This research is an inquiry into how online activism and images related to immigration justice can be manipulated and used to harass individuals or groups, spread false information, and incite hate. Over the span of a week, we collected and analyzed close to one million tweets related to the ongoing, contentious immigration debate in the United States and identified trends and case studies of image appropriation and re-contextualization patterns on social media. We enriched this analysis through interviews with WITNESS’ immigrant rights partners, which further demonstrate how the fight for social justice has both benefited from social media, as well as been disrupted by digital disinformation.

This study expands the current understanding on how rumors and alternative narratives propagate online; specifically, it contributes new knowledge with regards to the negative impact of digital content manipulation on the narratives surrounding immigrants and refugees within the United States. As more sophisticated, personalized photo and video manipulation techniques emerge, this research addresses the necessity to understand disinformation surrounding visual content and the ramifications on social justice movements.
INTRODUCTION

Amid the evolving mire of online information practices, visual content has surfaced as an exceptionally useful and reliable tool in protecting and defending human rights. Activists and the organizations supporting them, including WITNESS, have long demonstrated the value of visual media in the fight for justice. As accessibility to mobile phones only increases and social media plays a pervasive role in all of our lives, we have seen how media can serve as authentic documentation of violations and a catalyst for social change. These trends have enabled activists to circumvent hierarchical power structures and share the stories of marginalized groups, creating a new era of online human rights activism driven by visual content.

Meanwhile, visual mis- and disinformation, both in context and content, have become exceedingly prevalent online, including the threats of progressively sophisticated synthetic media, “deep fakes”, which pose potential social, economic, and political threat (World Economic Forum, 2017). If the authenticity of visual content is called into question, communities that rely on accurate documentation of human rights injustices -- such as the ones WITNESS engages with on a daily basis -- will find their lives and well-being at risk. While the participatory nature of social media has been proven to help individuals foster political will and organize with a collective purpose, bad actors have increasingly exploited the lower barriers of entry online to propagate new narratives and disinformation surrounding activist-generated visual content. Despite the increasing availability of effective and accessible means of visual content manipulation, journalistic and platform responses to ‘fake news’ issues have predominantly focused on text-based mis/dis/mal-information and text-based responses.

This project studies visual content created and shared by activists online in response to recent United States’ immigration policies and Immigration and Customs Enforcement (ICE) activities. This study intends to map the evolving contextual narratives surrounding visual content online as users, both those in support of and opposed to current immigration policies, negotiate the interpretations and contents of images and videos. The Twitter data poses limitations on the conclusions that can be drawn concerning the qualitative experiences of activists and immigrants using social media, so interviews were conducted to supplement the quantitative analysis. These interviews serve as anecdotal data to enrich the stories told by the tweet data alone.

This research intends to answer the following questions:

1. What common trends in conversational patterns on social media can we identify that develop around images and videos in a social movement context?

2. What are the more granular patterns in how individual users interact with these images and videos as they deliberate on the content and context of visual content, otherwise known as framing?1

3. How does misinformation develop and grow from the aforementioned framing negotiations of visual content?

---

1. In the study of social movements, framing is a process that enables groups of people to understand, remember, evaluate, and act upon a problem.
VISUAL CONTENT AND ONLINE POLITICAL COMMUNICATION

In an increasingly image-based society, online visual content has become a powerful tool for social rights advocacy. The communities that WITNESS works directly with not only depend on technology to expose injustices, but to share that content online to incite change as a part of larger social movements, including LGBTQ rights, fighting police brutality, and protesting immigration abuses. Visuals are fundamental in shaping public opinion (Grabe, 2009), and an increasing amount of the U.S. population, especially young people, rely on text-constrained, image-heavy social media to consume news.

Visual news consumption is not a new phenomenon, but the novel form of content being created and disseminated horizontally online is—eyewitness videos are being shared on social media platforms and supplementing breaking news in situations of danger and unrest where professional journalists don’t have the resources or access to work. In 2014, roughly one-in-ten social network users reported having posted news videos they took themselves (Mitchell & Page, 2014), and this can only be assumed to have increased as users of social media sites and the presence of video content in the U.S. have risen at least 10% in the last four years (Smith & Anderson, 2018). ‘Citizen journalist’ became a widespread term in December 2004, when first-person accounts of the South Asian tsunami by tourists on the scene were recognized as an extraordinary contribution to the coverage by mainstream journalists (Andén-Papadopoulos & Pantti, 2011). And in the U.S., the Boston Marathon bombings in 2013 were recorded by bystanders at the events, from initial blasts and to the final police chase. When Ukrainian protesters were massacred by riot police in Kiev’s Independence Square in 2014, we witnessed the threat and political influence that images hold; news organizations re-contextualized eyewitness content from the riots, creating new images that were contrary to the original intent of exposing police violence (Mitchell et al., 2014).

WITNESS equips activists and marginalized communities with technology and training to document injustices, and depends on the power of the image to spread awareness of human rights violations. As we still grapple with emerging uses of information technology, the process for how news organizations and citizens share content is still evolving, with questions about how to provide proper attribution, how to verify the accuracy of videos, and even how to archive them for the future (Mitchell et al., 2014).

