7 Exploring the Application of Computer Vision to Detect Faces and Emotions

An Analysis of Radical Right-Wing Leaders

Ignacio-Jesús Serrano-Contreras and Javier García-Marín

1. Introduction

The Internet was originally created to facilitate quick and global communication (Berners-Lee, 2000). From simple written messages to the rise of audiovisual platforms, all forms of media have undergone hybridization through this interconnection. This blending of media environments has not only democratized access to information, but has also empowered individuals to actively participate in content creation and distribution (Luengo, 2014), with social media standing out as a prominent example. The latter have played a pivotal role in shaping narratives and influencing public opinion, ultimately leading to a transformation in the dynamics between media, society and power (Maldonado, 2016). Nevertheless, there are divergent perspectives on the impact of these developments. On the one hand, they have been studied as potential catalysts for phenomena like disinformation, political polarization or populism (Gerbaudo, 2018). On the other hand, contrasting views exist (Lorenzo-Rodríguez & Torcal, 2022), and there is ongoing debate regarding the actual role of elements such as algorithms (recent contributions by García-Marín & Serrano-Contreras, 2023 or Ibrahim et al., 2023, offer insights into the role that computation can play in communication).

In the realm of politics, the digitalization of mass media has presented numerous opportunities for politicians to leverage and yield favorable outcomes. Platforms like Twitter and Facebook have been harnessed to bridge the divide between politicians and the general public effectively bypassing traditional media channels. This has allowed them to disseminate their messages, engage with their followers, mobilize their support bases and incite public action to champion their aspirations or to express discontent (Rahyadi et al., 2023; Segerberg & Bennett, 2011; Weismueller et al., 2022).

DOI: 10.4324/9781003457381-8

Today, the landscape appears to have evolved. While we have witnessed the professionalization of broadcasting, issues like the bi-directionality of communication, which shaped the digital agora, are no longer as central. However, a significant development in recent times has been the integration of audiovisual content into media campaigns alongside written messages. This shift has resulted in a transformation of the political landscape, with Twitter's prominence giving way to new strategies represented by new actors like TikTok (Zeng & Abidin, 2021). These new platforms support particular logics depending on the media (Altheide, 2020).

This digital age prompts questions about the evolution of both society and media. On the one hand, some of these questions have already been addressed, as studies exploring the scope and definition of the Internet and social networks reveal, although continuous reflection remains valuable (Sterrett et al., 2019). On the other hand, several challenges persist without robust or widespread solutions. As the Internet has undergone significant changes, scholars in the fields of social and human sciences have directed their research efforts toward comprehending the roles played by text and symbols, such as expressions and emojis (Barbieri et al., 2017; Debnath et al., 2020), as well as interactions between users. However, it is worth noting that most of these studies are yet to contribute significantly to the understanding of large-scale image and video broadcasting. Thus, there remains a need for a broad and experimental foundation to delve into complexities such as non-verbal language, the implications of color or the biases that can arise from different visual perspectives, which allow us to better connect the impact to specific social phenomena. Related to this, we wonder if using emotions in non-verbal language follows patterns we can analyze. As social networks have weakened the role of information filters represented by traditional media, it is interesting to see if the most extremist (and, therefore, populist) candidates take advantage of them to use specific strategies. Thus, our research question is whether we can detect patterns in the use of facial emotions in populist candidates.

2. Theoretical Framework

Various approaches have emerged to deal with the new drifts of mass communication. For example, text mining has equipped the social sciences and humanities with a potent tool capable of converting textual code into structured data. This process, known as "text as data" (Gentzkow et al., 2017; Grimmer et al., 2022), involves measuring sentiment in texts, assessing semantic distancing, or identifying the most frequently used verbs in a book, allowing for the transformation of qualitative information into quantitative values. Moreover, text mining represents just one of the initial phases in the broader field of Natural Language Processing (Serrano-Contreras, 2021), a discipline at the intersection of computer science, psychology and linguistics. However, the current media ecosystem necessitates methodologies that incorporate new tools capable of automatically extracting information from audio, images, or video content.

