively, we collected all videos related to each topic.[8] After de-duplication, our

Figure 1 Douyin Trending Page

dataset consists of 50,813 unique videos related to 9,946 topics. For each video, we collected the video itself, account information of the video creator (nickname, verification status and type),[9] account activity and popularity indicators (followers, followings, number of posts, total likes), the associated trending topic, the video description from the creator, video hashtags, creation date and time, duration of the video, and engagement meta-data (likes, comments, shares within the Douyin platform, and shares beyond Douyin to other social media platforms).

We focus on the Trending page because it is one of the main ways social media platforms direct user attention and engagement (Yang & Peng, 2020). As well, the Chinese regime may influence the appearance of topics on the Trending page. For example, in 2018, Douyin and the Ministry of Public Security collaborated to promote the topic "Salute my super hero" (#敬礼，我的　“超级英雄”) to call for support of the Chinese police and army.[10] During the Anniversary of the CCP in 2019, the topic "Red Memory" (红色记忆) was promoted by government departments (e.g., National Health Commission) and official media (e.g., *People's Daily*) on Douyin.[11]

Multi-Modal Approach

Douyin trending videos are accompanied by a wide variety of meta-data (e.g., account information) and textual data (topics, description, hashtags). Rather than discarding these non-video data, we adopt a multi-modal approach (see Figure 2), where we use video data for analysis of video-base features

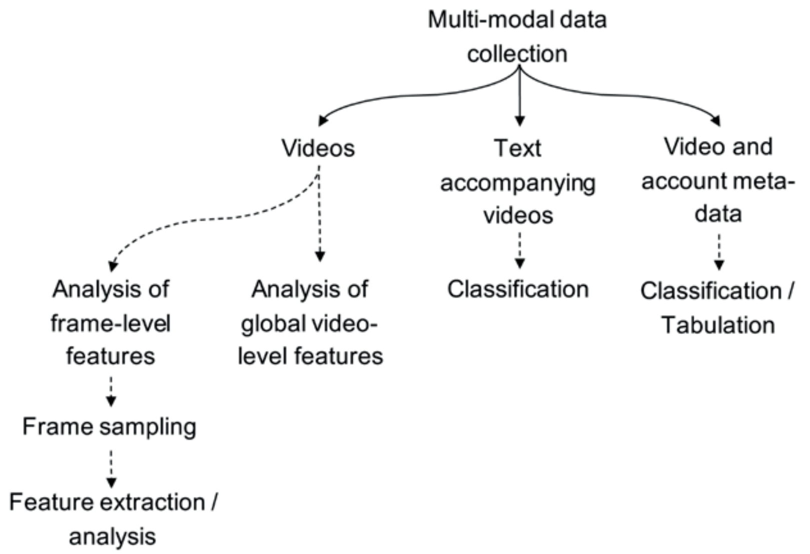

Figure 2 Multi-Modal Approach For Video Analysis

and text and meta-data for classification. In the analysis of the videos, we focus on features that, based on prior research, are likely to maximize attention. This is because our substantive question relates to how government propaganda efforts on social media compete for audience attention. We analyze the text accompanying the video topics to classify the content of video according to our substantive questions, and we analyze video-level and account-level meta-data to classify the type of video creators into those who are affiliated with the CCP regime and those who are not. A limitation of our method is that we do not analyze inter-frame temporal features of videos—for example, motion-related dynamics—or audio features, which we hope to remedy in future research.

Video Features

We analyze frame-level and global video features, which have been identified in prior research as those that maximize attention. Since videos consist of a series of consecutive frames, frame-level features are those that derive information from individual frames. For each video, we read all consecutive frames.[12] To reduce computational cost, we sample every sixth frame from all frames for analysis (Chen et al., 2018; Simonyan & Zisserman, 2014). For each sampled frame, we focus on three visual features: *brightness*, *entropy*, and *color dominance*, which are low-level features, and one higher-level feature, *face presence*.

Brightness: Brightness is fundamental to human visual perception and has been shown to play an important role in attracting attention and affecting visual working memory (Camgöz et al., 2004; Egusa, 1982; Qian et al., 2018). We operationalize brightness by measuring the luminance of pixels in the frame.[13] To calculate the brightness of the video, we take an average of scores across all sampled frames.

Entropy: The concept of entropy comes from information theory, where it refers to the uncertainty inherent in a random variable's possible outcomes (Sethna et al., 2006). Given a random variable X with possible outcomes x_i that occur with probability $P(x_i)$, the entropy of X is defined by the equation known as Shannon's entropy, where:

$$H(X) = -\sum_{i=1}^{n} P(x_i) \log_b P(x_i) \tag{1}$$

In the context of images, entropy can be conceptualized as the heterogeneity of pixels in the image (Brinberg et al., 2021; Reeves et al., 2019; Yang et al., 2019). Prior research shows that higher entropy can influence positive perception and aesthetic evaluation (Rosenholtz et al., 2005; Zheng et al., 2009).

To calculate entropy for a frame, we first convert pixels to grayscale to reduce dimensionality and increase computational speed.[14] Then, using Equation 1 with logarithm to the base 2, we calculate entropy where x is possible grayscale levels $i = 1,\ldots,n$, and $P(x_i)$ is the probability of the grayscale level i in a frame, which is the share of pixels in the frame with grayscale level i. This means a frame with all black pixels would have an entropy score at 0, while a highly textured frame would have higher entropy (8 is the maximum score). To calculate the entropy of a video, we take an average of entropy scores across all sampled frames.

Color Dominance: Besides general image complexity, different colors have been shown to generate different psychological impacts on individual affect and behaviors. For example, warm colors such as reds and yellows are thought to arouse more positive feelings and increase image popularity than cool colors such as greens and blues (Bakhshi et al., 2015; Elliot & Maier, 2014; Goethe & Eastlake, 2006; Peng & Jemmott, 2018). Therefore, we capture the dominance of specific colors from our video data.

For each frame, we extract the RGB color vector (e.g., [255, 0, 0]) from each pixel in a frame and calculate the vector that appears with the highest frequency, which we call the dominant color vector. Following the work of Peng and Jemmott (2018) and Van De Weijer et al. (2007), we match the dominant color vector to one of the 17 base color names (e.g. "red").[15] At the video level, warm color dominance is the proportion of frames dominated by warm colors (red, yellow, orange, maroon, olive), and cold color dominance is the proportion of frames dominated by cool colors (aqua, blue, green, navy, teal).

Face Presence: Empirical studies have shown how visual cues like human faces can deliver multi-faceted information and shape perceptions of objects in images (Joo et al., 2014; Joo et al., 2015; Peng, 2018). For example, Li and Xie (2020) find that pictures with human faces may induce engagement on Twitter. Kizilcec et al. (2014) find that videos including faces attract more visual attention and evoke positive affective responses. For each frame, we identify whether it contains any human *faces* or not.[16] We then compute the proportion of frames that contain human faces over all sampled frames as the rate of face presence of a video.