As social media platforms negotiate best practices for verification and provenance of user-generated videos, bad actors have taken advantage of the gaps in cohesive industry standards. Media manipulation, both in fabricated content and context, have been identified as two of seven types of mis- and disinformation (Wardle, 2017). In her report, Wardle describes ‘fake news’ not just as a news issue, but a fracture in the entire information ecosystem and identified motivations for disseminating manipulated visual content ranging from parody to propaganda and political influence. Researchers Marwick and Lewis in their report “Media Manipulation and Disinformation” found that images, specifically memes, designed to spread between individuals online, play into propaganda being shared through interpersonal ties rather than a top-down information flow. The threat of synthetic media has incited countless gatherings and convenings, including those hosted by WITNESS, of practitioners and academics in machine learning, journalism, activism, and other fields to discuss the risks and potential solutions for impending synthetic media attacks (Bakdash, 2018; Gregory, 2018). As it pertains to citizen journalists, the most prominent of concerns that surfaced from these convenings are gender-based attacks on credibility of human rights defenders and journalists, attacks on social movement narratives and credibility, and targeting of dissidents in authoritarian contexts with ubiquitous, unconstrained surveillance (Gregory, 2018).

BACKGROUND

While virtually anything can be a meme, in modern internet parlance, a meme is a visual trope that proliferates across internet spaces as it is replicated and altered by anonymous users. (Marwick & Lewis, 2017)
There is a rift in motivations of image creation and consumption. On one end, a new liberating platform for horizontal story-telling that tempers conventional narratives so that marginalized voices can rise to the surface; the other a new freedom to circumvent traditional information gatekeeping mechanisms and construct narratives that are malicious and threatening to the well-being of diverse communities and democratic infrastructure. As misinformation and conspiracy theories pander to anxieties of chaos in political and social order (Paul & Matthews, 2016), the emotional nature of the image makes it a particularly potent tool in stoking fears of different social groups. Past research points to an increasingly need for a deep inquiry into the online discursive structure around images targeting marginalized groups as they circulate on a social media platform. This study builds upon precedents and concepts of the political image while utilizing novel social network analysis techniques to examine these theories in a complex environment: the social media platform.

**NETWORKED ACTIVISM**

A well-documented process essential to social movement development, known as framing, has carried over from physical to online public spheres through social media channels (Meraz et al., 2013). A popular definition of framing comes from Entman, who specified that “to frame is to select some aspects of a perceived reality and make them more salient in a communicating text” (Entman, 1993, p. 52). Framing has been repeatedly applied to social network analyses3 (SNA) of online social movement organizing, using affordances specific to social media platforms—such as mentions, retweets, and hashtags— as a methodological technique to make sense of traditional political organizational processes like information gatekeeping online (Stewart et al., 2017; Meraz & Papacharissi, 2013).

The Arab Spring uprisings in 2011 were heavily studied political protests by social network researchers as they were the first to be largely facilitated by social media. Live-tweeting and live video feeds played an especially important role for activists, enabling them to share their experiences with a larger global audience, regardless of traditional media coverage. This produced eyewitness visual content that was disseminated online through retweeting and narrative frames for the movement were essentially crowd-sourced into prominence on a global stage (Starbird & Palen, 2012). Meraz & Papacharissi used previous seminal work on social movement framing as a lens to explain evolving patterns of communication on Twitter surrounding the Egyptian uprisings. Findings underscored the significant role of ‘ordinary’ users who “both rose to prominence and elevated others to elite status through networked gatekeeping actions” (Meraz et al., 2016).

Similarly, Kow et al. (2016) studied a more technologically complex situation during the Umbrella Movement, a series of student protests, that took place in Hong Kong in 2014. Across platforms, Kow, et al. identify four primary uses of social media for social movements: information seeking and dissemination, frame articulation, expression of solidarity, development of counter-narratives. This final purpose, development of counter-narratives, involves the expression of alternative perspectives, often in conflict with official reports (Al-Ani et al., 2012; Monroy-Hernández et al., 2013). Counter-narratives “rebult, undermine, or neutralize” a groups framing (Benford & Snow, 2000), a notion that is especially important for our analysis as we examine the network structure and motivations behind manipulated media. Networks around alternative media sites, which connect domains with the same users, have also been examined to triangulate the actors creating and propagating alternative narratives (Starbird, 2017) and have been used to track the patterns of online rumor propagation as a step towards building misinformation detection tools (Maddock et al., 2015).

This study extends previous research on online activism and rumor propagation by focusing on visual elements that are used to represent narratives and their subsequent counter-narratives. The visual content shared by social rights activists inextricably ties them to the current political social contexts and movements that they represent, making visual content both useful for framing and simultaneously constantly subject to reframing and recontextualization.

---

3. Social network analysis has roots in sociology and has more recently come into popularity due to advances in computational power and availability of social network big data. At its core, SNA aims to study the patterns in relationships that connect social actors through the use of network graphs and graph theory.
DATA COLLECTION

TWEETS FROM THE TWITTER API
Data was collected from the Twitter streaming application programming interface (API) from June 28th to July 5th, 2018. This time period was chosen due to the “End Family Separation” marches planned to occur across the U.S. on Saturday, June 30th. The zero tolerance and family separation immigration policies\(^4\) were already being largely debated after images and audio of detained young children from inside one of the children’s shelters had circulated earlier that day. Significantly, there was data loss caused by internet disconnection early Sunday, July 1st, the day following the marches, until the next day, July 2nd.