Spurred by the revolution brought about by the rise of television, streaming platforms, as well as video social media, generate an enormous amount of audiovisual content. In response to these transformations, the political arena has adapted by embracing trends of spectacularization (Gómez-García et al., 2023; Khoma & Kozma, 2022). This change places a strong emphasis on the mediatized image, surpassing text and therefore calm discourse in importance (Plazas-Olmedo & López-Rabadán, 2023). This symptom is one more factor that highlights the Americanization of politics, particularly in Europe, where visuals and audiovisual constructs can prove to be more effective in capturing citizens' attention and, in turn, mobilizing them (Durántez-Stolle & Martínez-Sanz, 2019; Tai & Turner, 2008). Thus, it becomes essential to design both descriptive and experimental analyses to empirically assess the extent to which these trends can impact audiences.

In recent years, algorithms in this field have been increasingly harnessing sophisticated developments, capitalizing on significant advances in computing. Presently, the concept of computer vision is emerging from disciplines in which machine learning models can make inferences from images and videos, including faces and objects. These new techniques bring forth fresh insights for a wide range of disciplines that were previously limited to manual analysis on a small scale. These innovations pave the way for such methodological advancements to be crucial in the understanding of contemporary societies (Schmøkel & Bossetta, 2022). Contributions in this domain help answer a wide array of questions in fields such as political communication, where content analysis can be enhanced through automation (Peng & Lu, 2023).

These advancements entail the ability to synthesize elements that were previously subject to the researchers' subjectivity. Machine classification, in particular, involves two main components. First, pre-trained models, developed largely through multidisciplinary research in which diverse fields in the social sciences and humanities establish guidelines for computerizing non-verbal cues, tones and rhetoric (Joo et al., 2019). Secondly, these automated algorithms can perform massive analyses, enabling the examination of hundreds of hours of video or thousands of images as an option. These breakthroughs enable the field of political communication to continue its exploration of theories like framing or priming, as well as to incorporate new areas of study into the analysis of political polarization and radicalization.

Our approach is specifically designed to open up new avenues for social inquiry. By analyzing large datasets, we aim to propose and explore fresh theoretical perspectives. In this pursuit, we concentrate on psychological factors to test whether these variables can detect distinct behaviors, thereby shedding light on potential theories of effects and concomitant structures within ideologies (Cichocka & Dhont, 2023). Our approach entails an analytical strategy capable of detecting sentiments in videos of politicians posted on YouTube. Our goal is to uncover, if they exist, common patterns among leaders framed as populist and radical right-wing. For this purpose, we propose the integration of computer vision techniques, specifically the sub-discipline known as FER (Face Emotion Recognition) (Khaireddin & Chen, 2021). FER employs models to detect emotions such as sadness, happiness or anger in facial expressions. Our mission is to identify possible patterns of non-verbal language among politicians, particularly among different leaders of the radical right, including figures such as Donald Trump, Giorgia Meloni, Santiago Abascal, Marine Le Pen or Javier Milei. Simultaneously, we intend to establish a control group featuring leaders from different ideological backgrounds.

3. Data and Method

In this research, we analyzed YouTube accounts belonging to political leaders and parties. Hence, our criteria are based on the video's popularity, not time or date. Our premise was that the official channels of politicians and parties are instrumental in conveying the image they wish to present to their voters. Therefore, we excluded all those audiovisual fragments shared by other profiles, as these may have edited the original samples and could have introduced biases or misunderstandings. In our endeavor to identify ideological patterns not necessarily tied to a specific society, we initiated our investigation by seeking the verified channels of several prominent leaders within the populist right. These leaders encompassed Donald Trump (USA), Giorgia Meloni (Italy), Santiago Abascal (Spain), Marine Le Pen (France) and Javier Milei (Argentina), in alignment with Pelinka's (2013) conception of populism. Although these profiles exhibit diversity, they were selected to test ideological patterns. We also aimed to include the president of El Salvador, Nayib Bukele, in our sample; still, his inclusion was discarded as he did not fit the predefined characteristics of our research design.