Video Length: We analyze one global feature of videos, the video length. The nature of Douyin as a short-video platform requires users to create content limited to a short time span (the limit for most users is 1 minute). Previous studies show that the duration of a video affects the engagement, with shorter videos being associated with higher engagement (Chen et al., 2021; Guo et al., 2014). Research from Zhu et al. (2020) finds that videos lasting less than 60 seconds are more successful on Douyin. Videos on Douyin are accompanied by duration meta-data, which we use to analyze video length.

Analysis of Non-Video Data

We make use of two non-video modes of data associated with Douyin trending videos in order to answer our substantive questions: account meta-data and textual topical descriptions of videos.

Account Meta-Data: For each video, we collect the name of the account creator, the verification status of the account, and detailed verification information about the account from Douyin. Douyin metadata tells us the account name of a video creator, e.g., Shanghai Fabu (上海发布), and in the verification information, Douyin also provides the name of the organization that is responsible for the account, e.g., Information Office of Shanghai Municipality (上海市政府新闻办官方抖音号). Using the account name and this verification information, four native Chinese speakers categorized all accounts into four types: 1) accounts affiliated with the CCP regime (regime-affiliated accounts), 2) accounts of ordinary users, 3) accounts of celebrities, and 4) accounts of non-official media. Regime-affiliated accounts are further broken down into: 1a) government/CCP accounts, 1b) mouthpiece newspaper accounts, 1c) non-mouthpiece newspaper accounts, and 1d) other official media accounts (see Table C1 in the Appendix for intercoder agreement).

Government/CCP accounts are those run directly by central or local government agencies or CCP organizations. This category includes government offices, government information/news center (政府新闻办公室， 新闻中心, or 融媒体中心), propaganda departments, Cyberspace Administration of China (CAC), and other government or CCP departments and entities. Mouthpiece newspapers, non-mouthpiece newspapers, and other official media accounts are those with an enterprise verification denoting them as state-controlled media organization—for example, *People's Daily*, Beijing TV, and Changcheng New Media Corporation (长城新媒体), a popular media outlet controlled by the Hebei provincial government). In China, broadcast media, including newspapers, radio, and television are for the most part owned and supervised by CCP committees; however, previous research has differentiated between newspapers that speak for the CCP—mouthpiece newspapers—and those that are more commercialized—non-mouthpiece newspapers (Qin et al., 2018; Stockmann, 2013; Zhao, 1998; Zhao & Guo, 2005). Following the definition used in Qin et al. (2018), we consider daily newspapers (日报) mouthpiece newspapers. The editorial policy of dailies is directly controlled by the corresponding Chinese propaganda department. Non-mouthpiece newspapers—for example, evening newspapers and metro newspapers, are those with fewer restrictions in their management. We define accounts from non-newspaper outlets, such as local TV stations, radio stations, and state-controlled media corporations, as other official media accounts.

Non-official media accounts are those verified as commercial media outlets, including 1) commercial newspaper and magazines (e.g. *Vogue*, *VistaStory*), 2) privately owned internet platforms (e.g. Youku, MGTV, Guancha.cn), 3) independent, self-media accounts or marketing accounts (e.g. Huihuo), and 4) accounts promoting one specific entertainment product like TV drama, variety shows, or sports shows (e.g. TV drama *Winter Begonia*, *NBA Headlines*).[17] The majority of accounts in this category are entertainment-based, though it does also include accounts such as *VistaStory*, which are not focused exclusively on entertainment. The term *celebrities* refers to accounts with individual verification as a celebrity (e.g. actors, singers, Douyin influencers) or an unverified user with more than 100,000 followers.[18] Ordinary users are accounts with no verification information and fewer than 100,000 followers.[19]

Textual Topic Descriptions: We use the topic associated with each video to place videos into two broad categories: propaganda and non-propaganda. Four native Chinese speakers categorized all topics, achieving high intercoder agreement. This coding scheme is driven by the fact that the common expectation is for authoritarian governments to disseminate propaganda. However, previous research suggests that to compete for attention on digital media, such governments may have to disseminate content unrelated to propaganda (Lu & Pan, 2020). We use an extremely broad definition of propaganda to provide the hardest possible test.

Traditionally, propaganda has been conceptualized as a form of persuasive communication where the sender (e.g., the government) aims to change the attitude, options, beliefs, or cognition of the receiver (Jowett & O'donnell, 2018). Others have adopted a narrower definition to argue that propaganda must entail deception through outright lies, omission, redirection, misrepresentation, or distortion (Bakir et al., 2019).

In this paper, we take the broadest definition and operationalize propaganda as any content pertaining to the Chinese regime and any content denigrating outgroups. Content pertaining to the Chinese regime includes any person, organization, idea, action, or outcome associated with the CCP and the Chinese government. This includes not only persuasive communication—any content that associates positive attributes such as competence or achievement with the CCP, CCP ideology, any level of government, as well as political leaders, officials, and party members—but also non-persuasive communication pertaining to the regime such as government announcements of programs, policies, procedures, and deadlines. Examples of propaganda topics include: "The Chinese Navy is the most trustworthy force for peace" (中国海军是最值得信赖的和平力量), which can be considered persuasive communication, and "Some roads in Beijing urban area will be temporarily controlled" (北京市区部分道路将临时管制), which is simply a government announcement that does not attribute positive a characteristic to any object. We also include in this operationalization of propaganda any content that denigrates—that is, associates with negative attributes—persons and organizations that the Chinese regime has designated as outgroups. This includes foreign countries, international organizations, and dissidents. For example, topics like "An old man in the U.S. was pushed to the ground by the police" (美国一老人被警察推倒在地), suggesting US police brutality, is considered propaganda because it falls into this category.

We disaggregate non-propaganda content into positive energy stories, human interest stories, breaking news, business news, and entertainment. Positive energy stories are stories about role models, societal and moral well-being, and happiness. "Positive energy" was initially a catchphrase circulated on the Chinese cyberspace to describe "any uplifting power and emotion, representing hope" (Yang & Tang, 2018). Starting from 2013, it has been commonly used by the state and CPC President Xi Jinping in his speeches to promote Chinese patriotism, moral behaviors, and support for the state ideology (Chen et al., 2020; Triggs, 2019).[20] On Douyin, different levels of governments have organized "positive energy" challenges to encourage users to create positive energy content.[21] For example, in our dataset, we captured positive-energy topics like "a delivery man bravely saved 6 people who fainted" (6人晕倒外卖小哥勇救人) and "I love everything about my mother" (我爱妈妈所有的样子). Although positive energy is clearly a key component of the information strategy of the Chinese regime, it does not fall into traditional conceptualizations of propaganda as its messages are unrelated to the regime.