Three neutral hashtags were used as keywords: “#immigration”, “#border”, “#HappeningNow.” Five pro-immigration hashtags were collected on: “#EndFamilyDetention”, “#WhereAreTheChildren”, “#AbolishICE”, “#FamiliesBelongTogether”, and “#OccupyICE.” Additionally, five anti-immigration hashtags were gathered: “#illegalimmigrants”, “#SendThemBack”, “#BuildTheWall”, “#ImmigrationReform”, “#SecureTheBorder.” Tweets were processed and stored in an SQLite database.

DESCRIPTIVE STATISTICS
In total, we analyzed 943,300 tweets and 53 pieces of metadata for each tweet, including tweet ID numbers, user ID numbers, ID numbers for media that is native, retweeted, or quoted, tweet ID of original content if it was retweeted or quoted, hashtags, and URLs. Of the near one million tweets, 759,397 were retweets (68%), 111,800 were original (10%), and 243,153 were quotes of another tweet (22%). Some tweets were retweets of a quote, explaining the combined total of these metrics being greater than total tweets in the dataset.

925,787 tweets (98%) came from unverified sources (including less prominent accounts such as eyewitness activists), and the remaining 17,513 (2%) came from verified sources, or more prominent sources such as journalists, organizations, or celebrities. Admittedly, this is not a flawless distinction, since the denomination was originally reserved for only ‘elite’ users, but Twitter has since made the application available to anyone willing to put the time into the lengthy verification process. Nonetheless, Twitter maintains that there are accounts of “public interest” and generally “maintained by users in music, acting, fashion, government, politics, religion, journalism, media, sports, business, and other key interest areas”\(^5\). 52% of the verified tweets were retweets, 9% were quotes, and 39% were original. Comparatively, for the unverified tweets, 81% were retweets, 8% quotes and 11% original. Overall, 22,898 tweets had at least one native piece of visual content (e.g. photo, movie, GIF), and the number of tweets with more than one piece of media decreased respectively until the maximum fourth image allowed per tweet. The total visual content count was comprised of 21,036 photos and 1,600 videos, and the remaining 262 were GIFS. The dataset also contained 80,030 tweets that were retweets of visual content and 57,715 quotes of visual content.

\(^4\) The Trump administration mandated a “zero-tolerance” policy by executive order on June 20, which calls for the prosecution of all individuals who illegally enter the United States. This policy has the effect of separating parents from their children when they enter the country together, because parents are referred for prosecution and the children are placed in the custody of a sponsor, such as a relative or foster home, or held in a shelter. (Politifact)

In order to capture collective narratives around visual content posted in the dataset, we conducted a process of iterative quantitative exploration of structural data and qualitative inquiry into products of the preceding quantitative network analysis. This process allowed us to narrow in on specific prevalent visual content frames circulating during data collection and identify discursive and organizational patterns common for users when interacting and framing visual content.

**Q1: CONSTRUCTING THE NETWORKS**

To investigate our first research question and construct a larger picture of the visual content networks, network graphs were generated by querying the database in R for quoted and retweeted unique media IDs associated with the original piece of visual content. The distinction between retweets and quotes was made to distinguish when a narrative about the shared content was being pushed, or when a new (potentially counter-) narrative was being created. Then, images (photos, videos, and GIFs) shared by the same users were connected and visualized utilizing Gephi, an open-source network analysis and visualization software. Each node represents the original tweet of visual content and are connected by edges that represent how many users retweeted or quoted both images—the thicker the edges the more the shared users engaged with both images and tied the nodes closer together.

The Gephi modality algorithm (Vincent D Blondel, 2008) estimated communities of images within the network for further analysis. Past studies have identified hashtags within clusters as important mechanisms in framing for social movements online (Stewart et al., 2017), and accordingly the most popular hashtags from these clusters were computed and used to give an overview of the most prevalent original narratives associated with each image cluster.

**Q2: EVOLVING NARRATIVES**

Qualitative analysis identified key original narrative stances and content analysis, using an existing framework, surfaced intent around each original image or video in the graph. This analysis further corroborated what the hashtags from Question 1 thematically suggested about prevalent narratives in each cluster. All coding was done by the first author, who was immersed in this work. Content analysis of the original content drew from Kow et. al (2016)’s four identified uses of social media for social movements as a framework to understand the intent behind each prevalent narrative. The codes were:

- **Information seeking and dissemination:** This included images of ‘protester rights’ flyers and videos posted by the ACLU.
- **Frame articulation:** This included tweets that were adding new angles to existing immigration rights frames, such as the ‘no one is illegal on stolen land’ frame.
- **Expression of solidarity:** This included tweets that were mainly composed of hashtags already established for the movement.
- **Development of counter-narratives:** This included visual content opposing an already existing frame, such as an image of a house with open front doors mocking the stance of open borders.

Then, subsequent discussion around a subset of the top 5 images from each of the 4 clusters were analyzed by identifying the quoted tweets with the most volume, or highest number of retweets (Arif et al., 2016). The high-volume quotes were determined by either being one of the top 10 most retweeted quotes for each image or until a quote had at least 2 retweets (in the cases where there were less than 10 quote tweets with high volumes). These popular narratives were compared to the original image narrative to better understand the evolution from narrative to counter-narrative.
Q3: MAPPING DISCURSIVE PATTERNS
Drawing on Maddock et al.’s (2015) work on rumor propagation, qualitative and visual analysis was applied to understand the origin and evolution of prominent counter-narratives identified in Q2. Specific patterns over time through visual signatures between narratives and counter-narratives were visualized and characteristics in users generating. While examining the dynamic structure of narratives and their respective counter-narratives we took specific notice of the reach and volume previous identified as important markers in rumor prorogation (Arif et al., 2016).