The final sample was compiled by selecting all those videos in order of popularity, with a maximum of 10 for each channel, provided that the duration of each was less than 5 minutes. We understood that longer videos would introduce unnecessary complexity to the research without yielding significant information. Additionally, the selected videos had to include a range of shots, from the wide cowboy shots to close-up shots, to ensure that the visibility of the faces was unambiguous and not compromised by the presence of third parties, the quality of the video or profile camera shots. The total sample size ($n = 50$) was then expanded to include leaders considered non-populist in the same countries: Joe Biden (USA), Giuseppe Conte (Italy), Pedro Sánchez (Spain), Emmanuel Macron (France) and Alberto Férnandez (Argentina). As a result, the total number of videos analyzed, comprising both the analysis sample and the control group, amounted to 100.

Subsequently, we proceeded to select one image per second from the videos, considering that they later were encoded at a rate of 25 frames per second. This resulted in a total of 6,981 images. Next, employing the Python programming language, we implemented a face detection algorithm. In our case, we opted for face-recognition due to its simplicity in Python and its effectiveness in face recognition with minimal training required. For the training of the algorithm, we gathered images online by conducting specific searches on DuckDuckGo (using the politician's name and the image search tool). We selected only those images that exclusively featured the politician's face, without any other additional images that could interfere with the detection process. Then, we designed the code to perform a series of steps; firstly, it detected the faces in the images from the videos (using OpenCV, which is practically a standard tool for this purpose); then, it recognized the detected faces using the face-recognition library; finally, it determined the emotions expressed in the faces using FER. The final output was a DataFrame (DF) containing the name of the recognized candidate, the timestamp of the detection and the video file. The entire process is outlined in Figure 7.1.

Note that the tools employed can detect several faces simultaneously, along with their associated emotions. Additionally, the algorithm can label

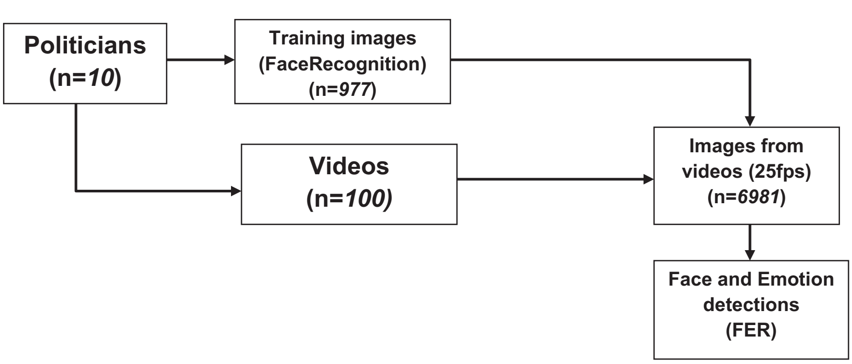

Figure 7.1 Flowchart.

a face as "unknown" when it cannot recognize a face from its training set, although it does continue to analyze its emotion. It is important to acknowledge that the algorithm can occasionally make errors, which means that some manual verification has to be carried out. Nonetheless, the incidence of errors was relatively low (only 5 out of over a 100 detections checked).

4. Results

Table 7.1 presents the raw results of the analysis, including the unknown encodings (both in the location of faces and emotions). This helps gauge the degree of effectiveness of the tools employed. The slight discrepancy between the total number of encodings (6,970) and the total number of images (6,981) represents localized errors, a small and acceptable number. As mentioned earlier, the overall effectiveness is sufficiently high for use in research settings; nevertheless, expanding the training sets, while relatively straightforward for public figures, may present challenges with other types of actors.

The data in Table 7.1 reveal several intriguing findings. First, there is a significant difference between the number of codings by name and those related to sentiment. This variance in codings by name can be attributed to two factors: the differing lengths of the videos and the prominence of the politicians features in each of them. It was also important to analyze the proportion of other faces in the videos, and the data yields some striking insights. Notably, the videos of Conte and Fernández exhibit the highest proportion of unknown faces (.17 and .16, respectively); while Trump's, Abascal and Meloni's videos feature the lowest proportion (.006, .03 and .03, respectively). On average, for politicians considered populist, the proportion of unknown faces is .04; whereas, for other candidates, the mean is .10; this represents a significant difference (with $\bar{x}$ - = .07). However, at the same time, this data also speaks to the utilization of personalism as a political strategy, and it is challenging to extrapolate broad conclusions when multiple variables may influence this relationship, from the political system (presidential or parliamentary), or the recognition of the individual and their personal style of politics.

