Understanding Voter Fraud Misinformation Videos during the 2024 Taiwan Election on YouTube

YANHENG LI, Renmin University of China, China
DA WANG*, Renmin University of China, China
YUPING WANG*, Renmin University of China, China

Although voter fraud misinformation in the U.S. has been extensively studied, fewer research studies have investigated voter fraud misinformation in Taiwan's elections. In this study, we present a mixed-method analysis of video-based voter fraud misinformation during the 2024 Taiwan election on YouTube. We develop a computational pipeline to identify video accounts with different political leanings that may produce election-related misinformation, collect the videos uploaded by these accounts, and automatically determine which videos are related to ballots with the assistance of Large Language Models. Subsequently, we manually identify voter fraud discussion videos using a codebook we developed. Following this, we conduct a comprehensive analysis of the identified videos. We find that video accounts aligned with the Democratic Progressive Party (DPP) produce the highest number of videos discussing voter fraud misinformation. Additionally, videos discussing such misinformation tend to receive more comments but fewer likes compared to videos without this content. We also observe that videos associated with the DPP are quite distinct from those linked to the Kuomintang (KMT) or Taiwan People's Party (TPP), with unique characteristics that may be further revealed through their audio and video features. Finally, we conduct case studies to examine different patterns in videos either supporting or refuting voter fraud misinformation. Among accounts promoting voter fraud claims, traditional media outlets often include misinformation in their news programs, implying tacit endorsement, while grassroots media may present suspicious vote-counting scenes as evidence to spread misinformation. In contrast, traditional and grassroots media accounts associated with the DPP tend to refute misinformation through news programs or influencer commentary. Our work sheds light on the discourse surrounding voter fraud in Taiwan and offers valuable insights into strategies for mitigating the spread of voter fraud misinformation videos globally.

CCS Concepts: • **Security and privacy → Social aspects of security and privacy**; • **Human-centered computing → Empirical studies in HCI**.

Additional Key Words and Phrases: Misinformation; Election; Social Media; YouTube; Multimodal; Videos

ACM Reference Format:
Yanheng Li, Da Wang, and Yuping Wang. 2025. Understanding Voter Fraud Misinformation Videos during the 2024 Taiwan Election on YouTube. *Proc. ACM Hum.-Comput. Interact.* 9, 7, Article CSCW516 (November 2025), 40 pages. https://doi.org/10.1145/3757697

*Corresponding Authors and both corresponding authors contributed equally to this research.

Authors' Contact Information: Yanheng Li, yanheng@ruc.edu.cn, Renmin University of China, Beijing, China; Da Wang, dawang@ruc.edu.cn, Renmin University of China, Beijing, China; Yuping Wang, yupingw@ruc.edu.cn, Renmin University of China, Beijing, China.

ACM 2573-0142/2025/11-ARTCSCW516
https://doi.org/10.1145/3757697

1 Introduction

Election integrity is a cornerstone of democracy, as it upholds public confidence in electoral outcomes and ensures the legitimacy of elected officials [50]. Despite efforts by election administrations worldwide, allegations of voting manipulation still appear occasionally after various elections [20, 26, 39, 51, 56, 67].

In the Taiwan's 2024 election, held on January 13, voters selected the next leader of Taiwan. Three candidate pairs ran: Lai Ching-te and Hsiao Bi-khim (Democratic Progressive Party, DPP), Hou Yu-ih and Jaw Shaw-kong (Kuomintang, KMT), and Ko Wen-je and Wu Hsin-ying (Taiwan People's Party, TPP). The final vote shares were 40.1% for the DPP, 33.5% for the KMT, and 26.5% for the TPP.[1] This competitive election triggered widespread discussion, including numerous allegations of voter fraud and concerns about the fairness of the voting process [24]. For example, they doubt the fairness of Taiwan's unique voting system, which relies on paper ballots and sets the stage for voter fraud, differing from many other regions globally.[2]

Moreover, the impact of social media, particularly YouTube, is significant, as it serves as an essential medium for both political messaging and public discourse on the election process and voter fraud controversies [24, 45]. In Taiwan, YouTube is frequently used by politicians and influencers to communicate directly with voters, aiming to affect voter perceptions and even election outcomes [45]. Past research highlights this trend, indicating that politicians often appear alongside popular influencers to amplify their reach and appeal to diverse voter groups [23]. This trend also mirrors the media landscape in other parts of the world, such as the United States, where election fraud controversies fueled media attention and public debate in the 2020 presidential election [58, 67]. Similarly, the Taiwan's 2024 election became contentious for the voter fraud scandal, when some social media influencers (e.g., "Alisasa", one of the most prominent online supporter of Ko Wen-je and the Taiwan People's Party) accused authorities of voter manipulation shortly after the election, despite being quickly fact-checked and debunked as misinformation by the Taiwan FactCheck Center (TFC) [24].

Unlike platforms like TikTok, which predominantly target younger audiences [42] , YouTube in Taiwan attracts a diverse demographic, including older adults and younger viewers alike [23]. Taiwan's YouTube channels span from traditional mainstream media outlets to grassroots channels. Traditional media accounts are affiliated with large, established media corporations and often have access to traditional communication channels, such as TV and radio. In contrast, grassroots channels typically consist of smaller, independent creators with no connections to major media companies, relying solely on digital platforms to publish their content. However, it remains unclear how these distinct media types address sensitive issues, such as voter fraud.

In this paper, we investigate how different media outlets on YouTube respond to allegations of voter fraud in the Taiwan's 2024 election. Specifically, we compiled a dataset[3] of 5,641 videos from three main parties and examined certain videos that are more inclined to discuss voter fraud. We then conduct research focusing on these three research questions:

RQ1: What is the distribution of voter-fraud videos among different parties?

RQ2: Do videos including voter fraud discussions have higher engagement than those without voter fraud discussions?

RQ3: What is the difference among voter-fraud videos from *(a) different parties* and *(b) different media types*, considering their audiovisual features?

[1] https://www.bloomberg.com/graphics/2024-taiwan-election/

[2] https://tfc-taiwan.org.tw/fact-check-reports/migration-10153/

[3] The code and data are available at: https://github.com/RUC-AI-for-Journalism-and-Communication/Voter-fraud-in-2024-Taiwan-Election

This research offers insights into how both traditional and grassroots media outlets frame and discuss issues related to election integrity. Notably, the Taiwan FactCheck Center (TFC) officially conducted extensive fact-checking, refuting voter fraud allegations as misinformation, underscoring the significance of our study in understanding and addressing political misinformation.

Ethics. In this paper, we only analyze publicly available YouTube data, and we do not investigate any personally identifiable data or have any interactions with users who generated these data. Therefore, our work has an exemption from the Institutional Review Board (IRB) at Renmin University of China.

2 Related Work

2.1 Taiwan Politics & Atypical Politician

Taiwan shows polarized politics. Traditionally, Taiwan mainly had two parties "Kuomintang" (KMT) and "Democratic Progressive Party" (DPP), and they are polarized in policies ranging from cross-strait relations to LGBT rights [15]. However, a third party is becoming a pivotal political force in Taiwan politics recently, which is called "Taiwan People's Party" (TPP). As such, some recent investigations dug into the three-party politics in Taiwan. Chang and Fang [7] found that the posts relevant to the KMT and DPP candidates are associated with more engagements on topics such as geopolitical issues, while posts related to the third-party leader associated with more engagements on domestic policy topics by analyzing nearly 1 million posts on Facebook and Instagram. Building on this line of inquiry, Ho [22] analyzed the role of social movements in shaping party agendas during the 2024 Taiwan election. Ho emphasized how social issues re-emerged in campaign discourse and influenced the policies of the KMT, DPP, and TPP. His study also highlighted how movement-based parties struggled to maintain electoral relevance, while the TPP strategically absorbed elements of social activism to consolidate its political position.

Particularly, the TPP candidate Ko Wen-je, who was a famous physician in Taiwan previously, came as an atypical Taiwanese politician [78]. In political studies, an atypical politician refers to a political figure who deviates from unconventional political trajectories and exhibits distinctive characteristics in their background, influence, and decision-making style. This concept can be understood through the following four dimensions:

Background and Political Trajectory. An atypical politician often enters politics from a non-traditional career path rather than through conventional political or bureaucratic routes. Examples include Donald Trump, who built his career as a businessman before becoming the U.S. president [61]; Ronald Reagan, a former Hollywood actor who later became president [62]; and Volodymyr Zelensky, a comedian and television personality who transitioned into the presidency of Ukraine [30].

Personal Charisma and Online Influence. Compared to traditional politicians, atypical politicians often possess a strong sense of personal charisma, which plays a crucial role in their political appeal. In the digital age, they tend to leverage social media and online platforms to cultivate a fan-based following. For instance, while both DPP and KMT used social media mainly to mobilize their supporters to vote [40, 71, 72], Ko Wen-je strived to become an online star in politics through his social media presence and rhetoric, which gained widespread support from his fans [77].

Decision-Making Style and Societal Impact. Atypical politicians often disrupt established political norms and decision-making processes, reflecting broader societal shifts. For example, Trump's presidency marked a potential turning point for the post-Cold War liberal international order led by the U.S., while Reagan's policies ushered in a new era of neoliberalism [59].

Election Campaign Strategies and Public Mobilization. Their campaigns often involve unconventional strategies and heightened emotional engagement from supporters. A notable case is the aftermath of the 2020 U.S. presidential election, when Trump's loss triggered the U.S. Capitol riot in 2021, illustrating the intensity of political mobilization surrounding atypical politicians [27].

Although existing studies have begun to examine the rise of atypical political figures and their use of social media, scant attention has been paid to the controversies surrounding such figures during the vote-counting phase. However, such controversies have emerged as a salient global phenomenon in recent years, with notable occurrences in the United States. Our study addresses this gap by specifically focusing on this underexplored aspect of atypical political actors.

Methodologically, prior research on atypical political figures has predominantly relied on qualitative approaches and descriptive data. In contrast, our study leverages the integration of large-scale data and large language models to construct a multimodal dataset. This approach not only offers new empirical insights into the dynamics of atypical political figures, but also provides a novel methodological pathway for future investigations into the structural and communicative mechanisms behind their rise.

2.2 Misinformation

Misinformation is often defined as a piece of information that is not accurate or is misleading [76], which does not need to be created or spread with malicious intent (i.e., *disinformation*) [75]. Given that some of the voter fraud spreaders may be genuinely skeptical towards election integrity, we focus on the concept *misinformation* in this paper.

According to Shu et al. [64], misinformation research has two important research directions: *detection* and *characterization*. In the direction of detecting misinformation, numerous papers have proposed ways to identify misinformation by using data mining [64]. Although data mining proves to be efficient and fast, an important drawback is that the results often fluctuate across different datasets [5]. In addition, it lacks interpretability, which is required in trustworthy artificial intelligence [3, 32]. Novel methods including automatic fact-checking approaches have been proposed to detect misinformation with more interpretability [19], especially by using Large Language Models [10], which requires further refinement.

Another research direction is to characterize propagation of misinformation on social media [81]. Numerous studies have investigated misinformation, including text misinformation and image-based misinformation. This type of research can be conducted from several perspectives, including quantitative and qualitative [69]. Some recent investigations include the U.S. 2020 elections. Moore et al. [48] examined exposure to untrustworthy websites during the 2020 U.S. election. Starbird et al. [67] discussed three cases of misleading claims of voter fraud during the 2020 U.S. election to discover the nature of participatory misinformation. The CSCW community has also emphasized working around misinformation. Some recent work includes COVID-19 misinformation [1, 73].

Similar to misinformation research elsewhere, researchers also focus on false narratives around the Taiwan elections. Lin [39] conducted a misinformation analysis of the 2018 Taiwan local elections, highlighting the interplay between various social media platforms and emphasizing the significance of fact-checking in combating misinformation. Hung et al. [24] highlighted the threat of AI-generated disinformation on the integrity of the Taiwan election. Our work is distinct from Hung et al. [24] in that, although the authors in [24] acknowledged the existence of voter fraud misinformation videos, they did not perform a deeper empirical analysis to assess their influence on social media or their impact on perceptions of election integrity. In contrast, we compile a dataset of videos discussing voter fraud misinformation from both traditional and grassroots media outlets on YouTube. Based on this dataset, we analyze user engagement metrics and examine the tactics used to present these videos through a combination of quantitative audiovisual analysis and qualitative case studies.

2.3 Election Integrity

Election integrity has been investigated from two aspects. One is to audit the voting infrastructures, appearing in security communities. For instance, Kohno et al. [34] identified several significant security vulnerabilities, which include allowing voters to cast multiple ballots using simple tricks of a widely used electronic voting system "AccuVote-TS" by analyzing its source code. Amid a voter fraud accusation against the voting results of the 2020 U.S. presidential election in Antrim County, Michigan, Halderman [20] conducted an independent investigation into the whole voting process by analyzing election system data and proved the election integrity in Antrim County. Debant and Hirschi [12] pinpointed vulnerabilities in the electronic French legislative election voting system by applying reverse engineering and proposed relevant methods to fix these issues.

Another research direction is to investigate how narratives around voter fraud spread on social media. By delving into three case studies, scholars demonstrated how election disinformation is amplified and disseminated on X, formerly known as Twitter [58, 67]. Based on millions of tweets and thousands of Facebook posts, Benkler et al. [4] found that mail-in ballot disinformation campaigns on Twitter and Facebook were driven by elites and manipulated by mass media. Jakesch et al. [26] discovered that the Bharatiya Janata Party (BJP) ran a series of Twitter trending hashtag manipulation campaigns in the 2019 Indian general election by forming a network of WhatsApp groups, in which the campaign organizers coordinated the BJP supporters to post tweets with designated hashtags intensely to promote the BJP politicians. As argued by the authors, such operations are in the intersection of manmade popularity and coordination to express support, whose legitimacy needs to be further reviewed [26].

Our work continues the discussion surrounding voter fraud in the 2020 U.S. election [58, 67] and may also have implications for elections worldwide in the future.

2.4 Multimodal Approach

Traditionally, researchers focus mainly on analyzing textual content on social media. Recently, with the rise of online platforms for sharing images and videos [52], investigating multimodal content has become increasingly crucial. One line of research applies perceptual hashing to group visually similar images into clusters, which are then annotated by scholars to facilitate the interpretation and understanding of their meaning. In this research direction, Zannettou et al. [80] find that hateful memes are popular on fringe Web communities (e.g., /pol/, Gab, and The_Donald subreddit) and Wang et al. [74] show that contradictory to previous findings in [70], tweets with images containing misinformation messages may involve higher engagement compared with tweets with images containing non-misinformation messages.

By similar approaches, Resende et al. [60] and Garimella and Eckles [17] respectively demonstrate how images with misinformation were widely disseminated in public Brazilian and Indian WhatsApp groups during elections. Lee et al. [38] used techniques introduced by Poco and Heer [57], demonstrating how COVID-19 skeptics analyze visualized COVID-19 data to promote false COVID-19 claims.