INTERVIEW ANALYSIS
Anecdotal evidence was collected through interviews with five separate members of the online activism community that regularly rely on social media to defend immigration rights. We conducted a grounded theory-based iterative and inductive analysis of the interview notes using coding and memoing (Corbin & Strauss, 1998) to sift out themes across interview data. The codes were reduced using axial and selective coding to the themes presented in the findings sections of this paper.
HIGH-LEVEL NETWORK STRUCTURE

Structures of framing and refraining

The first graphs constructed were networks of visual content with high levels of engagement: the first of tweets containing visual content that were retweeted at least once (Figure 1) and the second of tweets containing visual content that were quoted at least 20 times in the dataset (Figure 2). As in past research, retweeting signifies support and contribution to a growing narrative, whereas the process of quoting visual content entails additional context that shapes and tries to further propagate support or articulate a counter-frame for the visual content (Meraz et al., 2013).

Construction of the networks of shared users between this content, distinguishing between shared quoters or shared retweeters, visualizes patterns in creation and support of narratives around visual content. Communities within each network were computed and color-coded for each graph to distinguish visual content clusters.

Figure 1. Network of visual content with shared retweeting users from immigration rights protests

Table 1. Most used hashtags in clusters of retweeted visual content

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Hashtags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>BuildTheWall (63), MAGA (26), FamiliesBelongTogether (16), KAG (9), AbolishICE (8), buildthewall (7), AmericanFirst (7), ICE (6), OccupyICE (5), Trump2020 (4)</td>
</tr>
<tr>
<td>Multi-colored</td>
<td>FamiliesBelongTogether (2459), AbolishICE (663), FamiliesBelongTogetherMarch (450), familiesbelongtogether (252), EndFamilyDetention (192), abolishICE (127), FreeOurFuture (73), OccupyICE (73), KeepFamiliesTogether (67), abolishice (62)</td>
</tr>
</tbody>
</table>

Figure 2. Network of visual content with shared quoting users from immigration rights protests

Table 2. Most used hashtags in clusters of retweeted visual content

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Hashtags</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>MAGA (8), BuildTheWall (7), AbolishICE (7), FamiliesBelongTogether (4), KAG (2), Illegal (2)</td>
</tr>
<tr>
<td>Purple</td>
<td>FamiliesBelongTogether (32), AbolishICE (12), EndFamilyDetention (4), WomenDisobey (2), womendisobey (2), FreeOurFuture (2), FamiliesBelongTogetherMarch (2)</td>
</tr>
<tr>
<td>Blue</td>
<td>FamiliesBelongTogether (24), FamiliesBelongTogetherMarch (5), EndFamilyDetention (3), AmericaWelcomes (2)</td>
</tr>
<tr>
<td>Red</td>
<td>AbolishICE (13), FamiliesBelongTogether (4), abolishICE (3), OccupyCELA (2), OccupyICE (2)</td>
</tr>
</tbody>
</table>
The first shared users graph (Figure 1) is shared retweets of images (two images connected by the same user if the user retweeted both images) and shows two distinct super clusters of visual content: a single cluster to the right (black) and another on the left (multi-colored). The top 10 hashtags the right cluster and left multi-cluster were examined and (Table 1) revealed homogenous, clear left-leaning/right-leaning political divide in, and siloed consumption of, visual content. This siloed or ‘echo chamber’ behavior is a common characteristic noted by past network analysis of the spread of contentious content online (Starbird, 2017; Stewart et al., 2017; Arif et al., 2015). Comparatively, the second shared users graph (Figure 2) of shared quotes for images (two images connected by the same user if the user quoted both images) illustrates more intertwined clusters of images, indicating more connected images and shared users. Hashtags for clusters were more varied (Table 2), especially the green cluster, which notably has the least number of images but the highest number of unique hashtags. This implicates more diverse users interacting with visual content outside of their cluster boundaries, contrary to the more isolated spread of visual content in the retweet graph. Further inspection of hashtag use in the quoting graph showed that co-opting also was prevalent, a strategy previously identified (Stewart et al., 2017) that is commonly used to disrupt frame development on social media. Notably, this phenomenon is not one-sided, in several instances the MAGA hashtag was appropriated by march attendees. Two examples from this cluster are:

Figure 3. Two examples of hashtags co-opting concerning the End Family Separation movement