Human interest stories are stories unrelated to the regime that are attention grabbing because of their novelty or because they elicit strong, positive and negative, emotions. Examples include "The biggest Super Moon of the year will appear tonight" (全年最大超级月亮今晚上演) and "14-year-old boy fell off the building after eating breakfast" (14岁男孩吃完早饭坠楼). Breaking news includes reports of domestic and international events that affect relatively large numbers of people, such as fires and accidents. For example, videos covered a forest fire in southwest China in March 2020 with topics like "Xichang mountain fire killed 19 local firefighters" (西昌山火致19名地方扑火人员牺牲). Business news reports on the operations and performance of firms—for example, "Tencent's market value surpasses Alibaba" (腾讯市值超越阿里巴巴) fit in this category. Entertainment news covers domestic and international celebrities, movies, television, sports, music, and fandom. For example, topics like "Jay Chou's new song Mojito" (周杰伦新歌Mojito) are coded as entertainment topics. We only place content in these five categories if they do not pertain to the Chinese regime in any way. If there was a story about a government official as a positive role model, that content would be classified as propaganda instead of positive energy. If there was a story about a movie about CCP history, that would be classified as propaganda instead of entertainment.

Because our data collection period coincided with the spread of Covid-19 in China, we also categorize whether a topic relates to Covid-19, and then remove such content from our main analysis. A large volume of videos from regime-affiliated accounts pertained to Covid-19 in March and early April (see Figure B1 in the Appendix). Because we want to capture the "normal" operations of regime-affiliated accounts, we exclude Covid-19 content. However, our results do not change when Covid-19 content is included (see Appendix Section B).

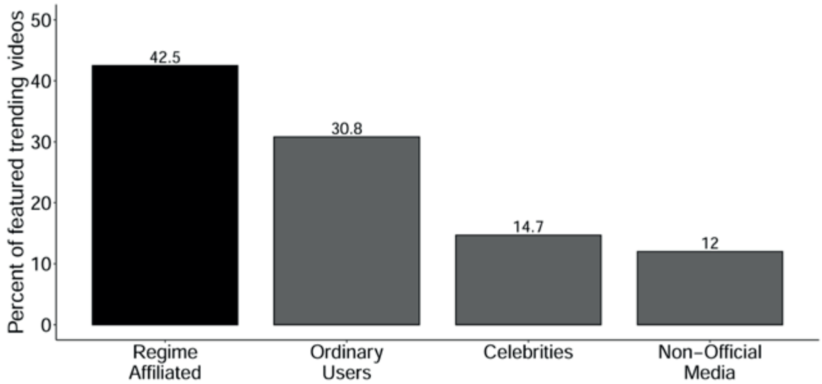

Figure 3 Share of Featured Trending Videos by Account Type

We use the text of topics to classify videos manually rather than classify videos themselves to increase efficacy and accuracy. To classify all videos in our dataset manually, coders would have had to watch the entire video before placing the video in a category, which would have taken approximately 300 hours. By contrast, classifying topics by text for all videos took research assistants around 120 hours total. An alternative to having human coders watch all the videos in our dataset would have been to manually annotate a small, random sample of videos and then classify all videos by using the human annotation to train a model such as a two-stream convolutional network (Simonyan & Zisserman, 2014), 3D convolutional network (Ji et al., 2012; Tran et al., 2015), or other state-of-the-art deep learning algorithm (Feichtenhofer et al., 2019). However, not only is computation costly but the Top-1 accuracy—how the highest probability answer returned from the model matches the ground-truth answer—for large datasets is less than 80% for most existing models (Xie et al., 2018). In other words, machine classification would have generated more errors than human coding of textual topics and likely required similar amounts of time for valiation.

Results

RQ1: Presence among trending videos: We find that videos featured in the Trending page on Douyin are dominated by content from regime-affiliated accounts. As Figure 3 shows, 42.5% of trending videos come from regime-affiliated outlets. Of the remaining trending videos during our time period,

Table 1 Regime-Affiliated Accounts

Type	No. of Accts.	Trending Video / Acct.
Government / CCP	665	2.8
Mouthpiece Newspaper	160	8.9
Non-Mouthpiece Newspaper	216	12.1
Other official media	1,005	8.4

30.8% come from ordinary users, 14.7% from celebrities, and 12.0% from non-state media accounts.

Featured trending videos come from a large number of regime-affiliated accounts, especially government/CCP accounts and other official media accounts (Table 1). In the study period, trending videos came from 665 government/CCP accounts, 160 mouthpiece newspapers, 216 non-mouthpiece newspapers, and 1,005 accounts of other types of official media. Table 1 shows that, on average, government/CCP accounts had an average of 2.8 videos featured on the Trending page, mouthpiece newspapers 8.9, non-mouthpiece newspapers 12.1, and other official media 8.4.

RQ2: Video features: Videos made by regime-affiliated accounts exhibit particularly high values of several features shown in previous research to attract attention and exhibit levels of these features that are more similar to celebrity and ordinary user content than content from non-official media accounts.

Figure 4 shows estimates of frame-level features with bootstrapped 95% confidence intervals. Regime-affiliated accounts create videos with higher levels of brightness than any other type of account. Regime-affiliated content has high entropy, close to that of celebrity accounts. Videos from regime-affiliated accounts are also very short in duration, similar to the length of videos from ordinary users. Figure 4 shows that videos of regime-affiliated accounts use warm color dominance to a degree similar to that of celebrity videos. The areas where videos of regime-affiliated accounts deviate from findings of research on attention is cold color dominance and face presence. Videos from regime-affiliated accounts exhibit lower face presence than videos from all other types of accounts, though is closest to the average face presence of videos from ordinary users. Videos from regime-affiliated accounts also exhibit higher levels of cold color dominance than ordinary users and celebrity content, similar to levels seen in content from non-official media.

RQ3: Content: The analysis of the textual description of trending videos shows that propaganda content does not represent the majority of content from regime-affiliated accounts. The left panel of Figure 5 shows that among

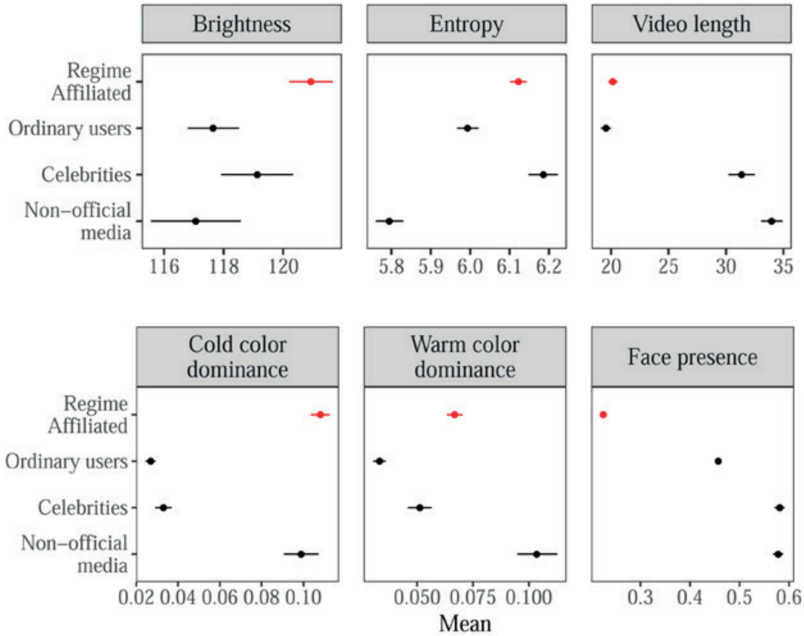

Figure 4 Frame-level Features

the featured trending videos from all regime-affiliated accounts, slightly over a third (33.7%) is propaganda. Among the remaining content, 8.1% is positive energy, 30.0% is human interest, 14.0% is breaking news, 3.8% is business news, and 10.4% is entertainment. Compared to ordinary users, non-official media, and celebrities, regime-affiliated accounts contain much less entertainment content, but regime-affiliated accounts contain human interest stories at levels similar to ordinary users and celebrities.