Another research direction is to use neural networks to analyze images and single frames of videos. In particular, Peng [54] found that images with subtle differences adopted by news outlets with partisan political leanings show imperceptible political bias. A complete tutorial can be found in [29]

3 Dataset & Methods

For this research, we have designed a computational pipeline for data collection and multimodal analysis, which is shown in Figure 1. We will first describe how we collect data, then we will show how to annotate the collected videos and how to extract audiovisual features from videos, respectively.

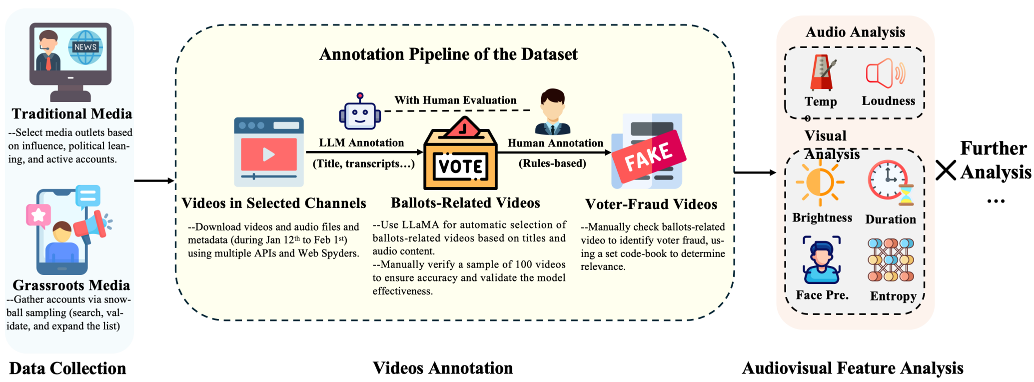

Fig. 1. Overview of Computational Analysis Pipeline

3.1 Data Collection

In this subsection, we will first explain how we determine the time frame. Next, we will illustrate the procedures to select media outlets in our research. Finally, we will describe the implementation of our data collection.

Determining the time frame. We decide to collect data within the time range from January 12th to February 1st, 2024. The selection of this justified period is based on the following considerations.

For the starting point, we verify it by cross-referencing a previous study on voter fraud misinformation in the 2024 Taiwan election [24], and the date that the election was held.

(1) The Taiwan 2024 election was held on January 13th, 2024, and the DPP candidate Lai Ching-te won the election.
(2) In the evening of January 13th, a popular YouTuber “Alisasa,” who was also a supporter of TPP,[4] posted several videos on Instagram with suspicious scenes that some poll volunteers were cheating when they were counting votes for the 2024 Taiwan election. Based on these videos, Alisasa claimed there was “voter fraud” in the election. Although soon debunked by the TFC, these posts led the voter fraud controversy to be widely discussed [24], sparking this controversy. Therefore, we mark January 13th as the beginning of this controversy.
(3) With a one-day buffer before January 13th, our selected time frame begins on January 12th.

For the ending point, we justify our selection on the date of the last relevant TFC fact-checking report and supplemental manual verification. Additional details regarding how the ending point is determined are provided in Appendix A.1.

(1) The ending point is supported by Taiwan’s media regulatory framework. Under the authority of the "National Communications Commission" (NCC), the dissemination of misinformation is restricted once identified by TFC. Since the final fact-checking report was released on the morning of January 31st, it is unlikely that significant voter fraud discourse occurred afterward.

[4]https://www.storm.mg/lifestyle/4986095

(2) This is further corroborated by our manual check, in which the authors confirmed that few videos discussing voter fraud were published in the following three days by any of the accounts in our dataset.
(3) With a one-day buffer after January 31st, our selected time frame concludes on February 1st.

Channels Selection. To better understand public opinions towards voter fraud issues, we divide YouTube accounts into traditional media, which are channels from TVs, radios, and newspapers, and grassroots media, which are created spontaneously.

For traditional media. To comprehensively obtain a list of traditional media outlets, we check the list of TV channels and radios from the "National Communication Committee" (NCC) of Taiwan[5] and the list of newspapers from the "Taiwan Periodical Literature."[6] Combining these two lists above gives us a complete list of Taiwan's officially recognized traditional news media.

Our goal is to select news outlets that have influence over the whole of Taiwan, and in addition to that, some of the media outlets are excluded for the following reasons.

(1) The media outlet does not have an active YouTube account, i.e., the account was terminated before July 2024, either due to voluntary deactivation or enforced removal.
(2) Though some media outlets are popular around Taiwan, either they focus on specific topics, e.g., entertainment and sports, instead of politics, or they target specific populations, e.g., people who speak Hakka Chinese.

At the end of this process, we obtain a list of 16 media groups and each group contains several separate YouTube accounts. For each group, we also identify its political leaning (Please see Appendix A.2.1 for details in determining the political leaning for each channel). This leads to four DPP media outlets, six KMT media outlets, and six media outlets that do not have any clear political leanings, which are denoted as "Other" in the list. We further confirm that all the YouTube accounts associated with these media outlets are included, and we eventually identify seventeen accounts for DPP, sixteen accounts for KMT, and thirteen accounts for "Other," respectively. These selected accounts mainly focus on topics that are highly related to politics.

Please note that since the Taiwan People's Party (TPP) was established long after the other two parties, it does not own any news outlets by itself, which was acknowledged by the then TPP leader Ko Wen-je [33]. Instead, they operate a large number of grassroots media accounts.[7]

For grassroots media. To the best of our knowledge, there is currently no comprehensive source listing YouTube grassroots accounts related to Taiwan politics. To compile such a list, we construct a relatively thorough set of grassroots accounts for the DPP, KMT, and TPP using a snowball sampling approach, described below.

(1) We begin by identifying grassroots media accounts affiliated with the DPP, which is generally considered to have the most developed media presence among all parties in Taiwan. Firstly, we use the Google search engine to collect as many relevant online articles and news reports as possible from credible sources, with each article listing various grassroots user accounts associated with the DPP.
(2) Notably, Taiwan's political parties and influencers maintain social media accounts across several popular platforms, including YouTube, Facebook, and X. In addition, some YouTube accounts may no longer be active, as the source articles cover a broad time period. Therefore, we have manually verified each account mentioned in the articles to determine the following two questions:
Q1: Does it maintain an active YouTube presence? That is, whether the account remained

[5] https://www.ncc.gov.tw/english/index.aspx
[6] https://tpl.ncl.edu.tw/NclService/
[7] For more information, please see https://www.tpp.org.tw/newsdetail/3138

active or was terminated before July 2024.

Q2: Was it active during the 2024 Taiwan election period? That is, whether the account published any videos between January 12th and February 1st, 2024.

If both criteria are met, the account is added to our grassroots list. Ultimately, we identify 10 qualified grassroots YouTube accounts affiliated with the DPP. When combined with 17 traditional media accounts, this brings the DPP's total to 27 media accounts. For a balanced comparison in subsequent analysis, we set 27 as the target number of total media accounts for both the KMT and TPP.

(3) We repeat the same process to sample accounts for the KMT and TPP. As a result, we obtain 11 grassroots accounts for the KMT and 27 for the TPP. In this way, all three parties reach a numerical balance in terms of selected media accounts.

Statistics of collected channels are shown in Table 1, and we also provide additional details and an link containing the complete list of media account names in Appendix A.2

Implementation of data collection. We then use the YouTube API[8] to retrieve the list of video IDs and metadata for all videos uploaded during this period. By combining a Python Web Spider package [63] with APIs,[9] we download the highest-quality video and audio files for each video. The data collection occurred during the last week of August 2024. In total, we obtained 5,641 videos. The overall number of videos is summarized in Table 2.

Table 1. Number of media accounts for different parties.

	KMT	DPP	TPP	Other
Traditional	16	17	0	13
Grassroots	11	10	27	0
Overall	27	27	27	13

Table 2. Number of videos across different parties and channel types.

	KMT	DPP	TPP	Other
Traditional	1,886	1,898	0	873
Grassroots	269	117	598	0
Overall	2,155	2,015	598	873

3.2 Videos Annotation

After obtaining the 5,641 videos, our goal is to identify those discussing voter fraud. This process is carried out in two main steps: 1) We use a Large Language Model (LLM) to identify all videos related to ballots. 2) We then manually inspect these ballot-related videos and label them as either voter-fraud-relevant or not. The detailed procedure is described below.

Phase I: Labeling ballot-related videos Labeling all 5,641 videos manually would be a daunting task. Therefore, we leverage an LLM to assist in identifying ballot-related videos. We develop an automated classifier using the open-source LLM, LLaMA 3.1 to extract all ballot-related videos. Rather than directly identifying voter fraud content, we first target ballot-related content, as videos discussing voter fraud are a subset of these, and LLMs more reliably identify this broader category.

The classifier is designed through extensive prompt engineering and human validation to ensure its effectiveness. It performs multimodal analysis by integrating both textual and audio information and operates in a two-step mechanism: *1) Title-Based Classification*: Each video is first evaluated based on its title. If the title alone is sufficient to determine the video as ballot-related, it is flagged. *2) Audio-Based Classification*: If the title is inconclusive, we convert the video's audio to text using

[8]Here, we use YouTube API V3 to collect data (https://developers.google.com/youtube/v3).

[9]Here, we use a combination of YouTube API V3 and Google Custom Search API (https://developers.google.com/custom-search/v1).

the VOSK speech recognition model [49]. The transcribed text is then analyzed by the LLM to determine whether it pertains to ballot-related content.

For the prompt setting, we begin with the following prompt template (translated from traditional Chinese):

> "Now, assume you are a news analyst, and please determine whether the following content is relevant to casting votes, voting scrutiny, and voter fraud. If so, please answer 'yes' and 'no' otherwise. \n The content is shown as: \n {text_content}"

However, we find that this prompt often leads the LLM to avoid giving definitive answers, possibly due to alignment-related hesitations. To address refusal issues, e.g., "Sorry, I can't answer this question...", we revise the prompt as follows:

> "Please determine whether the following content is relevant to casting votes, voting scrutiny, and voter fraud based on its words and content. If so, you must answer 'yes' and 'no' otherwise. \n Please answer as precisely as possible (especially when your answer is yes), and give your reasons briefly. The content is shown as: \n {text_content}"

To rigorously evaluate the validity of the LLM-based annotation process, we first construct a human-annotated golden dataset. Two authors independently annotate a randomly sampled set of 100 videos (i.e., each sampled video received two human annotations) according to the video titles and transcripts that aligned precisely with LLM labeling. The codebook of ballot-relevance annotation for human annotators is included in Appendix A.3.3. We then compute the inter-annotator agreement between the two human coders, which demonstrates high consistency. The inter-coder Fleiss's Kappa is 0.976, indicating substantial agreement [36]. Following this, the authors engage in a discussion to reconcile any discrepancies and reach consensus, resulting in a finalized golden dataset for validation.

We then apply the prompt to the same 100 videos using the LLM and compared its outputs to the golden dataset. We compute standard classification metrics to assess the LLM's labeling performance, including Accuracy (0.910), Macro-F1 (0.899), Precision (0.884), and Recall (0.926). The results validate the performance of LLM in annotating videos [37], supporting its use for large-scale annotation.

Based on this validation, we proceed to apply the LLM to the full video dataset, resulting in the identification of 2,771 ballot-related videos. Additional implementation details for this initial phase of annotation are provided in Appendix A.3.1.

Phase II: Labeling voter-fraud videos In the next step, the three authors manually review all videos labeled as relevant to ballots. A video is labeled as **voter-fraud** on the basis of whether it *contains the voter-fraud narrative* (i.e., the claim that ballots were stuffed, stolen, or otherwise manipulated). Based on preliminary observations of collected videos, the authors develop a codebook with the following rules to determine whether a video is relevant to voter fraud (the detailed codebook is provided in Appendix A.3.3):

(1) If any on-screen text in the video is related to voter fraud, the video is labeled as voter-fraud relevant.
(2) If any spoken audio in the video is related to voter fraud, the video is labeled as voter-fraud relevant.
(3) If one or more individuals in the video are shown engaging in voter fraud or are highly suspicious of doing so, the video is labeled as voter-fraud relevant.
(4) If the video mentions or implies that an entity (a person or organization) is attempting to commit voter fraud, it is labeled as voter-fraud relevant.
(5) If the video includes only viewer comments discussing voter fraud, without any connection to the video content itself, it is NOT considered a voter-fraud video.

To validate the annotation consistency, we randomly selected 100 videos from the 2,771 ballot-related videos. All three authors independently annotated these videos based on the codebook (i.e., each video received three annotations). We then used Fleiss's Kappa to assess inter-coder reliability, and the resulting average score was 0.923, indicating a high level of agreement among the three annotators.

Following this validation, the remaining 2,771 videos are divided among the authors, with each annotating approximately 923 videos independently. Since Rule (4) in the codebook may be more subjective or context-dependent, any video potentially falling under this rule is discussed collectively among all authors to reach a consensus. Through this process, we identify a total of 170 videos as related to voter fraud. More implementation details and annotated examples of phase II are supplemented in Appendix A.3.2

To ensure a comprehensive analysis, we also compile a list of all accounts that published these voter-fraud videos. For a balanced comparison in answering our research questions, we further include 3,667 non-voter-fraud videos published by the same accounts within the same data collection period. These videos form the final dataset for subsequent analysis.

3.3 Automated Video Features Extraction

In the next step, we need to analyze the videos. We employ a systematic approach to multimodal analysis, leveraging both audio and video processing techniques to extract key features from content data files within a specified directory for each video. The analysis framework includes individual media analysis workflows, feature extraction, and data aggregation, as detailed below [9, 44].

Video Analysis To achieve computational efficiency, the analysis samples video frames at specified intervals (i.e., every 30 frames) to capture representative data. This sampling approach reduces resource demands while preserving key visual patterns across the video. The video analysis workflow calculates multiple visual and structural characteristics of the video content using the OpenCV library in Python [11]. This multi-step workflow enables a general analysis of the video's visual properties, including brightness, entropy, face presence, and duration.

(1) *Brightness*: Using the LAB color space [2], brightness is calculated by averaging the intensity values within the Luminance (L) channel across sampled frames. This approach isolates brightness from color. Values scale between 0-255, with higher ones typically indicating well-lit scenes, while lower ones can suggest darker or even without light.
(2) *Entropy*: To represent the complexity and randomness within the video frames, video entropy is calculated using Shannon entropy [31]. The Shannon entropy for a sampled video frame is defined as follows:

$$H(X) = -\sum_{x \in X} P(x) \log_2 P(x), \tag{1}$$

where: $H(X)$ is the Shannon entropy of the pixel intensity distribution X in the sampled frame, $P(x)$ is the probability of occurrence of pixel intensity x, $\log_2$ represents the logarithm to the base 2. For implementation, we first convert pixels to grayscale to reduce dimensionality and decrease computation consumption, under which circumstance X takes values from 0 to 255 (the intensity range in an 8-bit image). Applying Equation 1, the output entropy scaling between 0 and 8 quantifies the uncertainty associated with the pixel intensities in the sampled frame. A higher entropy indicates more complexity and more various details, while a lower entropy suggests a more uniform or predictable content in selected frames (e.g., a frame with all black pixels would be entropy scored at zero). For the entropy of a video, we take the average of entropy scores across all sampled frames.