Node size is determined by the degree, or how many edges connect it to another node, of the visual content. Accordingly, degree is a robust measurement of image popularity and telling of the main narratives in each cluster, since the more users that retweeted/quoted that photo or video and another image in the graph, the larger the degree. Large nodes can be seen more prominently at the center of a cluster or acting as a bridge between two clusters. The evolving narratives around these popular images will be discussed further in proceeding sections, but these graphs alone indicate notable images acting as bridges, or a piece of content that connect two otherwise disparate clusters of images. For Figure 1, these shared retweeters are much less common, creating the siloed clusters as made visible by the graph, in which the same narratives are perpetuated through a refraining process without interruption from counter-narratives (Meraz et al., 2013). In the case of the shared quoting users, in which text can be appended to visual content and shared with new opinion or context, the bridges represent a negotiation of framing around images. The shared quoter graph is a visual representation of a constant negotiation and framing around images where the retweet graph represents images and their respective frames surfacing through repetition and shared ideology with little interruption.
Trends in engagement behavior
The networks represents patterns across all visual content in the entire dataset. A more granular examination of interactions with visual content builds a clearer picture of online behavior beneath the nodes and edges of the graphs, including difference in behavior around types of media and interaction with formal and informal sources. Each node in the graph is a piece of visual content, but the analysis shows not all media is interacted with equally. Photos were over 13 times more likely to be shared in the dataset, but the engagement with video, specifically retweets, was significantly higher. For each cluster, the majority of media was photos, ranging from 72-81% photos in each group.

Patterns around the binary ‘verified’ category, assigned to each user, were examined to distinguish content sharing behavior between more prominent sources (such as journalists, organizations, or celebrities) versus less prominent accounts, including the accounts of eyewitness activists. Admittedly, this is not a flawless distinction, since the denomination was originally reserved for only elite users, but Twitter has since made the application available to anyone willing to put the time into the lengthy verification process. Nonetheless, Twitter maintains that there are accounts of “public interest” and generally “maintained by users in music, acting, fashion, government, politics, religion, journalism, media, sports, business, and other key interest areas”. Verified users represented only 2% of the dataset yet, they shared proportionally twice as much visual content than unverified sources. Both types sources were equally likely to share videos in the dataset, but verified sources were more than twice as likely to share photos. Subsequent tweeting and retweeting of visual content was also more likely to happen with content from verified sources (Figure 4).

Past research suggests that online activism is made possible largely because of the breakdown of ‘gatekeepers’ or prominent actors controlling the narratives. Contrary to this, while there were significantly more pieces of original unverified content (20,179) as compared to verified content (2,792), verified content was over twice as likely to be quoted and nearly 5 times as likely to be retweeted than content shared from unverified sources. While there is opportunity for narratives around visual content to be crowd-sourced to prominence, content generated by verified sources, such as more prominent figures, are more likely to be engaged with and contribute to social movement frame formed with visual content. This is not necessarily negative, since content from verified sources is less likely to reflect misinformation and mal-intent, but illustrates the tension between horizontal communication and verified, reliable content often confronted by activists online.

CHARACTERIZATION OF THE CLUSTERS OF ORIGINAL VISUAL CONTENT

A complex right-leaning fringe (Figure 2 green cluster). Popular visual content in this cluster came from the account AlwaysActions that consists of mugshots from ‘illegal aliens’ that had committed horrendous crimes. Accordingly, common narratives that surfaced through content analysis, both of the original text and the subsequent quotes, revolved around safety, truth, and the fear of open borders. The original frames of these images remained generally uncontested in proceeding quotes of the image, 95% of high-volume preceding quotes were coded as solidarity with the original narrative. Only one quote from these images was a counter-frame, being a video of police lined up for a march and one user pointing out the other duties, such as protecting the city from crime and murder, that their time could be better spent on.

From these instances, the cluster of original frames and appears generally homogenous in narrative and user ideology, but of the top 5 most popular images, only 2 were aligned with the right-leaning narratives mentioned above. Other popular tweets of visual content had originated from left-leaning activists and were co-opted by this cluster, to the extent that they were grouped into the same (green) community. For example, the bridge between this clusters and the others is a photo of protesters from an account with the username NYC DSA, the New York Democratic Socialist Association. This was the most quoted image in the cluster and the entire dataset, alone this image was quoted 4,904 times. Of the top 10 most popular quotes of this image, all were counter-narratives to the original post. Popular quotes of the image ranged from civil discourse on socialist positions to claims that the photo depicted a modern Maoist uprising. Notably, all high-volume quotes reframing this image were verified, whereas all other sets of high-volume quotes of images in the cluster were at least 75% unverified.

PROMINENT NARRATIVES AND COUNTER-NARRATIVES

The metadata (tweet text, user, description, URLs, etc.) surrounding each of the 120 pieces in the shared-quoter network of original visual content and then the top-five quotes of popular images in each cluster were collected and analyzed. In each cluster common frames around immigration were found and the applied framework from Kow et al. (2016) revealed distinct motivations behind the use of Twitter for activism. Framing of each original piece of visual content aligned with popular hashtags discovered in the network graph section, both in content and homogeneity. Conversely, though, frames generated by other users through popular quoting of the original images did not always align with the intent of the original content and often ran contrary to the initial framing.
Closely tied protest clusters (Figure 2 blue and purple cluster). These two clusters represent similar communities of images and narratives in an already well-articulated movement. Frames surfaced around humane treatment of immigrants, dissent against the current administration, mobilizing marchers for protesting, and democracy. These clusters had premeditated frames for the End Family Separation marches, and thus solidarity was the most common type of engagement (75%), mostly reiteration of existing march hashtags, then information dissemination (15%), and frame articulation (10%). These clusters were largely homogenous and high-volume quotes of the original visual content was engaged with to show solidarity or disseminate information for protesters. Only about 10% of the high-volume quotes of these images were counter-narratives, explaining the homogeneity. In other words, users that were quoting these photos and videos were showing agreement and support of what the visual content represented.