When we examine the topic of featured trending videos by different types of regime-affiliated accounts, the right panel of Figure 5 shows that while government/CCP accounts and mouthpiece newspapers contain relatively more propaganda than non-mouthpiece newspapers and other official media—45.3% and 41.2% vs. 29.5% and 31.2%, respectively—the relative share of propaganda content always remains less than half of the content disseminated by regime-affiliated accounts.

A natural question that arises from these results is whether the pattern we observe—of a relatively low share of propaganda content—is because Douyin users are drawn to non-propaganda videos. The concern is that Douyin users prefer non-propaganda, so propaganda videos appear less frequently in the trending videos. If Douyin trending videos only featured topics and discussions

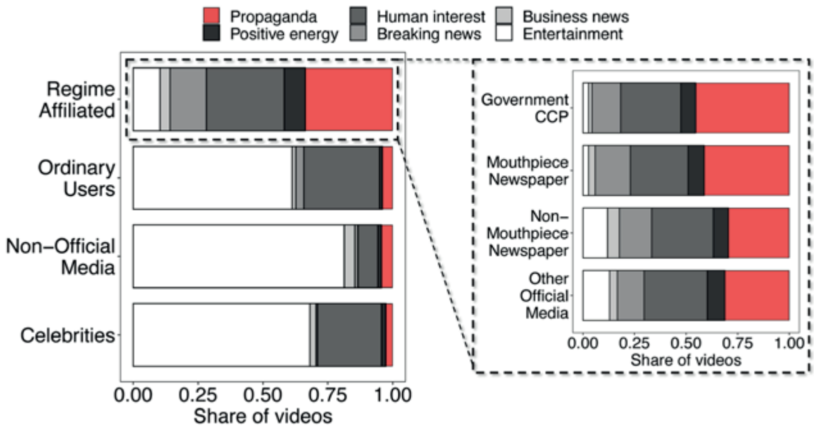

Figure 5 Content of Videos

that are organically viral—i.e., content attracting outsized attention on social media before being featured in the trending topics—then the patterns we observe would reflect user interest rather than government intention. However, we know that Douyin trending topics are not all organically viral because Douyin may include certain topics in the trending videos topic list at the behest of the regime or in collaboration with the regime. Thus, if we can focus solely on videos related to topics mandated by the Chinese regime for inclusion in trending topics, then the content would reflect the intentions of the regime rather than the interests of the audience. We do this by analyzing the content of featured trending videos that are only discussed by regime-affiliated accounts. Among trending topics in our study time period, a sizable portion of topics are only discussed by regime-affiliated accounts while another substantial proportion are never discussed by regime-affiliated accounts, resulting in a bimodal distribution (see Appendix Figure D1).

When we analyze the content of trending videos falling into topics that are only discussed by regime-affiliated accounts, we find that while the share of propaganda content increases, propaganda remains less than half of all content produced by these accounts (see Figure 6). Propaganda content represent 49.8% of all content in trending videos for government/CCP accounts, 45.5% for mouthpiece newspapers, 36.8% for non-mouthpiece newspapers, and 37.4% for other official media. Among videos in topics only discussed by regime-affiliated accounts, human interest stories represent a sizeable share of videos (28.6% for government/CCP, 29.0% for mouthpiece newspapers, 33.2% for non-mouthpiece newspapers, and 33.9% for other official media).

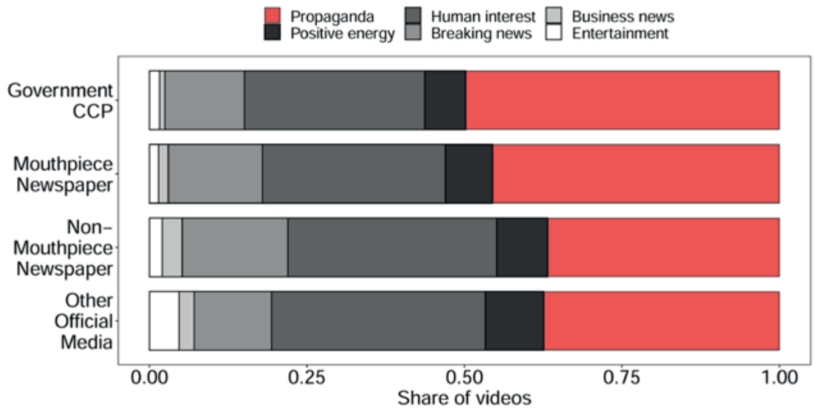

Figure 6 Content of Videos in Topics Only Discussed by Regime-Affiliated Accounts

Examples of propaganda content among featured trending video topics that are only discussed by regime-affiliated accounts include: "Gong Zheng appointed acting mayor of Shanghai" (龚正任上海市代市长) and "The Ministry of Foreign Affairs declares China has no interest in interfering in the U.S. election" (外交部称中国没兴趣干预美国大选).

Examples of human interest content among featured trending video topics that are only discussed by regime-affiliated accounts include: "Electric car helmets increase in price" (电动车头盔涨价) and "Man who saved a child's life was injured and admitted into the ICU" (男子救轻生女被砸受伤进ICU).

Conclusion

This paper shows the pervasive presence of the Chinese regime in Douyin trending videos. There are more videos from regime-affiliated accounts in the study period than from celebrities, ordinary users, or non-official media. Videos produced by regime-affiliated accounts exhibit visual features known to capture attention, such as brightness, entropy, warm color dominance, and short duration. Furthermore, the relative levels of these attention-maximizing features found in videos from regime-affiliated accounts are more similar to those found in videos from celebrities and ordinary users than non-official media. Propaganda consistently represents less than half of the content disseminated by regime-affiliated accounts, including government/CCP accounts, mouthpiece newspapers, non-mouthpiece newspapers, and other official media.

These findings contribute to our understanding of the information strategies of authoritarian regimes in the era of digital media. Autocrats must compete for audience attention, and this paper shows that they do so by producing digital content with visual features that maximize attention and by producing content other than propaganda. This deviates from prior research on information control in authoritarian regimes, which tends to assume that autocrats, including the Chinese regime, would censor objectionable content (using strategies ranging from censorship to repression) and produce propaganda (Huang, 2015; King et al., 2013; Pan & Siegel, 2020). What are the implications of regime-affiliated channels producing content that falls outside of most definitions of propaganda? Producing non-propaganda content may very well serve the goals of autocrats. If non-propaganda is appealing to audiences, it can increase the size of the audience for propaganda and redirect public attention away from topics the regime does not want discussed—a more general phenomenon that has also been called distraction (King et al., 2017) and information channeling (Earl et al., 2021). However, the production of non-propaganda content—in our case, content that is completely unrelated to the regime—may also dilute the authority and voice of regime-affiliated outlets. It may be more difficult for an outlet to be authoritative if a sizable share of their content does not appear serious. Users may also gloss over propaganda content in favor of non-propaganda content.