(3) *Face Presence*: The presence of human faces is quantified by calculating the ratio of frames containing faces to total frames, derived from the face_recognition library in Python [18]. Using the pre-trained face recognition model, the ratio scaling 0-1 provides a measure of the appearance of human subjects within the video content.
(4) *Duration*: Video duration is determined by calculating the ratio of frame count to frame rate, providing an accurate length measurement in seconds.

Audio Analysis To maintain compatibility across analysis tools, audio files are standardized to '.wav' format through an automatic conversion process. The analysis workflow centers on extracting rhythmic and amplitude-based features, specifically tempo and loudness:

(1) *Loudness (RMSE)*: Loudness is calculated based on Root Mean Square Energy (RMSE), providing an objective measure of audio intensity. The loudness calculation begins with loading the audio data, which is represented as a digital time series of sound pressure levels. Using the librosa library, RMSE is computed directly from this time-series data. It provides a measure of the energy contained in the signal, calculated as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}, \tag{2}$$

where x_i represents the amplitude of each sample, and N is the number of samples in the audio signal. By squaring each amplitude value before averaging, RMSE captures both the peaks and valleys of the waveform, resulting in a comprehensive measure of signal strength. Values are subsequently normalized to a 0–1 scale, which closer to 1 indicate higher loudness levels, while nearer to 0 suggests lower intensity or quieter segments.
(2) *Tempo (BPM)*: A measure of beats per minute (BPM), is calculated using the Librosa library in Python [47]. This process involves computing the onset envelope, which represents rhythmic events in the audio, and applying beat tracking to quantify the tempo.

We merge the results of audio and video analysis into a unified dataset with the metadata derived for each video, especially for engagement metrics (numbers of likes, views, and comments), enabling comprehensive data storage and subsequent analytical applications.

4 Results

4.1 RQ1: What is the distribution of voter-fraud videos among different parties?

For RQ1, we first present the temporal trend of the number of total included videos and voter-raud videos posted on YouTube throughout our data collection period in Figure 2, and then illustrate how the number of posted voter-fraud videos from different parties varies in Figure 3.

Figure 2 shows that voter-fraud videos account for a small proportion of all videos throughout the entire period, with a notable increase from January 13th to January 19th, approximately a week after the opening day of the election ballots. Voter-fraud videos almost disappear after January 30th, which aligns with the defined endpoint of our data collection period.

As shown in the left panel of Figure 3, we observe that the parties "DPP," "TPP," "KMT," and "Other" have posted 88, 44, 24, and 14 voter-fraud videos, respectively. By the right panel, Figure 3 also reveals the temporal trends of camp-specific voter-fraud videos. The "DPP" shows a significant spike in voter-fraud videos from January 13th to 18th, 2024, closely associated with the opening date of the election poll on January 13th, which likely triggered increased attention and engagement. During this period, the number of videos peaks above 15 per day on January 18th. By contrast, despite showing a general and slight increase from January 13th to 15th, "other" exhibits more modest occurrences of videos on the topic.

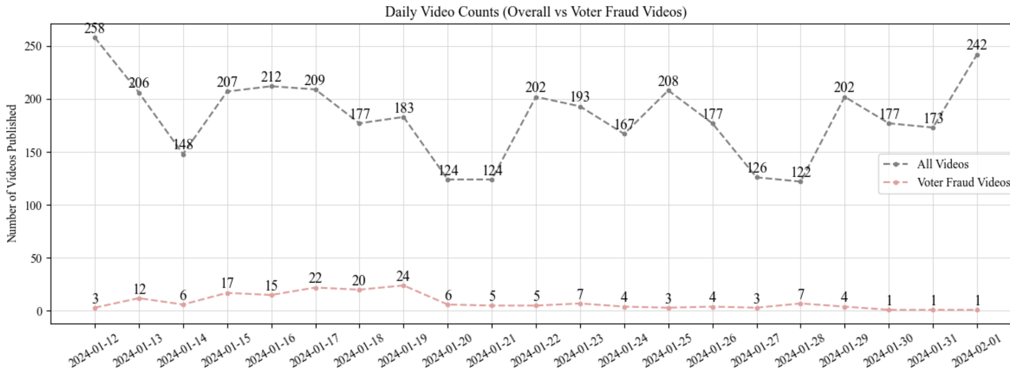

Fig. 2. Temporal Trend of the Number of Total Videos Included and the Number of Voter-Fraud Videos

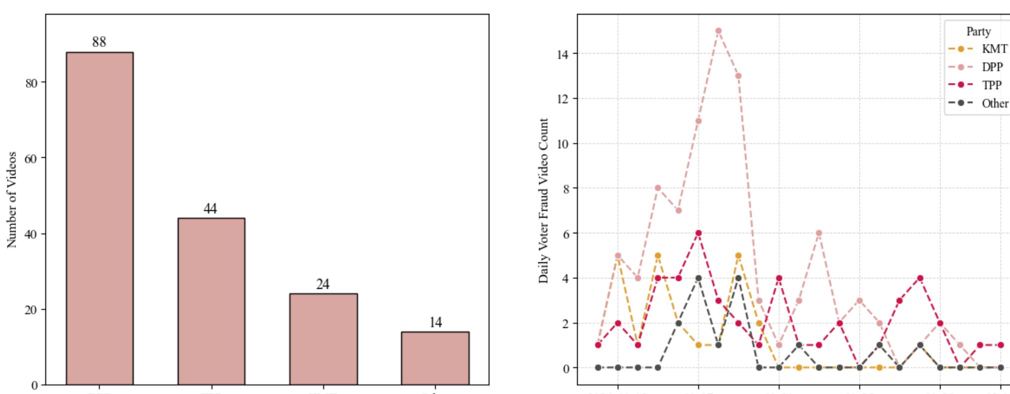

Fig. 3. Total and Daily Counts of Voter-Fraud Videos across Different Parties

Generally, voter-fraud videos consistently represent a small fraction of daily videos, with most days showing fewer than twenty voter-fraud videos compared to a much higher overall video count. This disparity underscores the narrow focus of voter-fraud content, suggesting limited general interest or a concentrated discussion primarily within certain political circles, notably the "DPP" and "TPP" parties. These patterns suggest that voter fraud concerns are not a predominant theme across overall video content, instead peaking in relevance around specific dates likely tied to particular election events.

Takeaways of RQ1 From RQ1, we find that "DPP" has the highest number of voter-fraud videos, followed by "TPP," "KMT," and "Other," respectively. The "DPP" camp posted the most voter fraud videos on January 18th.

4.2 RQ2: Do videos including voter fraud discussions have higher engagement than those without voter fraud discussions?

In RQ2, we mainly focus on three engagement metrics: "views," "likes," and "comments." We calculate and present the raw statistics of all three engagement metrics, comparing two groups: (1) "voter-fraud" videos—videos that discuss topics related to voter fraud during the 2024 Taiwan election; and (2) "non-voter-fraud" videos—videos that do not discuss voter fraud but were published by the same accounts during the same data collection period, serving as a local baseline to control for account-specific feature preferences.

With respect to these three metrics, we present the *empirical cumulative distribution functions (ECDFs) for both "voter-fraud" and "non-voter-fraud" videos in Figure 4, and the two-sample Kolmogorov-Smirnov (KS) test results in Table 3* for engagement variables [41, 70]. An explanation of the two-sample Kolmogorov-Smirnov (KS) test can be found in Appendix A.4.

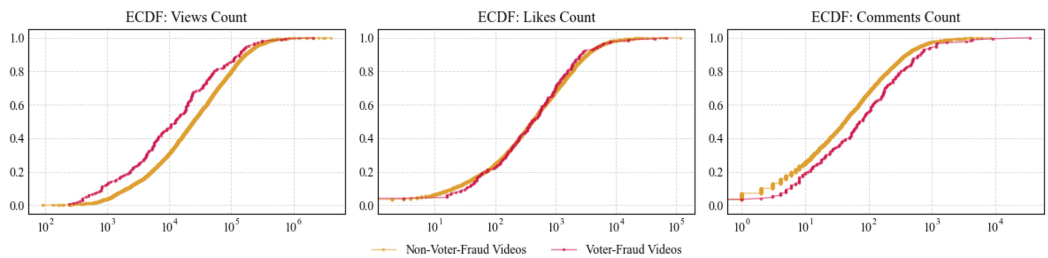

Fig. 4. Comparison of Engagement Variables for Voter-Fraud and Non-Voter-Fraud Videos: ECDFs for Views Count, Likes Count, and Comments Count with log transformation

Table 3. KS Test Results Comparing Engagement Variables of Voter-Fraud and Non-Voter-Fraud Videos

	Views Count	**Likes Count**	**Comments Count**
KS Statistic	0.176	0.043	0.142
P-Value	<0.001	0.905	0.003

Note: The cell is colored according to the video type color (as shown in Figure 4) corresponding to the type with the higher median value for that engagement attribute. The level of transparency is set according to the value of the KS Statistic.

From the results, we observe the following:

(1) **Views Count** presents an interesting divergence. Compared to non-voter-fraud videos (median = 24,332), voter-fraud videos show lower variability in views (median = 12,558.50), which may reflect limited public focus on specific political topics, with some videos receiving significant attention while others remain relatively unnoticed. The KS test result (statistic = 0.176, p-value < 0.001) suggests a statistically significant difference, indicating that voter-fraud videos receive a lower level of engagement in terms of "views count" compared to non-voter-fraud videos.

(2) **Likes Count** shows no significant differences in distribution between voter-fraud (median = 413) and non-voter-fraud videos (median = 402) with the result of the KS test (statistic = 0.043, p-value = 0.905).

(3) **Comments Count** demonstrates a skewed distribution, with voter-fraud videos (median = 76) showing a broader range compared to non-voter-fraud videos (median = 42). The higher number of comments on the former may indicate that these videos generate more discussion, potentially due to controversial or attention-grabbing narratives. The KS test (statistic = 0.142, p-value = 0.003) shows a statistically significant difference between the groups, suggesting that voter-fraud videos receive a higher level of engagement in terms of "comments" compared to non-voter-fraud videos.

Takeaways of RQ2 From RQ2, we find that the "views count" for voter-fraud videos is lower than that of non-voter-fraud videos, while the "comments count" for voter-fraud videos is often higher than that of non-voter-fraud videos.

4.3 RQ3: What is the difference among voter-fraud videos from *(a)different parties* and *(b)different media types*, considering their audiovisual features?

RQ3(a) aims to examine the differences in audiovisual features of voter-fraud videos across the three main parties and the "Other" category. We conduct the analysis by presenting the ECDFs for audiovisual features across the four groups in Figure 5 and the KS test results in Table 4.

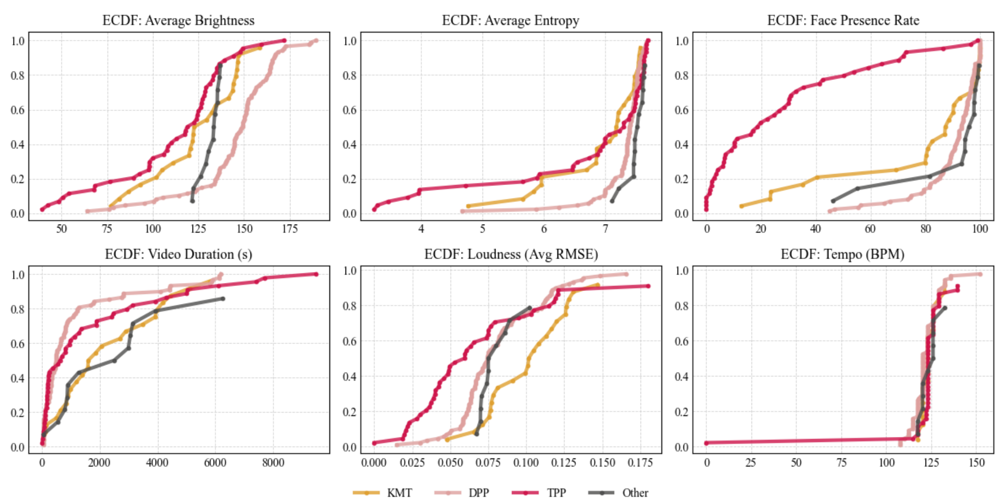

Fig. 5. Comparative Analysis of Empirical CDF of Audiovisual Features across Parties

Table 4. KS Test Results Comparing Audiovisual Features of Voter-Fraud Videos across Parties

	Brightness	**Entropy**	**Face Rate**	**Duration**	**Loudness**	**Tempo**
DPP vs TPP	0.636***	0.307**	0.864***	0.171	0.386***	0.386***
DPP vs KMT	0.504***	0.421**	0.311*	0.485***	0.409**	0.201
TPP vs KMT	0.227	0.273	0.655***	0.360**	0.534***	0.197
Other vs DPP	0.739***	0.328	0.221	0.516**	0.375*	0.166
Other vs TPP	0.500**	0.455*	0.773***	0.383	0.591***	0.344
Other vs KMT	0.387	0.423	0.351	0.149	0.512*	0.155

Note: We apply the KS test to all audiovisual features to compare each pair of parties. Each cell displays the KS statistic (rounded to three decimal places) along with its significance level (*p < .05; **p < .01; ***p < .001). The cell is colored according to the camp color (as shown in Figure 5) corresponding to the group with the higher median value for that attribute within the pair.

From the results, we observe the following:

(1) **Brightness** is highest in videos from the "DPP" camp (median = 149.290), followed by "KMT" videos (median = 122.720), with "TPP" videos (median = 120.720) exhibiting lower brightness. This suggests that "DPP" videos may use brighter visuals for emphasis or attention.

(2) **Entropy** is higher in "DPP" videos (median = 7.395) than in those from "KMT" videos (median = 7.180) and "TPP" videos (median = 7.300). This implies that "DPP" videos may incorporate more visual complexity, potentially to sustain viewer interest or present richer information.
(3) **Face Presence Rate** is highest in the videos from "DPP" camp (median = 93.280), indicating a focus on personal or direct communication—likely through interviews or monologues. In contrast, "TPP" (median = 19.230) has the lowest face presence rate, suggesting a different content style.
(4) **Duration** is longest in videos from the "KMT" camp (median = 1,602.300), while "DPP" videos (median = 489.960) and "TPP" videos (median = 658.345) tend to have shorter durations. This may reflect differences in content style and strategic communication preferences across parties.
(5) **Loudness** is lowest in the videos from "TPP" camp (median = 0.051), while "KMT" videos (median = 0.101) and "DPP" videos (median = 0.074) have higher values. This suggests that accounts supporting "DPP" or "KMT" may employ louder audio to attract attention or create a more engaging viewer experience.
(6) **Tempo** does not vary significantly across videos from the four categories, possibly due to standardized content production practices or the neutral nature of video pacing.

For partisan media outlets, "DPP" stands out as the most distinct camp in terms of audiovisual characteristics. Statistical results in Table 4 show that videos from "DPP" are significantly different from those by accounts supporting "TPP" in the features of *Tempo*, *Loudness*, *Brightness*, *Entropy*, and *Face Rate*, as well as significantly different from videos from"KMT" in *Loudness*, *Duration*, *Brightness*, *Entropy*, and *Face Rate*.