An organized Occupy movement (Figure 2 orange cluster). Of all clusters, the orange used social media engagement to disseminate information (23%) for organizational purposes and was most characterized by Occupy and Abolish ICE narratives. As these were more information-based and generally not used to promote an agenda, these tweets had the least amount of subsequent quoting, and instead demonstrated a higher percentage of retweets. As with the blue and purple clusters, solidarity was most, but the second most common engagement was Occupy/Abolish using photos and videos to communicate where and when sit-ins were happening and information on arrests of protesters that were taking place. The Occupy cluster, although also representing far-left ideology, seemed less vulnerable to online attacks because they used social media to spread information rather than share a political stance more than the other clusters.

QUALITIES OF PROMINENT COUNTER-NARRATIVES

The interdependence of image narrative and counter-narrative. For every narrative that was coded in the original set of visual content, there was a counter narrative, such as Democracy/Snowflakes protesting, Immigration/Open Borders, and Legal/Illegal immigrants. For example, in response to the NYC DSA’s image of protesters holding “Abolish ICE” signs during the End Family Separation march, one user quoted the photo and responded with a warning of communism:

(June 30, 6:54pm): We are witnessing a Maoist uprising in realtime [link]

The “Impending Communist” description was a common theme in many counter-narratives and easily played into the “Snowflake” narrative that discredited protesters and young marchers expressing their socialist-leaning views. This communist reframing of a socialist protest image was retweeted 257 times in this dataset alone.

Stoking fear with images. The long history of image campaigns targeting social groups has carried over into the digital sphere. Counter-narratives and the photos and videos they are used to contextualize play up existing stereotypes like the history of anti-immigration propaganda has done before. As mentioned above, central to the green cluster was an account, AlwaysActions, that claimed to share “breaking news and developing stories” in their description. This was a highly active account, with 106 tweets in the week span of collection alone, and appeared to target several social groups, including Muslims, African Americans, and immigrants, through tweets and retweeting of conspiracy theories and out-of-context images. In our dataset this account shared 4 images, each depicting the mugshot of a Latino man and detailed that he was seeking asylum in the U.S. from a horrendous crime the he had committed as an ‘illegal immigrant’. For example, the following tweet was accompanied by a photo of an immigrant:

(July 3, 1:54pm) #Mexican Killer Illegal alien #Sneaks into the United States 7 times #MAGA #BuildTheWall [link].
While this sounds like a fringe incident, these images were quoted and retweeted 632 just within the dataset and can be assumed to be disseminated further given the 47,594 followers the accounts had at the time of collection. The ensuing comments of these tweets were comprised of users genuinely scared for their safety and calling for a wall to be built at the Mexican border that could prevent immigrants like the one pictured from entering the U.S. These posts also prompted more gruesome, violent comments that called for physical punishment of the immigrant and those like him.

Harassment. Racial slurs, sexism, and calls for violent actions against protesters and immigrants were also noted in the counter-narratives. Several attacks on protesters were geared at sexual harassment of women and used photos to invite further commenting and misogyny. One user posted a video of lined up police in full-armor preparing to Occupy protesters which baited conversation of violence against the peaceful protesting. Others used derogatory terms for Latino immigrants, one of which was retweeted over 100 times in the dataset alone. While it is widely known that these unfortunate and demeaning attitudes exist online, the normalized harassment of immigrants and activists play an important role in deterring participation in online advocacy spaces, as activists recounted in interviews covered in the anecdotal section.

DEVELOPMENT OF MISINFORMATION AROUND IMAGES

This section analyzes who spread the narratives uncovered in the preceding section and how successfully they spread. To understand the underlying processes needed to develop misinformation about visual content requires a more granular examination of discursive patterns for individual visual content, more specifically how narratives and subsequent counter-narratives about immigrants are co-created and developed over time. This section focuses on counter-narratives identified by previous analysis that rose to prominence over the original tweet while co-opting the visual content (Table 3). Notably, only one of these prominent counter-narratives were articulated around a video.

Next, iterative quantitative and visual analysis of dynamic quoting and tweeting of these high-volume counter-narratives illustrated how new context for visual content rose to prominence. The reach, or follower size of each user (Arif et al., 2016) was taken into consideration, and 100% of co-opters had a broader reach than the poster of the original tweet. This is in line with Arif et al.’s (2016) research that found reach and volume to be just as important to rumor spread as the actual content. Interestingly, while verified sources were twice as likely to share native visual content than unverified sources, high-volume counter-narratives of the same content were over twice as likely (70%) to be created by unverified sources than verified.

Table 3 Immigration rights narratives and the counter-narratives that rose to prominence over them

<table>
<thead>
<tr>
<th>Narrative</th>
<th>Counter-narrative</th>
<th>Narrative Audience</th>
<th>Counter-Narrative Audience</th>
<th>CN Prominence Increase</th>
<th>Media Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissent</td>
<td>Communism</td>
<td>17,876</td>
<td>27,282</td>
<td>10%</td>
<td>Photo</td>
</tr>
<tr>
<td>American values</td>
<td>Working Class</td>
<td>2,370</td>
<td>23,258</td>
<td>40%</td>
<td>Photo</td>
</tr>
<tr>
<td>Freedom</td>
<td>Illegals</td>
<td>5,056</td>
<td>157,985</td>
<td>75%</td>
<td>Photo</td>
</tr>
<tr>
<td>Humane Treatment</td>
<td>Criminals</td>
<td>8,407</td>
<td>18,734</td>
<td>5%</td>
<td>Photo</td>
</tr>
<tr>
<td>Dissent</td>
<td>Authority</td>
<td>2,309</td>
<td>24,627</td>
<td>48%</td>
<td>Video</td>
</tr>
</tbody>
</table>