Beyond these substantive implications, this paper contributes to the emerging literature on visual analysis in communication research by demonstrating a multi-modal approach to analyzing video as data. This multi-modal approach allows us to increase accuracy and efficiency—e.g., using text rather than video for classification—while allowing us to use frame-level and global video features to answer substantive questions of interest.

Acknowledgements

Our thanks to Xin Jin, Yitian Liu, Yuyang Shi, and Yiwei Wang for research assistance; Ben Lee, Haoran Li, Yansong Tang, and participants of the Chinese Political Communication Forum and SICSS-Beijing for many helpful comments and suggestions; and to the National Science Foundation (Grant #1831481) and the Stanford Graduate Fellowship for research support.

Appendix A Douyin platform

Figure A1 Homepage Comparison Across Platforms

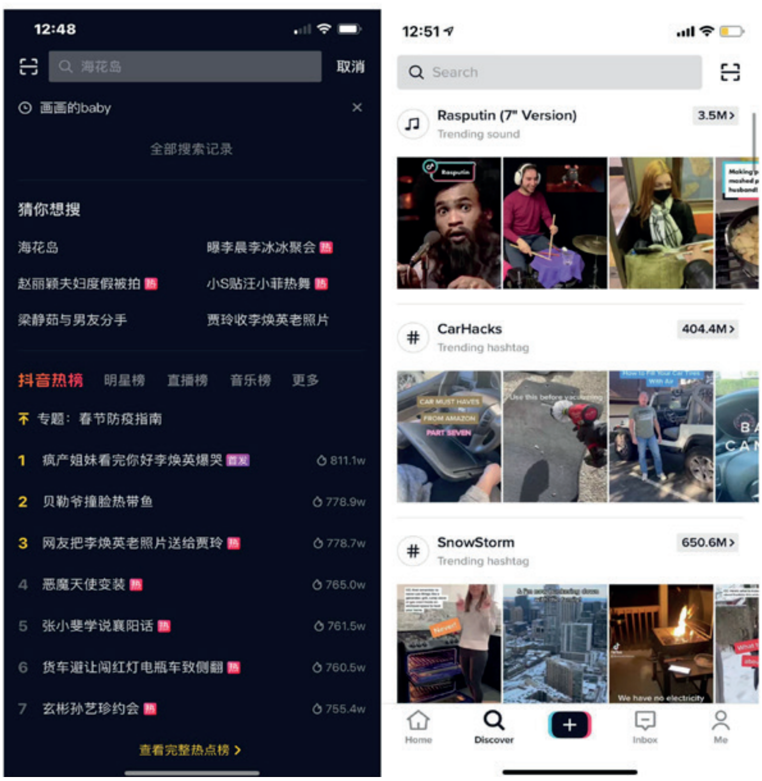

Figure A2 Discover Page of Douyin (left) and TikTok (right)

Appendix B Content related to Covid-19

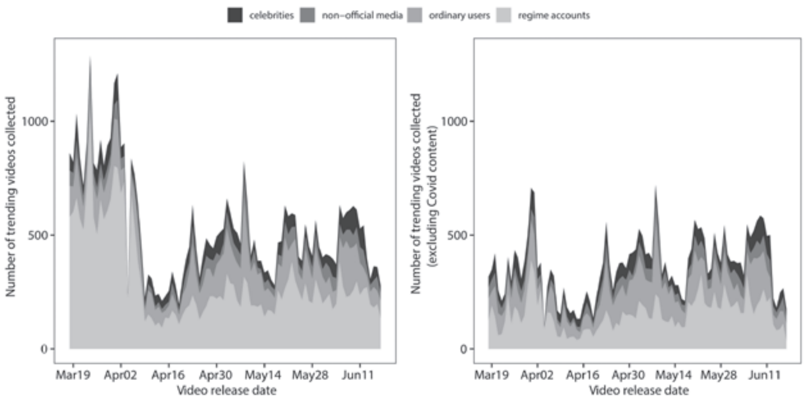

Figure B1 Content related to Covid-19 represents a large share of content from regime-affiliated accounts in March and April

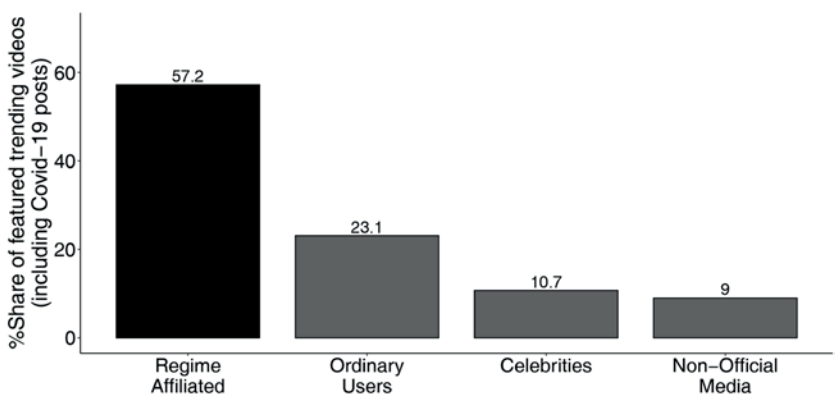

Figure B2 Share of Featured Trending Videos by Account Type (including Covid-19 content)

Table B1 Regime-Affiliated Accounts (including Covid-19 content)

Account type	Number of Accounts (Excluding Covid-19 content)	Number of Accounts (Including Covid-19 content)
Government / CCP	665	845
Mouthpiece Newspaper	160	190
Non-Mouthpiece Newspaper	216	242
Other official media	1,005	1,201

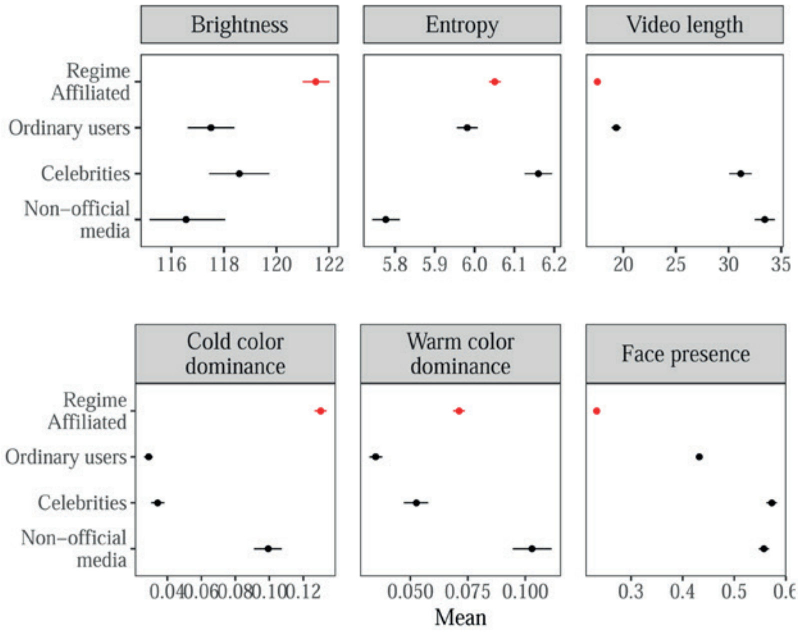

Figure B3 Frame-level Features (including Covid-19 content)