Comparing these three parties with the baseline "Other," the pair "Other vs DPP" shows the greatest divergence. This is consistent with previous findings that videos from "DPP" accounts exhibit the most distinctive characteristics. Notably, "TPP" is also statistically different from "Other" in the features of *Loudness*, *Brightness*, *Entropy*, and *Face Rate*. This distinction can be attributed to the different media sources: "Other" includes only traditional media accounts, whereas "TPP" consists entirely of grassroots media accounts, suggesting a clear difference in presentation styles.

RQ3(b) aims to examine the differences in voter fraud videos between traditional media and grassroots sources. We conduct an analysis similar to RQ2 by examining empirical cumulative distribution functions (ECDFs, Figure 6) and Kolmogorov-Smirnov (KS) test results (Table 5) for audiovisual features.

Table 5. KS Test Results Comparing Audiovisual Features of Voter Fraud Videos from Traditional Media and Grassroots

	Brightness	**Entropy**	**Face Rate**	**Duration**	**Loudness**	**Tempo**
KS Statistic	0.514	0.405	0.625	0.210	0.277	0.307
P-Value	<0.001	<0.001	<0.001	0.050	0.003	<0.001

Note: The cell is colored according to the media type color (as shown in Figure 6) corresponding to the type with the higher median value for that attribute. The level of transparency is set according to the value of the KS Statistic.

From the results, we observe the following:

(1) **Brightness**: Traditional media videos (median = 144.725) tend to have higher brightness levels than grassroots videos (median = 120.580), suggesting a preference for clearer or more polished visuals in professionally produced content.

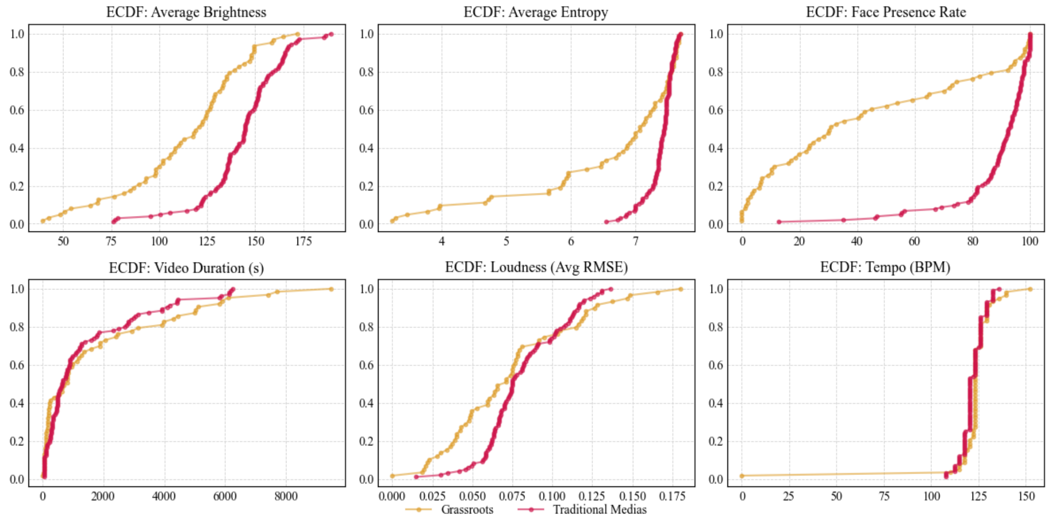

Fig. 6. Comparative Analysis of Empirical CDF of Audiovisual Features Between Voter Fraud Videos from Traditional Media and Grassroots

(2) **Entropy**: Traditional media videos (median = 7.450) display higher entropy than grassroots videos (median = 7.070), indicating a tendency toward more visually complex or information-rich content typical of professional production.
(3) **Face Presence Rate**: Traditional media videos (median = 92.905) exhibit a significantly higher face presence rate compared to grassroots videos (median = 30.810). This suggests that traditional media places greater emphasis on direct, personal communication—often through face-to-camera formats—while grassroots media may favor alternative styles with fewer face-centric visuals.
(4) **Duration**: Traditional media videos (median = 651.220) are generally shorter than those produced by grassroots media (median = 806.010), although this difference is not significant. This may reflect a consistent production style adopted by both groups, possibly due to content formatting norms.
(5) **Loudness**: Though traditional media videos (median = 0.075) appear slightly louder on average than grassroots videos (median = 0.071), the difference is not statistically significant.
(6) **Tempo**: The KS test reveals a significant difference in tempo between the two media types. Traditional media videos (median = 120.190), likely due to more standardized and structured production practices, show significantly lower variability in tempo than grassroots videos (median = 123.050).

In summary, traditional media accounts differ significantly from grassroots accounts in several audiovisual features. At the statistical significance level of $p = 0.001$, traditional media videos are significantly different from grassroots videos in the features of "Tempo (BPM)," "Average Brightness," "Average Entropy," and "Face Presence Rate." Overall, traditional media accounts tend to exhibit a more standardized tempo, higher brightness, greater visual complexity (entropy), and higher face presence compared to grassroots accounts.

Takeaways of RQ3 The answers to RQ3 can be summarized as:

(1) Voter-fraud videos from the "DPP" camp exhibit the most distinct audiovisual features, particularly in brightness, entropy, and face presence, compared to other parties.

(2) Voter-fraud videos from traditional media show significantly higher brightness, entropy, and face presence than those from grassroots media, reflecting more standardized and professional production styles even when discussing voter fraud misinformation.

5 Case Studies

In this section, we present two case studies to further illustrate the narratives surrounding voter fraud discussions by partisan media outlets, and then we highlight several key divergences compared with voter fraud misinformation in the 2020 U.S. presidential election, particularly in terms of narrative strategies.

Firstly, we examine the narrative strategies used in voter fraud discussions in Taiwan. To do so, we analyze every voter fraud video in our dataset and categorize them based on whether they support or refute the claim of voter fraud. We follow similar approach in Starbird [65]. In practice, we first develope a codebook based on preliminary observations to guide the identification of each voter-fraud video's stance. The complete codebook is provided in Appendix A.3.3. Specifically, if the video promotes the misinformation, we code it as supportive. Conversely, we code it as refuting. If evidence supporting or refuting the voter-fraud claim is presented without discernible bias in a video, we code it as neutral. Two authors independently viewed every voter-fraud video, discussed their judgments, and reached agreement on the final coding through deliberation, following procedures similar to Wang et al. [73]. In terms of the two cases, case 1 focuses on narratives that support the voter fraud claim, which are primarily driven by TPP and KMT media outlets. Case 2 highlights narratives that discredit the claim, mainly driven by DPP media outlets. Additionally, we explore the discourse supporting misinformation about voter fraud in Taiwan, and we also analyze its similarities and differences with that in the U.S.. Drawing on prior studies that have extensively examined voter fraud misinformation during the 2020 U.S. presidential election (see [56, 67], etc.), we identify several recurring patterns to compare with the voter fraud misinformation observed during the 2024 Taiwan election.

Case 1: Narratives that support the voter fraud issue. We identify three main narrative strategies, with two predominantly used by grassroots media and one by traditional media.

For grassroots videos from KMT and TPP, the first strategy involves directly presenting suspicious vote-counting scenes. We find 38 videos that fall into this category. These videos typically show questionable recordings of ballot counting at various polling stations, often accompanied by captions or commentary questioning the integrity of the process. An example is shown in Figure 7, where a volunteer is seen removing the bottom panel of a ballot box, revealing hidden ballots—a violation of proper election procedures. The second strategy by grassroots media involves expressing support for the belief that voter fraud is plausible. An example is shown in Figure 8 (left panel), where a host questions the decision to switch to an acrylic ballot box, implying it could facilitate fraud.

The third strategy is frequently used by traditional media from KMT, which promote the voter fraud narrative through news programming. An example is shown in Figure 8 (right panel), where an anchor reports on voter fraud allegations while presenting potentially suspicious footage. Although the program later includes a statement from the oversight committee affirming transparency, it does not directly refute the claim, thus leaving room for further speculation and discussion on platforms like YouTube.

Case 2: Narratives that refute voter fraud misinformation. We classify the videos refuting voter fraud misinformation into two categories.

The first category includes traditional media outlets affiliated with DPP, which use news programs to debunk voter fraud claims. An example is shown in the left panel of Figure 9, where the program includes captions such as "More than 10 cell phones were filming to monitor the ballot stations. How would anyone have time to cheat? Ko's supporters are being ridiculous."

Note: Top text (on red background) "What did I just see?"

Note: Top text (on red background) "What did I just see?" Bottom text (on orange background) "The white bottom panel was lifted, a bunch of ballots fell out."

Fig. 7. A Voter-Fraud Video Case from TPP: Featuring the first strategy, where videos directly present suspicious vote-counting scenes. These images are screenshots extracted from videos in our compiled YouTube dataset.

Note: Top text (red and white background) "Before counting votes, ballots must not be sorted..." Bottom text: "Lawmaker Cheng Cheng-ling: Why must it be changed to an acrylic box?"

Note: Main headline (yellow and white text on black background) "Ballot-counting chaos, many irregularities in vote counting. Blank ballots turned into counted votes, people caught red-handed!"

Fig. 8. Voter-Fraud Video Cases: Left panel shows the second strategy (grassroots commentary) exemplified by a video from grassroots media of TPP ; right panel shows the third strategy (news programming) exemplified by a video from traditional media of KMT. These images are screenshots extracted from videos in our compiled YouTube dataset.

The second category consists mostly of grassroots videos from DPP. These videos provide more detailed explanations and often use a flexible, narrative-driven style to communicate. One example, also shown in the right panel of Figure 9, features a content creator who walks through the entire voting process to explain why allegations of voter fraud are implausible.

Comparison with narratives supporting voter fraud misinformation in the U.S. A notable similarity across both contexts is the central role played by grassroots media in initiating

Note: Subtitle "Ko supporters' traffic flow." Bottom headline: "10 mobile phones monitoring the vote counting on-site! Any time to manipulate votes!? Ko supporters are unreasonably causing trouble!?"

Note: Main headline: "Ballot stacking? Miscounting? Did someone really manipulate the votes?" Subtext with cartoon dog: "Not done talking yet, join the 'Topic Wishing Pool'!"

Fig. 9. Voter-Fraud Video Cases from DPP: Left panel shows traditional media refuting misinformation; right panel shows a grassroots video providing detailed commentary. These images are screenshots extracted from videos in our compiled YouTube dataset.

and amplifying voter fraud misinformation. Narratives alleging vote tampering were frequently disseminated in the U.S., with similar patterns observed in Taiwan, often accompanied by visual content such as photos or videos taken at polling stations. Starbird et al. [67] discussed cases in which grassroots accounts shared images of ballots during the 2020 U.S. election. Similarly, in the 2024 Taiwan election, Figure 7 exemplifies how grassroots media supported voter fraud claims. By presenting such seemingly direct evidence, these accounts aimed to enhance the credibility of their narratives and boost audience engagement.

However, a key divergence lies in the temporal framing and manipulation of visual evidence. In the U.S., misinformation narratives often repurposed images or videos from previous election cycles, falsely portraying them as evidence of fraud in the 2020 presidential election [67]. This recycling of past materials often went unnoticed by the public, especially when contextual metadata was removed. In contrast, voter fraud misinformation in the 2024 Taiwan election mostly utilized footage genuinely captured during the current election (Figure 7, 8). Rather than fabricating the temporal context, these narratives distorted or exaggerated interpretations of contemporary scenes—for instance, misrepresenting standard electoral procedures or a decision to use acrylic ballot boxes as suspicious activities—to construct allegations of fraud.

6 Discussion

In the discussion section, we begin by interpreting the key findings of our study. We then outline the implications of these results for misinformation research, highlighting how our analysis contributes to a deeper understanding of misinformation dissemination. Next, we explore how social media platforms can leverage our insights to improve content moderation and policy design. Additionally, we discuss the broader relevance of our findings for the CSCW community. Finally, we acknowledge the limitations of our study and suggest directions for future research.

6.1 Illustrating the dissemination of voter fraud misinformation during the 2024 Taiwan election

This paper presents an illustration of voter fraud misinformation controversy on YouTube during the 2024 Taiwan election. Generally, we find that the dissemination of voter fraud-related content was highly time-sensitive and clustered primarily within the first week after the election. This temporal concentration suggests that the voter fraud narrative was largely a reactionary discourse rather than a long-term strategic campaign. The two major parties driving this discourse were the

DPP and the TPP, both of which engaged in post-election media responses, though with markedly divergent tones and strategies. Notably, the TPP actively contributed to the circulation of such misinformation, amplifying public skepticism toward the electoral process. In contrast, the DPP primarily sought to counteract the dissemination of voter fraud misinformation, positioning itself as a defender of electoral integrity (**RQ1**).

However, voter fraud videos did not gain proportionate visibility or virality within the broader information ecosystem. Despite the political salience of the topic, such videos represented only a moderate share of daily video outputs and did not outperform non-fraud-related content in terms of views. This finding suggests that the voter fraud narrative was largely confined to a specific subset of politically engaged users. However, the relatively high comment counts on voter fraud videos reflect that within this niche audience, to whom the topic remained contentious and provoked polarized discussions. This dual pattern—limited reach but intense engagement—demonstrates that voter fraud misinformation may have a restricted influence on general Taiwan YouTube users but it can deeply affect a small group of politically engaged users (**RQ2**).

Furthermore, the analysis of audiovisual features reveals stark contrasts across both political and media lines- divergent communication strategies connected with different media types. Videos associated with the DPP were characterized by louder audio, higher brightness, greater visual complexity, and higher facial presence—features consistent with content produced by traditional media, whose outputs tend to cluster around a consistent speech rate with low variance, higher brightness, and standardized visual presentation—features indicative of professionalized production norms (**DPP-traditional media**). Conversely, TPP-affiliated videos exhibited traits commonly associated with grassroots or alternative media: lower volume, dimmer lighting, simpler compositions, and minimal facial appearance. Grassroots displayed more erratic speech patterns and looser visual structure, highlighting the decentralized and often improvisational nature of alternative content creation (**TPP-grassroots media**). DPP's reliance on traditional and professionalized channels provided it with a more controlled and coherent narrative presence, whereas TPP's dependence on decentralized grassroots media, though more spontaneous, appeared less capable of sustaining attention or breaking into mainstream discourse. (**RQ3**).

Finally, we conduct case studies on narratives both supporting and refuting the voter fraud misinformation. The grassroots media, especially TPP-affiliated grassroots media, were inclined to upload videos of verisimilar scenes where poll workers were manipulating the posts to initiate the controversy, which were further disseminated through traditional KMT media. This may indicate that there is a coordinated influence operation within the TPP-affiliated grassroots media or even coordination between the KMT and TPP. In contrast, DPP-affiliated media refutes the misinformation by casting doubt on the motivation of TPP supporters who claimed the existence of voting fraud or by explaining the voting process comprehensively to justify the fairness of the election. Explanation in a proper way could be useful for future fact-checking agencies since vivid and interesting videos may help the ordinary people better understand the whole process of voting and build confidence for the election results (**Case Studies**).