There was a mix of patterns identified for counter-narratives rising to prominence. In the instance of the New York DSA image, the major bridge that connected the right-leaning and left-leaning clusters, co-opting the counter narrative consistently overpowered the message of the original photo (Figure CN1) within hours. The instance in Figure CN1 shows one of three counter narrative that rose to prominence of the original DSA photo, this one specifically targeted Democratic primary victor Alexandria Ocasio-Cortez for being a dues-paying member of the group.
Figure 4 Example of a growing counter-narrative rising to prominence over the original narrative

Figure 5 Example of an immediate drowning out of narrative by counter-narrative

Figure 6 Example of a resurgent counter-narrative
In several instances there were other attempts before to offer a reframing that did not reach prominence, but largely adopted counter-narratives took hold once more prominent figures created the necessary counter-frame for the visual content. For example, an account with only 13 followers and no original posts quoted an image and mentioned another user with a high audience reach (24,627 at the time) and the prominent user proceeded to quote the original image of a young woman protesting the police, which quickly overtook the popularity of the original video (Figure 5).

In all cases but one a narrative was consistently drowned out by a counter narrative. In the other instance, a dormant counter narrative that gained popularity after the original narrative had plateaued (Figure 6) and eventually rose to prominence over the original framing. This could be explained by a second prominent figure retweeting the counter narrative hours later. A user with close to 12,000 followers retweeted the counter narrative around Tuesday at 13:00 followed by another with close to 72,000 followers, explaining the new spike in the graph and supporting this hypothesis.

Experience of immigration rights activists online
The twitter data poses limitations on the conclusions that can be made concerning the qualitative experiences of activists and immigrants using social media, so interviews were conducted to supplement the network analysis. These interviews added anecdotal data to enrich the stories told by the tweet data alone. A diverse group of five immigrant rights activists that were recognized for their personal and/or organizational online presence were interviewed about their advocacy experiences online. Invariably, each activist iterated the importance of social media for immigrant rights activism work. The ability to reach millions of people in and outside immigrant audiences, to engage in policy-changing conversations, and share crucial information on legal issues is invaluable for these communities. Social media benefits always outweigh the harm done by bad actors perpetrating harassment on platforms. That said, there are four key themes that surfaced amongst the interviewees which emphasize the use and threats of immigrant rights activism online.

1. AFFORDANCES OF THE IMAGE
Interviewers noted transitions in the last few years towards image-based campaigns, especially for raising awareness of lived immigrant experiences. On a daily basis, activists utilized stories told through visual content to raise bond for immigrants, organize call-ins to advocate for the termination of a deportation, and spread information through formal videos on DACA applications and legal advice. One activist’s organization started engaging with their community through YouTube since immigrants were already sharing their stories with each other as videos on the platform. Another interviewee was raising money to support bail for unjust misdemeanors, which was difficult because they were seen as ‘criminals.’ Instead of using written words from people they told their story through video instead and were able raise $2,000,000, more than the expected $70,000. The transition to image-based campaigns, though, has invited new tactics for the targeting of immigrants and activists. Interviewers shared stories of digital “know your rights” flyers from their organization being photoshopped to spread contradicting information on ICE raids, video-taping activists and portraying them as ICE officers online, and reframing images to represent immigrant organizations as white supremacist supporters. Multiple interviewers iterated that perpetrators are tech-savvy and have brought their own cameras to live-stream immigrant event crashing, use social media to track and disrupt activism events, and have created memes of minorities looting after a natural disaster.
2. DIGITAL TO PHYSICAL SPACES
The evidence of physical violence, incited by false narratives, malicious tactics and mobilizing on social media, carrying over into physical spaces is the highest concern for activists. Groups are being tracked on social media and their information sessions and protests are being intruded on by bad actors with an intent to harass immigrants. Activists are seeking solutions around not with platforms to protect their communities from physical repercussions of online behaviors. One interviewee said “Monitoring hasn’t really happened, especially stuff around the anti-immigrant stuff and blatant lies about immigration not being addressed. There’s not much faith in corporations to provide this. The real question is: How can we provide that safety outside of these digital structures?”

The immigration crisis was and continues to be a confusing time for immigrant communities in the U.S. This uncertainty is intentionally amplified online where images are strategically used to sow further confusion about the identity of ICE officers and deter immigrants from attending informational sessions. This physical carry-over from digital spaces is what some interviews cite as the most dangerous and characterizing feature of disinformation.

3. ABSENCE OF FORMAL STRATEGY
At the time of the interviews, no major attacks were recounted by participants, but small-scale attacks were ubiquitous. In general, the most used strategy for human rights organizations interviewed is to drown out misinformation surrounding their content which was effective strategy given the most common attacks were accounts with small followings. To this end, no pre-emptive formal strategies were in place for these human rights organizations if a more orchestrated attack occurred, although these type of attacks from prominent actors showed to be the most effective in spreading misinformation, according to our network analysis. While organizations iterated that they felt comfortable in their ad hoc strategies for managing attacks, they recounted that often the individuals from their communities are more frequently targeted after engaging with their pro-immigrant content, which fettered overall participation and self-advocacy. Overall, no organization voiced trust in platform accountability and intervention as an option for misinformation management.