Appendix C Intercoder Agreement

Table C1 Intercoder Agreement for Account Categorization

Variable	Number of Accounts	Percent agreement	Kripp's alpha
Type of account	200	97.5%	0.93
Regime-affiliated account	200	98.5%	0.96

Appendix D Regime-Affiliated Content by Trending Topic

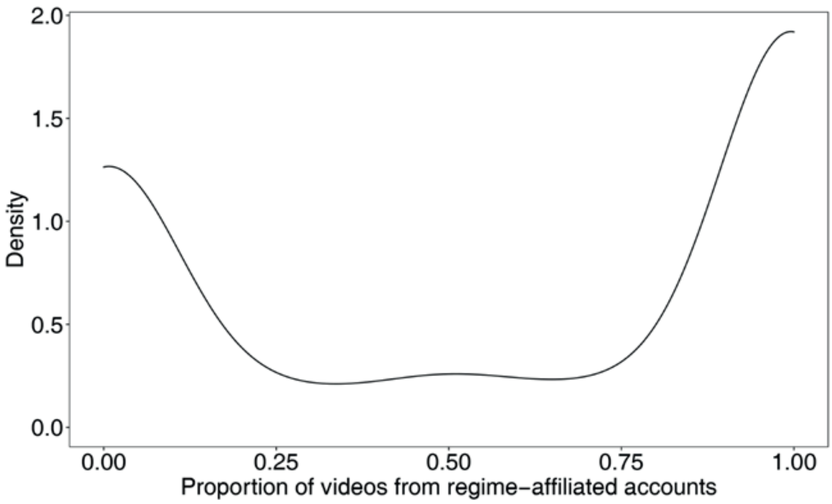

Figure D1 Density plot of proportion of videos from regime-affiliated accounts within each trending topic

Notes

1. See http://www.xinhuanet.com/tech/2021-01/05/c_1126948875.htm.
2. This contrasts with the quantitative analysis of textual data, which has grown rapidly over the past decade due to the availability of large-scale texts from social media, computational advances in automated text analysis methods, and established norms for integrating text into social science research (Blei, 2012; Grimmer & Stewart, 2013; Roberts et al., 2014).
3. For examples, see Edwardson et al. (1981), Graber (1990), Kizilcec et al. (2014), McLeod (1995), Newhagen and Reeves (1992), Pan et al. (2021), and Reese (1984).
4. See https://bit.ly/3bqnadN.
5. There is debate over whether and how content on TikTok is influenced by Chinese government directives, see https://wapo.st/3pMwBZU and https://bit.ly/3khOcbk. There is no debate that Douyin content is influenced by the Chinese government.
6. See http://media.people.com.cn/n1/2018/0904/c40699-30271586.html.
7. See http://www.199it.com/archives/828860.html.
8. We also collected videos in real time. However, the number of videos associated with each topic was extremely dynamic in real time, and some topics contained many videos while others contained very few. In December 2020, we retrospectively collected videos associated with each topic and found a much more stable number of videos in each topic. We did not observe sub-

stantive differences in features and topics when comparing real-time and retrospective videos and use retrospectively collected videos in our analysis because of their greater consistency.

9. Accounts can be verified as celebrity accounts or enterprise accounts.
10. See http://www.cac.gov.cn/2019-12/03/c_1576907933632994.htm
11. See https://www.hubpd.com/c/2019-07-04/856182.shtml.
12. We do this using the *VideoCapture* function of the OpenCV library (Bradski, 2000).
13. Loading frames by the OpenCV returns every frame in Blue, Green, Red (BGR) colorspace. We then convert each frame from BGR color space to CIELAB colorspace to extract the L channel (perceptual luminance) value (Baldevbhai & Anand, 2012).
14. For each sampled frame, we use the OpenCV function COLOR_BGR2GRAY to convert the frame from BGR colorspace to grayscale.
15. We use CSS2.1 to delineate base color, see https://www.w3.org/TR/CSS21/syndata.html#color-units. Dominant RGB vectors that could not be matched to any of the 17 base colors directly are matched to the closest base color by Euclidean distance based on the Webcolors library (see https://pypi.org/project/webcolors/).
16. To do this we apply an open-source Python library face_recognition (Geitgey, 2020). Face_recognition is built using the pretrained deep learning framework for face recognition from the C++ toolkit *dlib* (King, 2017), which reaches high benchmark performance.
17. Scholars often refer to evening and metro newspapers as "commercialized" newspapers. In this paper, because media accounts on Douyin include a greater variety of outlets—newspapers, TV, radio, magazines, etc.—we reserve the term "commercialized" to refer to media accounts not directly associated with the Party, and use the term "Non-mouthpiece Newspaper" to denote evening and metro papers.
18. The 100,000 threshold is commonly used by video-based social media companies to differentiate influencers and normal users when evaluating their platform, see http://www.199it.com/archives/958540.html.
19. We identified a small proportion of accounts (5.3%) belonging to non-government enterprises, public institutions, retail outlets, and other entities that we could not classify into these five categories. We excluded these accounts and their videos from our analysis.
20. See http://world.people.com.cn/n/2013/0116/c244926-20225251.html.
21. See https://zhuanlan.zhihu.com/p/38066659.

References

Bakhshi, S., Shamma, D. A., Kennedy, L., & Gilbert, E. (2015). Why we filter our photos and how it impacts engagement. *ICWSM*, 12–21.

Bakir, V., Herring, E., Miller, D., & Robinson, P. (2019). Organized persuasive communication: A new conceptual framework for research on public relations, propaganda and promotional culture. *Critical Sociology*, 45 (3), 311–328.

Baldevbhai, P. J., & Anand, R. (2012). Color image segmentation for medical images using l* a* b* color space. *IOSR Journal of Electronics and Communication Engineering*, 1 (2), 24–45.

Barry, A. M. (1997). *Visual intelligence: Perception, image, and manipulation in visual communication*. SUNY Press.

Bjola, C. (2017). Propaganda in the digital age. *Global Affairs*, 3 (3).

Blei, D. M. (2012). Probabilistic topic models. *Communications of the ACM*, 55 (4), 77–84.

Boussalis, C., & Coan, T. G. (2020). Facing the electorate: Computational approaches to the study of nonverbal communication and voter impression formation. *Political Communication*, 1–23.

Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*.

Brinberg, M., Ram, N., Yang, X., Cho, M.-J., Sundar, S. S., Robinson, T. N., & Reeves, B. (2021). The idiosyncrasies of everyday digital lives: Using the human screenome project to study user behavior on smartphones. *Computers in Human Behavior*, 114, 106570.