These findings illustrate a contested information space where traditional and grassroots media waged a post-election narrative battle over voter fraud. However, the limited diffusion and engagement of fraud-related content ultimately point to the relative inefficacy of the grassroots-media-driven misinformation push. The inability of the TPP's grassroots media network to sustain or escalate the voter fraud narrative reinforces the structural advantage of the DPP's institutionalized media apparatus. This also resonates with broader insights in misinformation research, which suggest that while grassroots media networks can generate momentary disruption, they often lack the infrastructural capacity to outpace more established communication systems—especially when countered by professionalized responses.

The results of our research could be valuable for election-related studies, particularly in the context of election integrity for future elections. With our annotated dataset in hand, researchers could easily replicate our workflow for the automated analysis of audiovisual features. Specifically, practical use in a new context is threefold: 1) Obtain or auto-generate subtitles using YouTube Auto Speech Recognition (ASR) or other open-source ASR. 2) If needed, machine-translate text to a language supported by the LLM, since our prompt is language-independent aside from label keywords. 3) Feed transcripts to run the LLM-assisted annotation pipeline and run the identical audiovisual-feature extractor. Conclusively, the pipeline is language-agnostic: given subtitles or transcripts, the same human–AI annotation workflow and multimodal feature extraction can be applied to other research contexts, enabling rapid comparative misinformation studies across regions.

6.2 Implications for Misinformation Research

Regarding our dataset construction pipeline, this research provides both valuable resources and annotation insights to support future misinformation studies. Firstly, we compile a comprehensive list of traditional media outlets and grassroots YouTube accounts in Taiwan, offering a useful starting point for CSCW researchers interested in Taiwanese politics but unfamiliar with the local media landscape. This curated list can serve as a foundation for building Taiwan-focused misinformation datasets on YouTube. Secondly, our dataset and accompanying codebooks offer practical value for developing automated and explainable techniques to detect multimodal voter fraud misinformation. This can facilitate future annotation tasks for video-based misinformation, similar to prior automated approaches in hate speech detection [46]. Furthermore, our use of LLMs to annotate misinformation highlights both the promise and limitations of such models, particularly in terms of alignment and reliability. This encourages future research to investigate prompt design and model bias using our released dataset. In addition, while content moderation has proven effective in limiting the spread of misinformation [28, 42, 79], most existing work focuses on English-language content [28, 79], leaving Chinese-language content relatively underexplored. Our annotated dataset provides a rare resource for future researchers to investigate moderation practices in Chinese and compare cross-linguistic differences in misinformation control.

As for illustrating and understanding misinformation controversies, our research offers a more comprehensive perspective that future misinformation studies may adopt: (1) a full-picture analysis of both supportive and oppositional media narratives; (2) a platform-specific investigation of user engagement patterns; (3) an audiovisual feature comparison grounded in multimodal misinformation frameworks; and (4) case studies suggesting the presence of coordinated information operations. Together, these components highlight the multifaceted nature of misinformation dynamics and offer actionable insights for future CSCW research.

A comprehensive analysis of both supportive and oppositional media narratives. Unlike previous studies on voter fraud misinformation [58, 67], which primarily focus on participants promoting misinformation, our analysis includes media accounts regardless of their stance. Our findings reveal that media affiliated with the Democratic Progressive Party (DPP, which won the election and actively refuted the voter fraud claims) and the Taiwan People's Party (TPP, which lost the election and was the primary source of those claims) shared the largest number of videos discussing voter fraud misinformation. Notably, DPP-affiliated accounts disseminated 88 videos refuting such claims—accounting for more than half of the total voter fraud misinformation-related videos and representing twice the number of videos posted by TPP-affiliated channels. This suggests that DPP-affiliated media played an active role in countering misinformation. These results highlight the importance of including both supportive and oppositional narratives in misinformation research, and suggest that future CSCW studies on voter fraud misinformation should consider incorporating media on both sides of the controversy to enable more holistic analyses.

A platform-specific investigation of user engagement patterns. Building on a bottom-up approach [73], our research further contributes to the literature on the relationship between misinformation and user engagement on YouTube. In contrast to [73], which finds that tweets containing COVID-19 misinformation images tend to receive more retweets and likes than those with random images, we observe a different pattern on YouTube: videos discussing voter fraud misinformation tend to gain fewer views but more comments compared to non-voter fraud videos. One possible explanation is that YouTube views are only counted after a user watches at least 30 seconds of a video,[10] while commenting also requires more cognitive effort and engagement than the relatively effortless acts of liking or retweeting on Twitter. This suggests that user engagement dynamics on Twitter may not directly translate to YouTube. Future CSCW misinformation researchers can build on our investigation to explore the psychological mechanisms that may underlie these cross-platform differences in engagement behavior.

An audiovisual feature comparison grounded in multimodal misinformation frameworks. Our research adopts multimodal misinformation frameworks proposed in prior work [9, 44, 55] to understand voter fraud misinformation videos on distinct audiovisual features. Among our findings, we observe statistically significant differences between the Democratic Progressive Party (DPP) and the Taiwan People's Party (TPP)—the two primary actors in this discourse—across several key audiovisual variables, including brightness, entropy, face rate, loudness, and tempo. This suggests that such frameworks enable researchers to uncover nuanced visual and auditory tactics employed by competing sides in a misinformation controversy, extending the analytical lens beyond textual or metadata-based comparisons. Building on prior studies [9, 44, 55], which show that these audiovisual characteristics may shape audience persuasion, future CSCW scholars can further investigate the media effects and psychological mechanisms under these multimodal tactics, using our findings as a foundation.

Case studies suggesting the presence of coordinated information operations. As part of our analysis of the voter fraud misinformation controversy, we conducted a series of case studies to investigate potential coordination patterns. Notably, we observed that probable scenes depicting alleged voting fraud—initially disseminated by TPP-affiliated grassroots media—were later amplified by KMT-affiliated traditional media outlets. This sequential dissemination pattern suggests a possible coordinated information operation between the TPP and KMT to undermine the perceived legitimacy of the election, aligning with prior definitions of coordinated disinformation campaigns [66]. Our findings thus offer preliminary evidence and analytical clues for future CSCW research seeking to identify and study coordinated information operations in the context of the 2024 Taiwan election [66].

6.3 Implications for platform moderation

This paper investigates the controversy surrounding voter fraud misinformation during the 2024 Taiwan election, highlighting the contrasting roles of partisan media: TPP-affiliated grassroots media were instrumental in promoting voter fraud narratives, while DPP-affiliated traditional media actively worked to defend the integrity of the electoral process. Similar to Donald Trump, TPP leader Ko Wen-je represents an atypical politician who emphasizes personal charisma. Following their respective electoral defeats, supporters of both figures turned to social media to disseminate unfounded claims about election fraud. To prevent the recurrence of such misinformation-driven

[10] See https://support.google.com/youtube/answer/2991785

dynamics in future elections, our study identifies key opportunities and challenges for social media platforms like YouTube to more effectively moderate and respond to election-related misinformation.

One opportunity lies in leveraging face presence rate for early detection. Firstly, we find that TPP-affiliated media—which consist solely of grassroots accounts—exhibit the lowest face presence rate among all selected parties. Given that TPP-affiliated grassroots media are key sources for disseminating voter fraud misinformation, this finding suggests that low face presence may serve as a potential indicator of misinformation-promoting content. One plausible explanation is that these grassroots account operators intentionally avoid showing their faces to remain anonymous and evade accountability. This audiovisual feature may therefore be integrated into content moderation tools to assist platforms like YouTube in identifying videos that potentially contain false claims about election integrity. Future work could further test the diagnostic power of face presence across elections and platforms before deploying it at scale.

Another opportunity is identifying early warning signs from cross-election patterns. Our case studies—interacted with the 2020 U.S. election—suggest that the emergence of false narratives about election integrity in grassroots media is particularly following the defeat of a candidate with traits corresponding to an atypical politician (e.g., Donald Trump, Ko Wen-je, both known for their strong personal charisma and highly devoted supporter bases.). This may serve as an early signal of broader misinformation controversies on social media. These narratives are often later amplified by other outlets, including traditional media. Accordingly, content moderators on platforms like YouTube should treat such patterns as early warning signs and prepare timely interventions to mitigate the spread of misinformation.

Besides, our investigation highlights several challenges and vulnerabilities in current platform moderation practices requiring further consideration. **A key challenge is the complexity of multimodal misinformation detection at scale.** Video-sharing platforms such as YouTube can utilize our released codebase to improve their capacity for recognizing voter fraud misinformation at scale. In the long term, we argue that platforms should invest in the development of automated multimodal detection techniques, which could be trained and evaluated using our dataset, to help curb the spread of such misinformation content. **Another challenge lies in navigating the blurred line between refutation and amplification.** This research also suggests that YouTube's content moderation policies may need to be reviewed to better address the nuanced ways election misinformation can spread. As our case studies show, even in the absence of direct confirmation, KMT-affiliated traditional media outlets aired voter fraud misinformation without providing proper refutations—possibly due to partisan interests—thus allowing the narrative to persist and circulate on the platform. To address this, platforms must consider the subtle distinction between actively refuting misinformation and passively enabling it through uncritical amplification.

Contribution to the CSCW community Overall, we have made several contributions to the CSCW community.

From the perspective of methodology:

- Similar to previous work published in CSCW [1, 58, 73], our study focuses on election integrity, a key topic within misinformation research. Our findings align with other CSCW research, such as Jakesch et al. [26], which emphasizes the need for further investigations into elections beyond those in the U.S.. To the best of our knowledge, this study presents the first empirical investigation into voter fraud misinformation during the Taiwan's 2024 election. Additionally, we compiled and annotated a comprehensive list of Taiwanese YouTube media accounts, providing a valuable resource for CSCW scholars interested in exploring voter fraud misinformation in the Taiwanese context.

- We establish a multimodal computational pipeline to detect videos discussing voter fraud misinformation at scale. Future CSCW scholars can replicate our results and even develop fully automated methods to detect videos with voter fraud misinformation.
- We further introduce the audiovisual analysis framework from [9, 44, 55] to the CSCW community. While prior CSCW research has explored misinformation in video content (e.g., [16, 25, 66]), such studies have primarily focused on metadata [16] or semantic content in videos [25, 66], without systematically examining the visual and auditory elements of the videos themselves. Our study addresses this gap by applying the audiovisual framework from [9, 44, 55] to analyze videos related to voter fraud misinformation. This approach enables us to identify and compare the visual and auditory strategies employed by different parties, as well as distinguish between traditional media and grassroots media tactics. In doing so, we offer a more nuanced and comprehensive understanding of how audiovisual cues shape the controversy surrounding voter fraud misinformation.

From the perspective of our findings:

- In the context of an electoral defeat for an atypical political figure, we examine how voter fraud misinformation is disseminated. Notably, we find that voter-fraud videos receive fewer views but generate more comments compared to non-voter-fraud videos. This contrasts with previous CSCW research on COVID-19 misinformation on Twitter, where misinformation content often garnered broader exposure [73]. These findings may offer new directions for future CSCW studies on user engagement with misinformation on different platforms.
- Our analysis also uncovers significant audiovisual differences between videos produced by the DPP and TPP, particularly in terms of brightness, entropy, face rate, loudness, and tempo. According to prior research [9, 44, 55], such features can influence media effects or psychological perceptions on audiences. These findings open up further avenues for CSCW scholars to explore how audiovisual presentation shapes audience perception in political communication.

From the perspective of CSCW Themes:

- *Sociotechnical Collaboration.* Misinformation constitutes a persistent societal challenge, posing significant risks to public health [68] and even triggering chaos in a country [21]. In response, CSCW researchers have developed diverse technical solutions, including AI assistants for fact-checking [43] and multimodal knowledge graphs to detect fauxtography [35]. Our work extends this research direction by designing a novel LLM-based computational pipeline to facilitate the analysis of videos that contain misinformation narratives at scale. Future CSCW scholars can apply this pipeline to address other video-based misinformation. Moreover, our research further reveals that combating voter-fraud misinformation requires layered cooperation among computer researchers, social scientists, fact-checking agencies, governments, and social media platforms. These intertwined collaborations exemplify the sociotechnical collaboration perspective central to CSCW.
- *Platform Governance.* Moderating harmful content, including hate speech [14] and misinformation [6] represents a critical research focus within CSCW communities, given these platforms' integral role in contemporary society. Our work contributes significant insights to this domain. Specifically, we find that the contrasting face-presence rates and narrative strategies uncovered in Section 5 expose how professional versus grassroots actors exploit YouTube's recommendation mechanism differently. These findings inform governance debates about whether automated moderation should weight audiovisual signals (e.g. face presence) when flagging content relevant to voter fraud, and caution that low-budget producers could be disproportionately impacted if such thresholds are not resource-aware.

- *Civic Information Infrastructure.* Prior CSCW research has devoted substantial attention to investigating how civic information infrastructures influence society. This includes exploring the application of information and communication systems to support earthquake rescue efforts [53] and examining the impact of smart city technologies on underserved communities [13]. Our work demonstrates its potential to inspire future CSCW studies in this area of research. Specifically, we identified a lack of official videos to counter videos alleging voter fraud. Future CSCW research could focus on designing systems to better monitor the vote-counting process. For instance, developing strategies for the effective placement of cameras in poll stations could be explored to enhance trustworthiness in voting systems.

Limitations of our study The first limitation of this paper is that unlike previous research in platforms like Twitter [8], where data was collected in a real-time streaming service, our research collects data after all the videos have been posted. As a result, some videos might have been deleted before our data collection process started, making them unable to be collected into our dataset. Future research investigating elections on YouTube may predetermine the accounts before the election starts and then collect videos immediately after the videos are posted to ensure every posted video is collected.

The second limitation of our study is that the choice of grassroots YouTube accounts may be influenced by biases introduced by the recommendation algorithms of Google because there is not an existing report including a complete list of grassroots YouTube accounts. Nonetheless, we release this useful list of Taiwan grassroots YouTube accounts, which will be of great help to future research about Taiwan politics on YouTube.

7 Conclusion

This research provides a detailed analysis of how different parties and media types on YouTube engaged with misinformation, especially voter fraud discussions during the Taiwan's 2024 election. We identify related traditional and grassroots YouTube channels, and then used a self-collected set of 5,641 videos published from January 12th, 2024 to Feburary 1st, 2024 to construct the dataset with a combination of LLM annotations and manual annotations, resulting in 170 videos tagged as voter fraud-related for further analysis.

The engagement metrics—views, likes, and comments—are examined to understand the reception of misinformation (i.e., voter fraud) content. Our results indicate that videos discussing voter fraud display lower views but higher comments than non-voter-fraud videos, suggesting that this kind of controversial or politically charged content often drives more discussion despite limited focus.