4. INCLUSIVE SOLUTIONS
Although platform responses have begun to take shape, immigration activists believe these solutions aren’t taking into account the most affected by malicious online practices. For example, one interviewee cited a recent Twitter account audit that not only terminated harmful accounts in June, but also a key organizer in the Black Lives Matter marches. Subsequently the march was sent into disarray and could have been avoided would major voices of social rights activist organizers have been present in the conversation. Unfortunately, although immigrant populations make use of social media, several interviewers noted that their audience is comprised mostly of white allies without the lived experiences of an immigrant, who are less subjected to harassment and intimidation while advocating for immigration rights online. If those most affected are reluctant to participate online to advocate for themselves, then allies that feel comfortable in these digital spaces are still running the conversations.

These interviews gave us a glimpse at the day-to-day hurdles of online immigration activists that the Twitter data alone cannot tell. Overall, activists voiced their daily processes in working around the systems in place since platforms aren’t doing enough to guarantee the safety of their immigrant communities. The users have developed their own tactics to deter bad actors, such as creating background patterns on flyers to make photoshopping more difficult, that provide safety outside of structures in place that aren’t working. Misinformation, if not on a mass scale, isn’t so obvious, especially to mediators of platforms that are controlling billions of users and their data at a time. But these interviews shed light on the tangible and daily realities overflowing from digital spaces that immigrant communities and their allies encounter.
DISCUSSION

The assertion that the objective camera lens will invariably convey truth is an unreasonable expectation for the evolving image online. In the same way the contents within a video or picture frame can mask or uncover realities, human fallibility, or worse, mal-intent, can shape the narrative around visual content as it circulates on social media. Those most vulnerable to mis-use of visual content are already marginalized voices, which can become subjected to harmful counter narratives playing on historical propaganda techniques like stereotypes and/or drowned out by more prominent figures on social media extolling these harmful narratives. Past research has suggested that opposing ideologies are generally siloed throughout the process of frame and counter-frame development when deliberating over text alone. The analysis of discursive patterns around visual content has shown to be more complex and encouraging of multiple perspectives to participate in the framing of images, although qualitative work has shown these negotiations to vary in civility and effectiveness.

Human rights activists have an equally complex relationship with social media—the ability to organize audiences that otherwise would be unreachable is indispensable and rallying around images and symbols of movement create a larger sense of purpose and identity has become vital, yet there is the inevitable cost of bad actors using the same affordances of social media.

In various forms, every finding in this study underlines the power of prominent figures in shaping how we perceive the images presented to us. Given the rapid exposure and quick emotional response to visual content, it’s clear that images particularly lend themselves to recontextualization and amplification in counter-narratives by key influencers online – whether or not that account is verified or the new narrative matches the image. This makes images especially predisposed for contextual misinformation. The following recommendations are for practitioners and activists who use social media for human rights initiatives, based on the evidence uncovered in this report.

1. To photo or video? There’s not enough evidence in this study to conclusively recommend one medium over the other. That said, videos receive significantly higher engagement and, of all the images that facilitated the rise of counter-narratives in this dataset, only one was video. Verified sources were also twice as likely to share videos over unverified sources, suggesting they are a preferred medium for more prominent sources.

2. Develop a plan for online attacks. All participants had experience with small-scale attacks, but strategies were premeditated for combating larger attacks. It’s important to know answers to questions like: When do you engage with the attacker? When should you join a conversation to support a community member? When do you report an account? Organizations addressed small-scale attacks ad hoc but given the quick spread of counter-narratives found in this study, a predetermined plan could temper an attack that could escalate in just a few hours.

3. Remain aware of radical accounts with large followings. Harmful attacks will come from accounts with large followings and have enough clout to convince people that their narrative about media is more correct than the original information. Solidarity with partner organizations and other users is the best tactic against misinformation.

4. Unverified sources are more likely to create counter-narratives around visual content and content generated by community members.

5. Protect individuals in the community. Since attacks often targeted users with small followings, most victims were single users and not large organizations. For immigrant rights organizations with a larger audience, if there is an individual whose media has been targeted or mis-contextualized online, simply correcting the context and misinformation on an organizational account can go a long way in combating misinformation.
6. **Less fear, more engagement.** If the target audience is not directly engaging, but instead allies that feel safer advocating online, then supporting and standing by individuals mentioned in #3 could lead to more community engagement. It’s important to have allies, but creating digital spaces where people feel safe to advocate for themselves leads to more diverse conversations and solutions.

7. **Communicate the risks of sharing controversial media online.** Given the rampant nature of sexism, racism, and calls for violence online, harassment around photos and videos are not exempt from this trend. Warn community members about the risk of harassment that may stem from sharing photos and videos about contentious topics like immigration rights and disseminate tips on how to address it. Several participants noted that attacks tend to develop around specific events, like marches or protests, and around social justice issues circulating in the news at the time.

8. **Join the conversation on platform misinformation solutions.** Solutions aren’t comprehensive until they represent those most affected, and take into account their needs. Community members and the organizations that represent them need to advocate for their digital well-being so the stories of marginalized voices can rise to prominence.
REFERENCES