Camgöz, N., Yener, C., & Güvenç, D. (2004). Effects of hue, saturation, and brightness: Part 2: Attention. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group* (*Great Britain*)*, Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur*, 29 (1), 20–28.

Chen, Q., Min, C., Zhang, W., Ma, X., & Evans, R. (2021). Factors driving citizen engagement with government TikTok accounts during the Covid-19 pandemic: Model development and analysis. *Journal of Medical Internet Research,* 23 (2), e21463.

Chen, X., Kaye, D., & Zeng, J. (2020). # Positiveenergy douyin: Constructing "playful patriotism" in a chinese short-video application. *Chinese Journal of Communication*, 1–21.

Chen, Y., Wang, S., Zhang, W., & Huang, Q. (2018). Less is more: Picking informative frames for video captioning. *Proceedings of the European Conference on Computer Vision* (*ECCV*), 358–373.

CNNIC. (2020). The 46th China statistical report on internet development.

CNNIC. (2021). The 47th China statistical report on internet development.

Confessore, N., & Wakabayashi, D. (2017). How Russia harvested American rage to reshape US politics. *New York Times*.

Dahmen, N. S. (2012). Photographic framing in the stem cell debate: Integrating eye-tracking data for a new dimension of media effects research. *American Behavioral Scientist,* 56 (2), 189–203.

Earl, J., Maher, T., & Pan, J. (2022). The digital repression of social movements, protest, and activism: A synthetic review. *Science Advances. Forthcoming*.

Edwardson, M., Grooms, D., & Proudlove, S. (1981). Television news information gain from interesting video vs. talking heads. *Journal of Broadcasting & Electronic Media*, 25 (1), 15–24.

Egusa, H. (1982). Effect of brightness on perceived distance as a figure—ground phenomenon. *Perception*, 11 (6), 671–676.

Elliot, A. J., & Maier, M. A. (2014). Color psychology: Effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology*, 65, 95–120.

Feichtenhofer, C., Fan, H., Malik, J., & He, K. (2019). Slowfast networks for video recognition. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 6202–6211.

Geitgey, A. (2020). Ageitgey/face_recognition: The world's simplest facial recognition api for python and the command line. https://github.com/ageitgey/face%5C_recognition

Goel, S., Anderson, A., Hofman, J., & Watts, D. J. (2016). The structural virality of online diffusion. *Management Science*, 62 (1), 180–196.

Goethe, J. W. v., & Eastlake, C. L. (2006). *Theory of colours*. Dover.

Graber, D. A. (1990). Seeing is remembering: How visuals contribute to learning from television news. *Journal of Communication*, 40 (3), 134–155.

Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21 (3), 267–297.

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of mooc videos. *Proceedings of the First ACM Conference on Learning@ Scale Conference*, 41–50.

Hameleers, M., Powell, T. E., Van Der Meer, T. G., & Bos, L. (2020). A picture paints a thousand lies? The effects and mechanisms of multimodal disinformation and rebuttals disseminated via social media. *Political Communication*, 37 (2), 281–301.

Hoogeveen, M. (1997). Toward a theory of the effectiveness of multimedia systems. *International Journal of Human-Computer Interaction*, 9 (2), 151–168.

Huang, H. (2015). Propaganda as signaling. *Comparative Politics*, 47 (4), 419–444.

Iyer, A., & Oldmeadow, J. (2006). Picture this: Emotional and political responses to photographs of the Kenneth Bigley kidnapping. *European Journal of Social Psychology*, 36 (5), 635–647.

Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3D convolutional neural networks for human action recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35 (1), 221–231.

Joo, J., Li, W., Steen, F. F., & Zhu, S.-C. (2014). Visual persuasion: Inferring communicative intents of images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 216–223.

Joo, J., Steen, F. F., & Zhu, S.-C. (2015). Automated facial trait judgment and election outcome prediction: Social dimensions of face. *Proceedings of the IEEE International Conference on Computer Vision*, 3712–3720.

Joo, J., & Steinert-Threlkeld, Z. C. (2018). Image as data: Automated visual content analysis for political science. *arXiv preprint arXiv:1810.01544*.

Jowett, G. S., & O'donnell, V. (2018). *Propaganda & persuasion*. Sage Publications.

Keller, F. B., Schoch, D., Stier, S., & Yang, J. (2020). Political astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication*, 37 (2), 256–280.

King, D. (2017). Dlib c++ library: High quality face recognition with deep metric learning. *Recognition Processing Time* (*one server vs. three servers*).

King, G., Pan, J., & Roberts, M. E. (2013). How censorship in China allows government criticism but silences collective expression. *American Political Science Review*, 107 (2), 326–343.

King, G., Pan, J., & Roberts, M. E. (2017). How the Chinese government fabricates social media posts for strategic distraction, not engaged argument. *American Political Science Review*, 111 (3), 484–501.

Kizilcec, R. F., Papadopoulos, K., & Sritanyaratana, L. (2014). Showing face in video instruction: Effects on information retention, visual attention, and affect. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2095–2102.

Li, Y., & Xie, Y. (2020). Is a picture worth a thousand words? An empirical study of image content and social media engagement. *Journal of Marketing Research*, 57 (1), 1–19.

Lu, Y., & Pan, J. (2020). Capturing clicks: How the Chinese government uses clickbait to compete for visibility. *Political Communication*, 1–32. https://doi.org/10.1080/10584609.2020.1765914

Marmolin, H. (1992). Multimedia from the perspectives of psychology. *Multimedia*, 39–52.

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. New York: Data & Society Research Institute.

McLeod, D. M. (1995). Communicating deviance: The effects of television news coverage of social protest. *Journal of Broadcasting & Electronic Media*, 39 (1), 4–19.

Mebane Jr., W. R., Wu, P., Woods, L., Klaver, J., Pineda, A., & Miller, B. (2018). Observing election incidents in the United States via Twitter: Does who observes matter? *Annual Meeting of the Midwest Political Science Association, Chicago*.

Messaris, P., & Abraham, L. (2001). The role of images in framing news stories. *Framing public life: Perspectives on media and our understanding of the social world*, 215–226. Routledge: Londong.

Newhagen, J. E., & Reeves, B. (1992). The evening's bad news: Effects of compelling negative television news images on memory. *Journal of Communication*, 42 (2), 25–41.

Nyhuis, D., Ringwald, T., Rittmann, O., Gschwend, T., & Stiefelhagen, R. (2021). Automated video analysis for social science research. *Handbook of Computational Social Science, Volume 2: Data Science, Statistical Modelling, and Machine Learning Methods*. Routledge: London.

Pan, J., Shao, Z., & Xu, Y. (2021). How government-controlled media shifts policy attitudes through framing. *Political Science Research and Methods*.

Pan, J., & Siegel, A. (2020). How Saudi crackdowns fail to silence online dissent. *American Political Science Review*, 114 (1), 109–125.

Pancer, E., & Poole, M. (2016). The popularity and virality of political social media: Hashtags, mentions, and links predict likes and retweets of 2016 US presidential nominees' tweets. *Social Influence*, 11 (4), 259–270.

Peng, Y. (2018). Same candidates, different faces: Uncovering media bias in visual portrayals of presidential candidates with computer vision. *Journal of Communication*, 68 (5), 920–941.