The audiovisual features of the videos are measured systematically and analyzed to explore differences in content style across parties and media types. Our results reveal that DPP videos, in particular, showed higher brightness and face presence, indicating a visually engaging and human-centered approach, which contrasts with the lower brightness and face presence in TPP videos. Traditional media, compared to grassroots media, had significantly higher brightness, entropy, and face presence, suggesting a more formal production style often associated with professional media.

Overall, our research studies how voter fraud discussions are strategically framed and differentiated across media types and parties on YouTube. These findings highlight the need for advanced moderation techniques, especially for multimodal and non-English misinformation, and call for more cross-platform investigations to better understand how election misinformation spreads and influences public discourse. Our research is potentially informing future studies on election misinformation in different political contexts, e.g., the U.S. 2024 election.

Acknowledgments

This work was supported by Renmin University of China under the grants 2024300111 and 202430061, by the grant 2023JZDZ033 from Ministry of Education of the People's Republic of China, and by the grant GZC20233083 from China Postdoctoral Science Foundation.

References

[1] Syeda Zainab Akbar, Anmol Panda, Divyanshu Kukreti, Azhagu Meena, and Joyojeet Pal. 2021. Misinformation as a Window into Prejudice: COVID-19 and the Information Environment in India. *Proceedings of the ACM on human-computer interaction* 4, CSCW3 (2021), 1–28.

[2] Patel Janakkumar Baldevbhai. 2012. Color Image Segmentation for Medical Images using L*a*b* Color Space. *IOSR Journal of Electronics and Communication Engineering* 1 (2012), 24–45.

[3] Yoshua Bengio, Geoffrey Hinton, Andrew Yao, Dawn Song, Pieter Abbeel, Trevor Darrell, Yuval Noah Harari, Ya-Qin Zhang, Lan Xue, Shai Shalev-Shwartz, et al. 2024. Managing extreme AI risks amid rapid progress. *Science* 384, 6698 (2024), 842–845.

[4] Yochai Benkler, Casey Tilton, Bruce Etling, Hal Roberts, Justin Clark, Robert Faris, Jonas Kaiser, and Carolyn Schmitt. 2020. Mail-in voter fraud: Anatomy of a disinformation campaign. *Berkman Center Research Publication* 2020-6 (2020).

[5] Lia Bozarth and Ceren Budak. 2020. Toward a better performance evaluation framework for fake news classification. In *Proceedings of the international AAAI conference on web and social media*, Vol. 14. 60–71.

[6] Lia Bozarth, Jane Im, Christopher Quarles, and Ceren Budak. 2023. Wisdom of Two Crowds: Misinformation Moderation on Reddit and How to Improve this Process—A Case Study of COVID-19. *Proc. ACM Hum.-Comput. Interact.* 7, CSCW1, Article 155 (April 2023), 33 pages. https://doi.org/10.1145/3579631

[7] Ho-Chun Herbert Chang and Yu Sunny Fang. 2024. The 2024 Taiwanese Presidential Election on social media: Identity, policy, and affective virality. *PNAS nexus* 3, 4 (2024), 130.

[8] Emily Chen, Kristina Lerman, and Emilio Ferrara. 2020. Tracking Social Media Discourse About the COVID-19 Pandemic: Development of a Public Coronavirus Twitter Data Set. *JMIR Public Health and Surveillance* 6, 2 (2020), e19273.

[9] Kaiping Chen, Sang Jung Kim, Qiantong Gao, and Sebastian Raschka. 2022. Visual Framing of Science Conspiracy Videos: Integrating Machine Learning with Communication Theories to Study the Use of Color and Brightness. *Computational Communication Research* 4, 1 (May 2022), 98–134.

[10] Eun Cheol Choi and Emilio Ferrara. 2024. Automated Claim Matching with Large Language Models: Empowering Fact-Checkers in the Fight Against Misinformation. Association for Computing Machinery, New York, NY, USA.

[11] I. Culjak, D. Abram, Tomislav Pribanic, Hrvoje Džapo, and Mario Cifrek. 2012. A brief introduction to OpenCV. 1725–1730.

[12] Alexandre Debant and Lucca Hirschi. 2023. Reversing, Breaking, and Fixing the French Legislative Election {E-Voting} Protocol. In *32nd USENIX Security Symposium (USENIX Security 23)*. 6737–6752.

[13] Jessa Dickinson, Mark Díaz, Christopher A Le Dantec, and Sheena Erete. 2019. " The cavalry ain't coming in to save us" Supporting Capacities and Relationships through Civic Tech. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–21.

[14] Jacob Erickson and Bei Yan. 2025. Content Moderation and Hate Speech on Alternative Platforms: A Case Study of BitChute. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW052 (May 2025), 18 pages. https://doi.org/10.1145/3710950

[15] Dafydd Fell. 2006. *Party politics in Taiwan: party change and the democratic evolution of Taiwan, 1991-2004*. Routledge.

[16] KJ Kevin Feng, Nick Ritchie, Pia Blumenthal, Andy Parsons, and Amy X Zhang. 2023. Examining the impact of provenance-enabled media on trust and accuracy perceptions. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–42.

[17] Kiran Garimella and Dean Eckles. 2020. Images and misinformation in political groups: Evidence from WhatsApp in India. *Harvard Kennedy School Misinformation Review* (2020).

[18] Adam Geitgey. 2020. Face recognition: The world's simplest facial recognition API for python and the command line. https://github.com/ageitgey/face_recognition

[19] Zhijiang Guo, Michael Schlichtkrull, and Andreas Vlachos. 2022. A survey on automated fact-checking. *Transactions of the Association for Computational Linguistics* 10 (2022), 178–206.

[20] J Alex Halderman. 2022. The Antrim county 2020 election incident: an independent forensic investigation. In *31st USENIX Security Symposium (USENIX Security 22)*. 589–605.

[21] Md Mahfuzul Haque, Mohammad Yousuf, Ahmed Shatil Alam, Pratyasha Saha, Syed Ishtiaque Ahmed, and Naeemul Hassan. 2020. Combating misinformation in Bangladesh: Roles and responsibilities as perceived by journalists, fact-checkers, and users. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW2 (2020), 1–32.

[22] Ming-Sho Ho. 2024. Taiwan 2024 election: A social movement perspective. *21st Century* 201 (2024), 35–42.
[23] Liang-Yi Hsu and Shao-Hong Lin. 2023. Party identification and YouTube usage patterns: An exploratory analysis. *Taiwan Politics* (2023).
[24] Chen-Ling Hung, Wen-Cheng Fu, Chang-Ce Liu, and Hui-Ju Tsai. 2024. AI Disinformation Attacks and Taiwan's Responses during the 2024 Presidential Election. *Thomson Foundation* (2024).
[25] Eslam Hussein, Prerna Juneja, and Tanushree Mitra. 2020. Measuring misinformation in video search platforms: An audit study on YouTube. *Proceedings of the ACM on human-computer interaction* 4, CSCW1 (2020), 1–27.
[26] Maurice Jakesch, Kiran Garimella, Dean Eckles, and Mor Naaman. 2021. Trend alert: A cross-platform organization manipulated Twitter trends in the Indian general election. *Proceedings of the ACM on Human-computer Interaction* 5, CSCW2 (2021), 1–19.
[27] Sandra Jeppesen, Henry Giroux, Michael Hoechsmann, Chenjerai Kumanyika, David VanDyke, Miranda McKee, et al. 2022. *The Capitol riots: digital media, disinformation, and democracy under attack*. Routledge.
[28] Shagun Jhaver, Alice Qian Zhang, Quan Ze Chen, Nikhila Natarajan, Ruotong Wang, and Amy X Zhang. 2023. Personalizing content moderation on social media: User perspectives on moderation choices, interface design, and labor. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW2 (2023), 1–33.
[29] Jungseock Joo and Zachary C Steinert-Threlkeld. 2022. Image as data: Automated content analysis for visual presentations of political actors and events. *Computational Communication Research* 4, 1 (2022).
[30] Paul Joosse and Dominik Zelinsky. 2023. Charismatic mimicry: Innovation and imitation in the case of volodymyr zelensky. *Sociological Theory* 41, 3 (2023), 201–228.
[31] Yeliz Karaca and Majaz Moonis. 2022. *Chapter 14 - Shannon entropy-based complexity quantification of nonlinear stochastic process: diagnostic and predictive spatiotemporal uncertainty of multiple sclerosis subgroups*. Academic Press. 231–245 pages.
[32] Davinder Kaur, Suleyman Uslu, Kaley J Rittichier, and Arjan Durresi. 2022. Trustworthy artificial intelligence: a review. *ACM computing surveys (CSUR)* 55, 2 (2022), 1–38.
[33] Wen-je Ko. 2023. *Ko Wen-je's Taiwan Notebook*. Business Weekly Publications, a division of Cité Publishing Ltd.
[34] Tadayoshi Kohno, Adam Stubblefield, Aviel D Rubin, and Dan S Wallach. 2004. Analysis of an electronic voting system. In *IEEE Symposium on Security and Privacy, 2004. Proceedings. 2004*. IEEE, 27–40.
[35] Ziyi Kou, Yang Zhang, Daniel Zhang, and Dong Wang. 2022. CrowdGraph: A crowdsourcing multi-modal knowledge graph approach to explainable fauxtography detection. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–28.
[36] JR Landis. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* (1977).
[37] Md Tahmid Rahman Laskar, M Saiful Bari, Mizanur Rahman, Md Amran Hossen Bhuiyan, Shafiq Joty, and Jimmy Huang. 2023. A Systematic Study and Comprehensive Evaluation of ChatGPT on Benchmark Datasets. In *Findings of the Association for Computational Linguistics: ACL 2023*, Anna Rogers, Jordan Boyd-Graber, and Naoaki Okazaki (Eds.). Association for Computational Linguistics, Toronto, Canada, 431–469. https://doi.org/10.18653/v1/2023.findings-acl.29
[38] Crystal Lee, Tanya Yang, Gabrielle D Inchoco, Graham M Jones, and Arvind Satyanarayan. 2021. Viral visualizations: How coronavirus skeptics use orthodox data practices to promote unorthodox science online. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–18.
[39] Chao-Chen Lin. 2020. The models of fake news and media convergence: An exploration on the 2018 Taiwan election. *Journalism Research Journal* 142 (2020), 111–53.
[40] Tsong-Jyi Lin. 2007. Social Networks, Political Discussions, and Voting Participations. *Journal of Electoral Studies* 14, 2 (2007), 1–24.
[41] Bernard Lindgren. 2017. *Statistical theory*. Routledge.
[42] Chen Ling, Krishna P Gummadi, and Savvas Zannettou. 2023. " Learn the Facts About COVID-19": Analyzing the Use of Warning Labels on TikTok Videos. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 17. 554–565.
[43] Houjiang Liu, Anubrata Das, Alexander Boltz, Didi Zhou, Daisy Pinaroc, Matthew Lease, and Min Kyung Lee. 2024. Human-centered NLP Fact-checking: Co-Designing with Fact-checkers using Matchmaking for AI. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW2 (2024), 1–44.
[44] Yingdan Lu and Cuihua (Cindy) Shen. 2023. Unpacking Multimodal Fact-Checking: Features and Engagement of Fact-Checking Videos on Chinese TikTok (Douyin). *Social Media + Society* 9, 1 (2023).
[45] Yueh-Chi Lu. 2020. Analyzing Taiwanese Politicians' Social Media Adoption On Facebook, YouTube, And Instagram. *Master's Thesis* 2020 (2020), 1–115.
[46] Enrico Mariconti, Guillermo Suarez-Tangil, Jeremy Blackburn, Emiliano De Cristofaro, Nicolas Kourtellis, Ilias Leontiadis, Jordi Luque Serrano, and Gianluca Stringhini. 2019. " You Know What to Do" Proactive Detection of YouTube Videos Targeted by Coordinated Hate Attacks. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–21.

[47] Brian McFee, Colin Raffel, Dawen Liang, Daniel Ellis, Matt Mcvicar, Eric Battenberg, and Oriol Nieto. 2015. librosa: Audio and Music Signal Analysis in Python. 18–24.
[48] Ryan C Moore, Ross Dahlke, and Jeffrey T Hancock. 2023. Exposure to untrustworthy websites in the 2020 US election. *Nat Hum Behav* 7 (2023), 1096–1105.
[49] Shmyrev Nickolay V and other contributors. 2020. Vosk Speech Recognition Toolkit: Offline speech recognition API for An- droid, iOS, Raspberry Pi and servers with Python, Java, C# and Node. https://github.com/alphacep/vosk-api
[50] Pippa Norris. 2014. *Why electoral integrity matters.* Cambridge University Press.
[51] Kabir Panahi, Shawn Robertson, Yasemin Acar, Alexandru G Bardas, Tadayoshi Kohno, and Lucy Simko. 2024. " But they have overlooked a few things in {Afghanistan:}" An Analysis of the Integration of Biometric Voter Verification in the 2019 Afghan Presidential Elections. In *33rd USENIX Security Symposium (USENIX Security 24)*. 2047–2064.
[52] Kostantinos Papadamou, Savvas Zannettou, Jeremy Blackburn, Emiliano De Cristofaro, Gianluca Stringhini, and Michael Sirivianos. 2021. " How over is it?" Understanding the Incel Community on YouTube. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–25.
[53] Shreyasha Paudel, Wendy Norris, and Robert Soden. 2024. Aftermath: Infrastructure, Resources, and Organizational Adaptation in the Wake of Disaster. *Proceedings of the ACM on Human-Computer Interaction* 8, CSCW1 (2024), 1–24.
[54] Yilang Peng. 2018. Same candidates, different faces: Uncovering media bias in visual portrayals of presidential candidates with computer vision. *Journal of Communication* 68, 5 (2018), 920–941.
[55] Yilang Peng, Yingdan Lu, and Cuihua Shen. 2023. An agenda for studying credibility perceptions of visual misinformation. *Political Communication* 40, 2 (2023), 225–237.
[56] Gordon Pennycook and David G Rand. 2021. Examining false beliefs about voter fraud in the wake of the 2020 Presidential Election. *The Harvard Kennedy School Misinformation Review* (2021).
[57] Jorge Poco and Jeffrey Heer. 2017. Reverse-engineering visualizations: Recovering visual encodings from chart images. In *Computer graphics forum*, Vol. 36. Wiley Online Library, 353–363.
[58] Stephen Prochaska, Kayla Duskin, Zarine Kharazian, Carly Minow, Stephanie Blucker, Sylvie Venuto, Jevin D West, and Kate Starbird. 2023. Mobilizing manufactured reality: How participatory disinformation shaped deep stories to catalyze action during the 2020 US presidential election. *Proceedings of the ACM on human-computer interaction* 7, CSCW1 (2023), 1–39.
[59] J. Rasmus. 2020. *The Scourge of Neoliberalism: US Economic Policy from Reagan to Trump.* Clarity Press, Incorporated. https://books.google.co.jp/books?id=L_W2wAEACAAJ
[60] Gustavo Resende, Philipe Melo, Hugo Sousa, Johnnatan Messias, Marisa Vasconcelos, Jussara Almeida, and Fabrício Benevenuto. 2019. (Mis) information dissemination in WhatsApp: Gathering, analyzing and countermeasures. In *The World Wide Web Conference*. 818–828.
[61] Robert C Rowland. 2018. Donald Trump and the rejection of the norms of American politics and rhetoric. , 189–205 pages.
[62] Lester M Salamon and Michael S Lund. 1984. *The Reagan presidency and the governing of America.* The Urban Insitute. 2–3 pages.
[63] SeproDev and other contributors. 2024. yt-dlp/yt-dlp: A feature-rich command-line audio/video downloader. https://github.com/yt-dlp/yt-dlp
[64] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter* 19, 1 (2017).
[65] Kate Starbird. 2017. Examining the alternative media ecosystem through the production of alternative narratives of mass shooting events on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 11. 230–239.
[66] Kate Starbird, Ahmer Arif, and Tom Wilson. 2019. Disinformation as collaborative work: Surfacing the participatory nature of strategic information operations. *Proceedings of the ACM on human-computer interaction* 3, CSCW (2019), 1–26.
[67] Kate Starbird, Renée DiResta, and Matt DeButts. 2023. Influence and improvisation: Participatory disinformation during the 2020 US election. *Social Media+ Society* 9, 2 (2023), 20563051231177943.
[68] Sharifa Sultana and Susan R. Fussell. 2021. Dissemination, Situated Fact-checking, and Social Effects of Misinformation among Rural Bangladeshi Villagers During the COVID-19 Pandemic. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 436 (Oct. 2021), 34 pages. https://doi.org/10.1145/3479580
[69] Rama Adithya Varanasi, Joyojeet Pal, and Aditya Vashistha. 2022. Accost, accede, or amplify: attitudes towards COVID-19 misinformation on WhatsApp in India. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–17.
[70] Soroush Vosoughi, Deb Roy, and Sinan Aral. 2018. The spread of true and false news online. *Science* 359, 6380 (2018).
[71] Ching-Hsing Wang and Chia-hung Tsai. 2023. Social media as mobilizer and demobilizer: Evidence from Taiwan. *SAGE Open* 13, 1 (2023), 21582440231156806.