Peng, Y., & Jemmott, J., III. (2018). Feast for the eyes: Effects of food perceptions and computer vision features on food photo popularity. *International Journal of Communication*, 12, 313–336.

Pieters, R., & Wedel, M. (2004). Attention capture and transfer in advertising: Brand, pictorial, and text-size effects. *Journal of Marketing*, 68 (2), 36–50.

Powell, T. E., Boomgaarden, H. G., De Swert, K., & de Vreese, C. H. (2015). A clearer picture: The contribution of visuals and text to framing effects. *Journal of Communication*, 65 (6), 997–1017.

Powell, T. E., Boomgaarden, H. G., De Swert, K., & de Vreese, C. H. (2019). Framing fast and slow: A dual processing account of multimodal framing effects. *Media Psychology*, 22 (4), 572–600.

Qian, J., Zhang, K., Wang, K., Li, J., & Lei, Q. (2018). Saturation and brightness modulate the effect of depth on visual working memory. *Journal of Vision*, 18 (9), 16.

Qin, B., Strömberg, D., & Wu, Y. (2018). Media bias in China. *American Economic Review*, 108 (9), 244224–76.

Reese, S. D. (1984). Visual-verbal redundancy effects on television news learning. *Journal of Broadcasting & Electronic Media*, 28 (1), 79–87.

Reeves, B., Ram, N., Robinson, T. N., Cummings, J. J., Giles, C. L., Pan, J., Chiatti, A., Cho, M., Roehrick, K., Yang, X., et al. (2019). Screenomics: A framework to capture and analyze personal life experiences and the ways that technology shapes them. *Human–Computer Interaction*, 1–52.

Roberts, M. E. (2018). *Censored: Distraction and diversion inside China's great firewall*. Princeton University Press.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., Albertson, B., & Rand, D. G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58 (4), 1064–1082.

Rosenholtz, R., Li, Y., Mansfield, J., & Jin, Z. (2005). Feature congestion: A measure of display clutter. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 761–770.

Sanovich, S. (2017). Computational propaganda in Russia: The origins of digital misinformation. *Computational Propaganda Research Project. Working paper*, (3).

Sethna, J. et al. (2006). *Statistical mechanics: Entropy, order parameters, and complexity* (Vol. 14). Oxford University Press.

Simonyan, K., & Zisserman, A. (2014). Two-stream convolutional networks for action recognition in videos. *Advances in neural information processing systems*, 568–576.

Sobolev, A., Chen, M. K., Joo, J., Steinert-Threlkeld, Z. C., et al. (2020). News and geolocated social media accurately measure protest size variation. *American Political Science Review*, 114 (4), 1343–1351.

Steinert-Threlkeld, Z. (2019). The future of event data is images. *Sociological Methodology*, 49 (1), 68–75.

Stenberg, G. (2006). Conceptual and perceptual factors in the picture superiority effect. *European Journal of Cognitive Psychology*, 18 (6), 813–847.

Stockmann, D. (2013). *Media commercialization and authoritarian rule in China*. Cambridge University Press.

Sundar, S. S. (2000). Multimedia effects on processing and perception of online news: A study of picture, audio, and video downloads. *Journalism & Mass Communication Quarterly*, 77 (3), 480–499.

Sundar, S. S. (2008). *The main model: A heuristic approach to understanding technology effects on credibility*. MacArthur Foundation Digital Media; Learning Initiative.

Sundar, S. S., & Limperos, A. M. (2013). Uses and grats 2.0: New gratifications for new media. *Journal of Broadcasting & Electronic Media*, 57 (4), 504–525.

Tran, D., Bourdev, L., Fergus, R., Torresani, L., & Paluri, M. (2015). Learning spatiotemporal features with 3D convolutional networks. *Proceedings of the IEEE International Conference on Computer Vision*, 4489–4497.

Triggs, F. (2019). The ideological function of "positive energy" discourse: A people's daily analysis. *British Journal of Chinese Studies*, 9 (2), 83–112.

Van De Weijer, J., Schmid, C., & Verbeek, J. (2007). Learning color names from real-world images. *2007 IEEE Conference on Computer Vision and Pattern Recognition*, 1–8.

Williams, N. W., Casas, A., & Wilkerson, J. D. (2020). *Images as data for social science research: An introduction to convolutional neural nets for image classification*. Cambridge University Press.

Woolley, S. C., & Howard, P. N. (2017). Computational propaganda worldwide. *Working Paper*, (11. Oxford, UK), Projecton Computational Propaganda.

Xi, N., Ma, D., Liou, M., Steinert-Threlkeld, Z. C., Anastasopoulos, J., & Joo, J. (2020). Understanding the political ideology of legislators from social media images. *Proceedings of the International AAAI Conference on Web and Social Media*, 14, 726–737.

Xie, S., Sun, C., Huang, J., Tu, Z., & Murphy, K. (2018). Rethinking spatiotemporal feature learning: Speed-accuracy trade-offs in video classification. *Proceedings of the European Conference on Computer Vision* (*ECCV*), 305–321.

Yang, P., & Tang, L. (2018). "Positive energy": Hegemonic intervention and online media discourse in China's Xi Jinping era. *China: An International Journal*, 16 (1), 1–22.

Yang, T., & Peng, Y. (2020). The importance of trending topics in the gatekeeping of social media news engagement: A natural experiment on Weibo. *Communication Research*, 0093650220933729.

Yang, X., Ram, N., Robinson, T., & Reeves, B. (2019). Using screenshots to predict task switching on smartphones. *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–6.

Zhang, H., & Pan, J. (2019). Casm: A deep-learning approach for identifying collective action events with text and image data from social media. *Sociological Methodology*, 49 (1), 1–57.

Zhao, Y. (1998). *Media, market, and democracy in China: Between the party line and the bottom line* (Vol. 139). University of Illinois Press.

Zhao, Y., & Guo, Z. (2005). Television in China: History, political economy, and ideology. *A Companion to Television*. Wiley: New York:, 521–539.

Zheng, X. S., Chakraborty, I., Lin, J. J.-W., & Rauschenberger, R. (2009). Correlating low-level image statistics with users-rapid aesthetic and affective judgments of web pages. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1–10.

Zhu, C., Xu, X., Zhang, W., Chen, J., & Evans, R. (2020). How health communication via Tik tok makes a difference: A content analysis of Tik tok accounts run by Chinese provincial health committees. *International Journal of Environmental Research and Public Health*, 17 (1), 192.

Zillmann, D., Knobloch, S., & Yu, H.-s. (2001). Effects of photographs on the selective reading of news reports. *Media Psychology*, 3 (4), 301–324.

About the author

Jennifer Pan (http://jenpan.com/) is an Associate Professor of Communication at Stanford University. Email: jp1@stanford.edu. ORCID: https://orcid.org/0000-0002-4818-0122.

Yingdan Lu (https://yingdanlu.com/) is a PhD Candidate at Department of Communication, Stanford University. Email:yingdan@stanford.edu. ORCID: https://orcid.org/0000-0002-9955-6070.