[72] Song-In Wang. 2006. The Influence of the Internet on Political Attitudes and Campaign Participation in the 2004 Legislative Election. *Taiwan Democracy Quarterly* 3, 4 (2006), 72–102.

[73] Yuping Wang, Chen Ling, and Gianluca Stringhini. 2023. Understanding the use of images to spread COVID-19 misinformation on Twitter. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–32.

[74] Yuping Wang, Fatemeh Tahmasbi, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, David Magerman, Savvas Zannettou, and Gianluca Stringhini. 2021. Understanding the use of fauxtography on social media. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15. 776–786.

[75] Tom Wilson, Kaitlyn Zhou, and Kate Starbird. 2018. Assembling Strategic Narratives: Information Operations as Collaborative Work within an Online Community. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 183 (Nov. 2018), 26 pages. https://doi.org/10.1145/3274452

[76] Liang Wu, Fred Morstatter, Kathleen M Carley, and Huan Liu. 2019. Misinformation in social media: definition, manipulation, and detection. *ACM SIGKDD Explorations Newsletter* 21, 2 (2019), 80–90.

[77] Yu-Shan Wu. 2024. Pendulum Stopped: Rise of Ko Wen-je and Its Impact on Taiwan Politics. *American Journal of Chinese Studies* 31, 1 (2024).

[78] Ho-Don Yan and Michael SH Heng. 2016. Political entrepreneurship and the Ko-P phenomenon in Taiwan's democratization. *Asian Journal of Political Science* 24, 3 (2016), 285–304.

[79] Savvas Zannettou. 2021. " I Won the Election!": An Empirical Analysis of Soft Moderation Interventions on Twitter. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 15. 865–876.

[80] Savvas Zannettou, Tristan Caulfield, Jeremy Blackburn, Emiliano De Cristofaro, Michael Sirivianos, Gianluca Stringhini, and Guillermo Suarez-Tangil. 2018. On the Origins of Memes by Means of Fringe Web Communities. In *Proceedings of the Internet Measurement Conference 2018*. 188–202.

[81] Xinyi Zhou and Reza Zafarani. 2020. A survey of fake news: Fundamental theories, detection methods, and opportunities. *ACM Computing Surveys (CSUR)* 53, 5 (2020), 1–40.

A Additional Details of Method

A.1 Collection Period Selection

The Taiwan FactCheck Center (TFC) was established in April 2018 through a collaboration between the Taiwan Media Watch Foundation and the Association for Quality Journalism. It began releasing fact-checking reports on July 31, 2018, aiming to combat digital misinformation. It is widely considered as the most official fact-checking agency for debunking digital misinformation in Taiwan. The voter fraud scandal during the Taiwan election of 2024 has been fact-checked and debunked as misinformation by TFC.

According to TFC, the first fact-checking report related to voter fraud misinformation was published on *January 13th, 2024*, and the last one on *January 31st, 2024*. Therefore, we decided to collect data within the time range from January 12th to February 1st. This time frame represents **a one-day buffer** before and after the justified period of January 13th to January 31st.

We have already explained the reasons for why we choose January 13th as the starting point in the main text, and now we will only provide more detailed reasons for why we choose January 31st as the ending point:

(1) The final fact-checking report related to voter fraud misinformation was published by TFC at 10:24 A.M. on *January 31st, 2024*. The report, titled "[Misunderstanding] The video: If this is not vote fraud, what is vote fraud?",[11] indicated that the voter fraud scandal continued to circulate before January 31st. However, since this is the last relevant fact-checking report, it suggests that the general trend of voter fraud misinformation dissemination had largely subsided by that time.
(2) The ending point is further supported by Taiwan's media regulation framework. Under the authority of the "National Communications Commission" (NCC), the dissemination of misinformation is officially curtailed once it is flagged by TFC. This regulation is highly effective for traditional media outlets. Once misinformation is identified by TFC, these outlets are prohibited from further spreading the content, or they may face legal action from the NCC. Although grassroots media are less directly controlled by the NCC, they still face scrutiny regarding media credibility and are monitored by digital platforms. Therefore, given that the last report was issued on the morning of January 31st, it is highly probable that voter fraud misinformation was minimal after this date.
(3) To further verify the ending point, we conducted a manual check, reviewing all videos released by the listed accounts between February 1st and February 3rd. We found only one video related to voter fraud misinformation during this period. This video was posted by a grassroots media account with a political preference toward the TPP, which strongly believed that TPP's loss in the 2024 Taiwan election was due to voter fraud. We noted that this account continued to publish videos related to voter fraud throughout 2024 by showing videotapes recorded from different polling places (small areas with walls or curtains on three sides where voters can cast their votes in private). Despite this outlier, we decided to set the justified ending point on January 31st, based on the overall trend of voter fraud misinformation dissemination and to maintain a balanced comparison with other channels.

We finally offer a one-day buffer for the justified period, resulting in the final period for data collection is ranging from January 12th to February 1st.

[11] https://tfc-taiwan.org.tw/fact-check-reports/migration-10259/

A.2 Chanel Selection

We include an link containing the complete list of media account names here:
https://docs.google.com/document/d/1vLEKTULHiJF8ZDJnZlkTHEEo5cedAe5qhxE7tG60HPU/edit?usp=sharing

A.2.1 Traditional Media. The selection of traditional media is based on official lists. We check the list of TV channels and radios from the "National Communication Committee" (NCC) of Taiwan [12] and the newspaper list from the "Taiwan Periodical Literature." [13] The results are almost complete and probing.

In order to determine the political leanings of traditional media outlets in Taiwan, we draw on existing reports and publicly available online information. Particular attention was given to the development histories of these media organizations, the political affiliations of their funders, and the partisan orientations of the parent companies. Based on these criteria, the selected media outlets were categorized into three groups: pan-blue, pan-green, and others.

Media outlets with clear affiliations to the pan-blue camp—typically associated with the **KMT** or pro-blue individuals—include TVBS, CTV, CTi, BCC, "China Times," and United Daily News. Conversely, outlets linked to the pan-green camp—generally associated with the **DPP** or pro-green individuals—include FTV, SET, and Liberty Times. This categorization is primarily informed by data from Wikiversity,[14] cross-verified with additional sources. Furthermore, ERA TV was also classified as pan-green due to its financial backing by Lien Tai-sheng, a figure widely recognized for his ties to the TPP camp.[15]

In contrast, media outlets such as the CTS and PTS, which are publicly funded and maintain a degree of institutional independence, were classified as **"Other."** Similarly, TTV and USTV were also placed in the "Other" category, as there is no substantive evidence suggesting control by financial groups with explicit political leanings. Besides, for EBC and Global News, there is considerable divergence in public opinion regarding their political orientation.[16] In light of this ambiguity, and after consulting additional sources (e.g., https://taiwan.huanqiu.com/article/9CaKrnK1yfH), this study conservatively chose not to assign these outlets to a particular political category, instead including them under "Other."

A.2.2 Grassroots Media. Then, we provide more details on the selection of grassroots media for each political camp.

For Democratic Progressive Party. We first use the Google search engine to query the grassroots media accounts of DPP, which is commonly considered to have the most mature media community among all parties in Taiwan. We then get a number of relevant articles and news reports from credible websites, which contain several listed grassroots user accounts of DPP. Three articles are highly informative and they are listed below.

(1) "The list is revealed! The "big names" of the Green Wing (DPP) are all in the list, and I can't be the only one to see them"[17] from *Yahoo!*.

[12] https://www.ncc.gov.tw/chinese/index.aspx
[13] https://tpl.ncl.edu.tw/NclService/JournalGuide?q%5B0%5D.i=%E6%8A%A5&q%5B0%5D.f=*&q%5B1%5D.o=0&q%5B1%5D.i=&q%5B1%5D.f=*&lang=&pys=&pms=&pye=&pme=
[14] https://zh.wikiversity.org/zh-sg
[15] https://www.chinatimes.com/realtimenews/20200503001477-260407)
[16] http://big5.taiwan.cn/taiwan/tw_SocialNews/201011/t20101113_1601009.htm
[17] https://tw.news.yahoo.com/%E5%90%8D%E5%96%AE%E6%9B%9D%E5%85%89-%E7%B6%A0%E7%87%9F%E5%81%B4%E7%BF%BC%E7%BE%A4%E7%B5%84-%E5%A4%A7%E5%92%96-%E9%83%BD%E5%9C%A8%E8%A3%A1%E9%9D%A2-%E5%BE%90%E5%B7%A7%E8%8A%AF-034600152.html

(2) "Help the Democratic Progressive Party to organize the list of flank cleaning after losing the election"[18] from *MOBILE01*.
(3) "Green Flanker Slacker Bag"[19] from *HackMd*.

Noticeably, listed accounts in obtained articles may belong to other social media platforms, e.g., Facebook, X, etc. Moreover, some accounts may be inactive now since some of our obtained articles are from several years ago. Therefore, we then check each mentioned account in the list to see whether they have an active account on YouTube and whether their account is active during the Taiwan 2024 election period. If both are so, then we include this account for our grassroots account list. It turns out that grassroots accounts are highly spared and most of the listed accounts in the obtained articles belong to Facebook, the platform that is most frequently used for propaganda in Taiwan. The qualified accounts we can obtain are restricted. For example, in the report (1) mentioned above (the *Yahoo!* report), only three of over 40 listed accounts are qualified, though this report is the most informative one among all the other reports we obtained for the selection of grassroots accounts for DPP. We finally confirm 10 qualified grassroots accounts in YouTube for DPP. This brings the DPP's total to 27 media accounts. For balanced comparison and subsequent analysis, we set 27 as the target number of aggregated media accounts for both the KMT and TPP.

For Kuomintang. Repeating the process, the most informative articles for confirming grassroots YouTube accounts of KMT is listed below.

(1) "The reasons for the marginalization of the Blue Camp's water army are revealed: they are loyal to Han Kuo-yu and ignore the KMT?"[20] from *People News*.
(2) "The Distance Between Political Micro-Netizens and Fake News"[21] from *Yahoo!*.
(3) "Blue and White don't get along? Blue and white flanker fans start a war on their own side, and the green flanker turns into a "knife fan" to join in the fun"[22] from *newtalk*.
(4) "[The Blue Camp's Battle of the Bulge] Criticizing the U.S. Pig is Forced to Be Re-recorded TV Celebrities Turn to the Internet and Refuse to be Silenced"[23] from *upmedia*.

Similar to DPP, grassroots accounts of KMT on YouTube are sparse and the qualified accounts are restricted. For example, in the article (1) mentioned above (the *People News* report), only four of all listed accounts are qualified, though this report is the most informative one among all the other reports of KMT. Moreover, the power of the grassroots accounts of KMT is not as strong as the ones belonging to DPP. They are mostly constituted by the fans of Han Kuo-yu (see article (1)), a famous and influential politician in Taiwan. We eventually get 11 grassroots media accounts on YouTube. This results in aggregated 27 accounts for KMT, which is a numeric balance with DPP channels.

For Taiwan People's Party. Known that since the Taiwan People's Party (TPP) was established long after the other two parties, the TPP does not own any news outlets by itself. [24] Instead, they operate a quantity of personal grassroots media accounts. Their accounts are more active and highly connected to each other, compared to those of KMT and DPP. The most informative articles for confirming grassroots YouTube accounts of DPP are provided below.

[18]https://www.mobile01.com/topicdetail.php?f=638&t=6701984
[19]https://hackmd.io/@kpsupkeepgoing2024/BkWjWe0uA?utm_source=preview-mode&utm_medium=rec
[20]https://www-peoplenews-tw.translate.goog/articles/bdaef4b06b?_x_tr_sl=zh-TW&_x_tr_tl=zh-CN&_x_tr_hl=zh-CN&_x_tr_pto=sc
[21]https://tw.news.yahoo.com/%E6%94%BF%E6%B2%BB%E5%BE%AE%E7%B6%B2%E7%B4%85%E8%88%87%E5%81%87%E8%A8%8A%E6%81%AF%E7%9A%84%E8%B7%9D%E9%9B%A2-230058391.html
[22]https://newtalk.tw/news/view/2023-10-16/892647
[23]https://www.upmedia.mg/news_info.php?Type=1&SerialNo=97810
[24] https://www.tpp.org.tw/newsdetail/3138

(1) "[Ko P's Self-Help Breakout 2] Five Million YouTubers Support Ko Wen-je at Their Own Expense, 'Influencer' Helps Fight Air War Against Blue-Green Clique"[25] from *Yahoo!*.
(2) "Ko made his first public appearance after his defeat in the election"[26] from *Rti*.
(3) "For every vote Hau gets over Ko, he gets a million dollars! Ko's fans' "this action" and Wu Tzu-chia's reaction is so shocking"[27] from *Yahoo!*.

We finally list 27 grassroots media accounts of TPP on YouTube, keeping an aggregated numeric balance with KMT and DPP.

A.3 Videos Annotation

In this part, our goal is to identify those discussing voter fraud. This process is carried out in two main steps: 1) We use LLaMA3.1:8B to identify all videos related to ballots. 2) We then manually inspect these ballot-related videos and label them as either voter-fraud-relevant or not. Except for the information provided in the main paper, we add more additional details here.

A.3.1 For LLM Annotation (Phase I: Labeling ballot-related videos). We first target ballot-related content, as videos discussing voter fraud are a subset of these, and LLMs more reliably identify the broader category.

The classifier was designed through extensive prompt engineering and human validation to ensure its effectiveness. It performs multimodal analysis by integrating both textual and audio information and operates in a two-step mechanism: *1) Title-Based Classification*: Each video is first evaluated based on its title. We derived the title by scraping the metadata of all collected videos using YouTube Data API V3. If the title alone is sufficient to determine the video as ballot-related, it is flagged. *2) Audio-Based Classification*: If the title is inconclusive, we convert the video's audio to text using the VOSK speech recognition model [49]. The transcribed text is then analyzed by the LLM to determine whether it pertains to ballot-related content.

The prompt template (translated from traditional Chinese) is developed as follows. We begin with this original version:

> "Now, assume you are a news analyst, and please determine whether the following content is relevant to casting votes, voting scrutiny, and voter fraud. If so, please answer 'yes' and 'no' otherwise. \n The content is shown as: \n {text_content}"

However, we found that this prompt often led the LLM to avoid giving definitive answers, possibly due to alignment-related hesitations. To address refusal issues, e.g., "Sorry, I can't answer this question...", we revised the prompt as follows:

> "Please determine whether the following content is relevant to casting votes, voting scrutiny, and voter fraud based on its words and content. If so, you must answer 'yes' and 'no' otherwise. \n Please answer as precisely as possible (especially when your answer is yes), and give your reasons briefly. The content is shown as: \n {text_content}"

To ensure the performance of the LLM, we conducted a manual validation before applying it at scale. We randomly selected 100 videos from the dataset and applied the prompt to have the LLM label them. Meanwhile, two authors independently annotated all 100 videos (i.e., each

[25] https://tw.news.yahoo.com/%E6%9F%AFp%E8%87%AA%E6%95%91%E7%AA%81%E5%9C%8D2-500%E8%90%ACyoutuber%E8%87%AA%E8%B2%BB%E6%8C%BA%E6%9F%AF%E6%96%87%E5%93%B2-%E7%B6%B2%E7%B4%85%E8%81%AF%E6%92%AD-%E5%8A%A9%E6%94%BB%E7%A9%BA%E6%88%B0%E6%8A%97%E8%97%8D%E7%B6%A0%E5%A4%BE%E6%AE%BA-215858057.html
[26] https://www.rti.org.tw/news/player/id/2193119
[27] https://tw.news.yahoo.com/%E4%BE%AF%E5%A4%9A%E6%9F%AF1%E7%A5%A8%E5%B0%B1%E9%80%81100%E8%90%AC-%E6%9F%AF%E7%B2%89-%E9%80%99%E5%8B%95%E4%BD%9C-%E5%90%B3%E5%AD%90%E5%98%89%E5%8F%8D%E6%87%89%E8%B7%8C%E7%A0%B4%E7%9C%BC%E9%8F%A1-005214059.html

video received two human annotations). We then calculated Fleiss's Kappa to evaluate inter-annotator agreement between two human annotators. The result was 0.976, indicating substantial agreement [36]. Following this, the authors engaged in discussion to reconcile any discrepancies and reach consensus, resulting in a finalized golden dataset for validation.

We then applied the prompt to the same 100 videos using the LLM and compared its outputs to the golden dataset. We computed standard classification metrics to assess the LLM's labeling performance, including Accuracy, Macro-F1, Precision, and Recall. The results in Table 6 validates the LLM's labeling performance [37], supporting its use for large-scale annotation.

Table 6. Validation results comparing LLM annotations to the golden dataset

Accuracy	**Macro-F1**	**Precision**	**Recall**
0.910	0.899	0.884	0.926

A.3.2 For Manual Annotation (Phase II: Labeling voter fraud videos). To validate the annotation consistency, we randomly selected 100 videos from the 2,421 ballot-related videos. All three authors independently annotated these videos based on the codebook mentioned in the paper (i.e., each video received three annotations). We used Fleiss's Kappa to assess inter-coder reliability, and the resulting average score was 0.923, indicating a high level of agreement among the three annotators. Details of the intercoder agreements are shown as the Table 7 below.

Table 7. Inter-annotator agreement among three human annotators using Fleiss's Kappa

Annotator	Human NO.2	Human NO.3
Human NO.1	100.00%	88.37%
Human NO.2		88.37%

To help readers better understand how voter-fraud videos are annotated, we present one representative example for each rule described in the paper. Figures 10 to 14 illustrate a distinct example corresponding to each annotation rule. (For English translations of Figures 10, 12, 13, please refer to the Case Study section of the paper.)

(1) If any on-screen text in the video is related to voter fraud, the video is labeled as voter-fraud relevant.
(2) If any spoken audio in the video is related to voter fraud, the video is labeled as voter-fraud relevant.
(3) If one or more individuals in the video are shown engaging in voter fraud or are highly suspicious of doing so, the video is labeled as voter-fraud relevant.
(4) If the video mentions or implies that an entity (a person or organization) is attempting to commit voter fraud, it is labeled as voter-fraud relevant.
(5) If the video includes only viewer comments discussing voter fraud, without any connection to the video content itself, it is NOT considered a voter-fraud video.

Fig. 10. On-screen text related to voter fraud:*Main headline (yellow and white text on black background) "Ballot-counting chaos, many irregularities in vote counting. Blank ballots turned into counted votes, people caught red-handed!" On-screen text related to voter fraud is highlighted. The image above is a screenshot extracted from a video in our compiled YouTube dataset.*

Fig. 11. Spoken audio referencing voter fraud: *Although there is no on-screen text or visual representation, its spoken audio discusses voter fraud at 8:21. The image above is a screenshot extracted from a video in our compiled YouTube dataset.*

Fig. 12. Suspicious scene of engaging in voter fraud: *A volunteer is seemingly removing the bottom panel of a ballot box, revealing hidden ballots—reveals a violation of proper election procedures. The image above is a screenshot extracted from a video in our compiled YouTube dataset.*

Fig. 13. Implication that an entity is attempting to commit voter fraud: *Bottom text: "Legislator Cheng Cheng-ling: Why must it be changed to an acrylic box?", implying that the change could be associated with an attempt to commit voter fraud. The image above is a screenshot extracted from a video in our compiled YouTube dataset.*

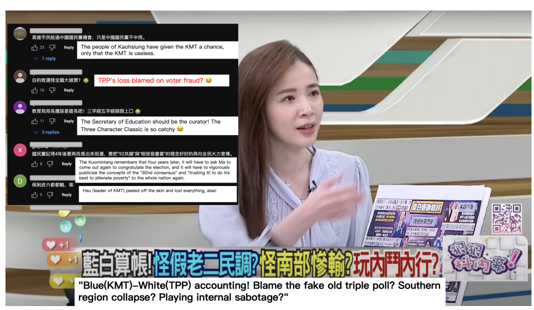

Fig. 14. Videos that only include viewer comments discussing voter fraud are **NOT** annotated as voter-fraud video (a screenshot taken by the author is displayed on the left side, showing comments that are not part of the original video content): *This video discusses the triple poll and makes no mention of voter fraud, although one user comment refers to it. The image above is a screenshot extracted from a video in our compiled YouTube dataset.*

A.3.3 Human Annotation Code-books.

Codebook for Ballot-Relevance Annotation (Human Annotators). To determine whether a video is relevant to voting processes, ballot scrutiny, or voter fraud, human annotators should apply the following rule. Each annotator is instructed to assess the video based on its **title and transcript**, which are provided in textual format. Relevance is defined as any substantive mention, description, or implication of the following:

- **Casting votes:** References to the act of voting, including ballot submission, polling procedures, or voter participation.
- **Voting scrutiny:** Mentions of vote counting, ballot handling, oversight mechanisms, election monitoring, or procedural integrity.
- **Voter fraud:** Allegations, implications, or evidence of fraudulent behavior related to elections, such as ballot tampering, illegal voting, or conspiracies involving vote manipulation.

Annotators must answer the binary question: **Is this video relevant to casting votes, voting scrutiny, or voter fraud?** Valid responses are yes or no.

Annotation Instructions:

(1) Carefully read the full title and transcript of the video.

(2) Determine whether the content substantively discusses or implies any of the three ballot-relevant categories listed above.
(3) If the content is relevant, record `yes`; otherwise, record `no`.
(4) If answering `yes`, briefly write a justification describing the specific content that triggered the decision. This justification should be concise and refer to relevant phrases or themes.
(5) Do not rely solely on superficial mentions (e.g., user comments or unrelated metaphors); focus on the main subject matter conveyed by the title and transcript.

This codebook (Table 8) ensures consistent interpretation of ballot relevance and supports high inter-annotator agreement through structured, normative guidance.

Table 8. Unified Codebook for Ballot-Relevance Annotation: Criteria and Instructions

	Ballot-Relevant Criteria
Casting Votes	References to the act of voting, including ballot submission, polling procedures, or voter participation.
Voting Scrutiny	Mentions of vote counting, ballot handling, oversight mechanisms, election monitoring, or procedural integrity.
Voter Fraud	Allegations, implications, or evidence of fraudulent behavior related to elections, such as ballot tampering, illegal voting, or conspiracies involving vote manipulation.
	Annotation Instructions
Step 1	Carefully read the full title and transcript of the video.
Step 2	Determine whether the content substantively discusses or implies any of the three ballot-relevant categories listed above.
Step 3	If the content is relevant, record `yes`; otherwise, record `no`.
Step 4	If answering `yes`, briefly write a justification describing the specific content that triggered the decision. This justification should be concise and refer to relevant phrases or themes.
Step 5	Do not rely solely on superficial mentions (e.g., user comments or unrelated metaphors); focus on the main subject matter conveyed by the title and transcript.

Codebook for Voter Fraud Relevance Annotation. To determine whether a video is relevant to voter fraud, the authors collaboratively developed a codebook based on an initial review of the annotated dataset. The following normative rules were applied during manual annotation:

(1) **On-Screen Text:** A video is labeled as "voter-fraud relevant" if any on-screen text explicitly or implicitly references voter fraud.
(2) **Spoken Audio:** A video is labeled as "voter-fraud relevant" if any portion of the spoken audio content (e.g., narration, interview, speech) includes claims or discussions related to voter fraud.

(3) **Visual Depiction of Acts:** A video is labeled as "voter-fraud relevant" if one or more individuals are shown engaging in acts that can reasonably be interpreted as voter fraud or that are highly suspicious thereof.
(4) **Allegations or Implications:** A video is labeled as "voter-fraud relevant" if it contains explicit allegations or clear implications that a specific individual or organization is attempting to commit voter fraud.
(5) **Comment-Only Content:** A video is labeled as "not voter-fraud relevant" if the only mention of voter fraud appears in the user comments section, with no indication of such content in the video's visuals, audio, or narration.

All annotations were conducted manually by the three authors, following this codebook (Table 9) to ensure consistent application of criteria across the dataset.

Table 9. Unified Codebook for Voter-Fraud Relevance Annotation: Criteria and Instructions

Criteria for Voter-Fraud Relevance	
Criterion 1	If any on-screen text in the video is related to voter fraud, the video is labeled as **voter-fraud relevant**.
Criterion 2	If any spoken audio in the video is related to voter fraud, the video is labeled as **voter-fraud relevant**.
Criterion 3	If one or more individuals in the video are shown engaging in voter fraud or are highly suspicious of doing so, the video is labeled as **voter-fraud relevant**.
Criterion 4	If the video mentions or implies that an entity (a person or organization) is attempting to commit voter fraud, it is labeled as **voter-fraud relevant**.
Criterion 5	If the video includes only viewer comments discussing voter fraud, without any connection to the video content itself, it is **not** considered a voter-fraud video.
Annotation Instructions	
Step 1	Watch the full video carefully, including all visual and audio components.
Step 2	Evaluate whether any part of the content—spoken audio, on-screen text, or visual scenes—satisfies any of the relevance criteria listed below.
Step 3	If at least one criterion is satisfied, label the video as `voter-fraud relevant`. If none are satisfied, label the video as `not relevant`.
Step 4	Record a brief justification for each positive annotation to facilitate inter-coder reliability checks.

Codebook for Narrative Stance Annotation. To help readers better understand how the stance of each voter-fraud video is coded for the analysis in Section 5, we include and explain our codebook here. Table 10 defines the three categories and typical cues for each narrative stance.

Especially, coders observe the following indicators to determine the stance (S–Supportive; R–Refute; N-Neutral).

Table 10. Unified Codebook for Stance Annotation: Definition and Instructions

Criteria for Video Stance toward Voter-Fraud Narrative	
Supportive	The video endorses, amplifies, or treats as plausible the claim that voter fraud occurred.
Refuting	The video rejects, debunks, or ridicules the voter-fraud narrative.
Neutral	The video mentions fraud but neither clearly supports nor refutes it, or mixes supportive and refuting cues without a dominant thrust.
Annotation Instructions	
Step 1	Watch the full video carefully, including all audio and visual elements.
Step 2	Examine the transcript, spoken audio, on-screen text, and visual scenes for evaluative cues related to voter fraud.
Step 3	Assign one of the three stance labels (Supportive, Refuted, Neutral) based on the dominant evidence using the definitions above.
Step 4	Record a concise justification explaining the basis for the label. Flag ambiguous cases for group adjudication.

S–1 On-screen text claims or implies voter fraud.
S–2 Spoken audio affirms or suspects that voter fraud occurrs.
S–3 Visual sequence is presented as proof of voter fraud.
S–4 Presenter asks leading questions assuming the existence of voter fraud.
R–1 Text or narrative explicitly denies or implies its denial of voter fraud.
R–2 Official fact-check or debunk is cited.
R–3 Step-by-step explanation shows fraud is impossible / nonexistent.
R–4 Rebuttal to voter fraud supporters.
N–1 Balanced coverage or fleeting mention only.

A.4 Kolmogorov–Smirnov (KS) Test

The Kolmogorov-Smirnov (KS) test is a nonparametric test used to compare two empirical cumulative distribution functions (ECDFs). It is commonly used to determine whether two samples come from the same distribution [41, 70].

Let $F_n(x)$ and $G_m(x)$ denote the empirical distribution functions of two samples of sizes n and m, respectively.

The test statistic is defined as:

$$D_{n,m} = \sup_x |F_n(x) - G_m(x)|$$

This statistic measures the maximum vertical distance between the two ECDFs.

Null hypothesis:

$$H_0 : F(x) = G(x) \quad \text{for all } x$$

Alternative hypothesis:

$$H_1 : F(x) \neq G(x) \quad \text{for some } x$$

The significance level (p-value) is calculated as:

$$p = Q_{KS}\left(\sqrt{\frac{nm}{n+m}} \cdot D_{n,m}\right)$$

where the Kolmogorov distribution is:

$$Q_{KS}(\lambda) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2\lambda^2}$$

In practice, this infinite sum converges quickly. A small p-value (e.g., $p < 0.05$) indicates that the null hypothesis should be rejected in favor of the alternative.

Received October 2024; revised April 2025; accepted August 2025