Nevertheless, Table 7.1 reveals another remarkable finding, namely, a pronounced prevalence of feelings of sadness in the sample under analysis. Initially, one might attribute this to a poorly calibrated algorithm. However, drawing from our experience with the tool, we have reason to believe otherwise. If we exclude this possibility, we are left with the assumption that politicians have a proclivity for expressing sadness in their communications. This tendency may be due to their routinely employed communication styles, which would explain its prevalence.

Table 7.1 Coded emotions per name

Name	*Unknown*	*Anger*	*Disgust*	*Fear*	*Happiness*	*Neutral*	*Sadness*	*Surprise*	*Total*
Abascal	18	69	0	9	213	132	**268**	18	727
Biden	34	4	0	0	4	7	**206**	2	257
Conte	11	35	0	7	144	150	**225**	5	577
Unknown	16	79	0	7	211	257	**278**	15	863
Fernández	7	30	0	16	65	81	**124**	7	330
Le_Pen	67	113	0	124	291	301	**665**	72	1,633
Macron	13	73	0	8	48	43	**120**	26	331
Meloni	3	26	0	46	25	35	**120**	76	331
Milei	9	41	0	13	50	34	**102**	15	264
Trump	130	339	1	28	78	190	**694**	12	1,472
Pedro_Sánchez	3	48	0	4	22	31	**75**	2	185
Total	311	857	1	262	1,151	1,261	2,877	250	6,970

Figure 7.2 illustrates the sentiments assigned to each political figure, with a grouping based on populism (represented by the shaded area). The figure vividly confirms our previous observations: All politicians intensively employ facial expressions that convey sadness, which is the most common emotion among all of them. Similar trends are observed with other emotions codified in our research (i.e., neutrality, happiness and anger). Notably, disgust is absent, and there are minimal instances of surprise and fear.

The differences observed can be attributed, once again, to their individual styles of expressing ideas, as well as to shared (probably global or, perhaps, Western) communication techniques. Thus, we can discern what, in principle, appear to be different styles: Meloni uses expressions of fear or surprise more frequently than any other individual, while Trump, Macron or Sánchez exhibit a higher prevalence of anger; happiness, on the other hand, seems to be more prominent in Abascal, Conte and Fernández. Can we categorize these results based on the populism variable? Figure 7.3 shows the result of such grouping in relative values. Although we do observe variations that might suggest specific patterns, especially in the use of happiness and neutrality, it is challenging to draw scientifically valid conclusions. So, as we suspected, the results do not provide substantial information, thus reinforcing the notion that we are dealing with individual styles rather than trends that can be attributed to factors such as populism.

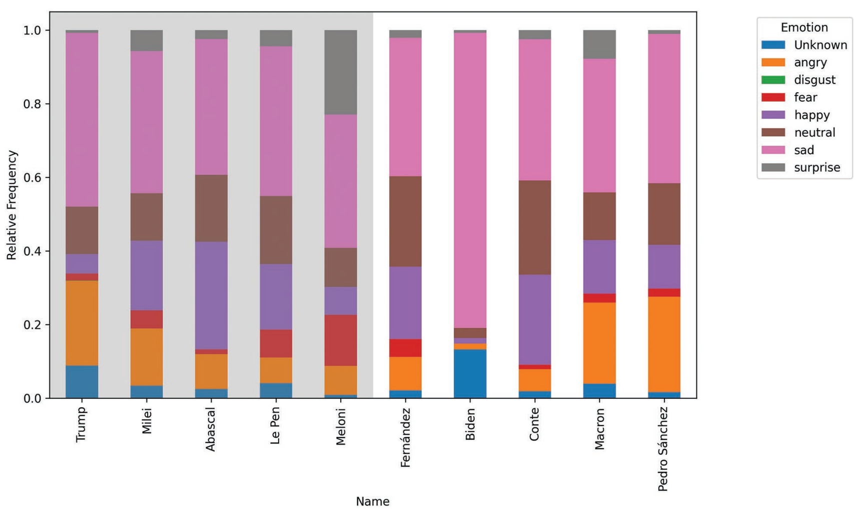

Figure 7.2 The sentiments assigned to each political figure, with a grouping based on populism (represented by the shaded area).

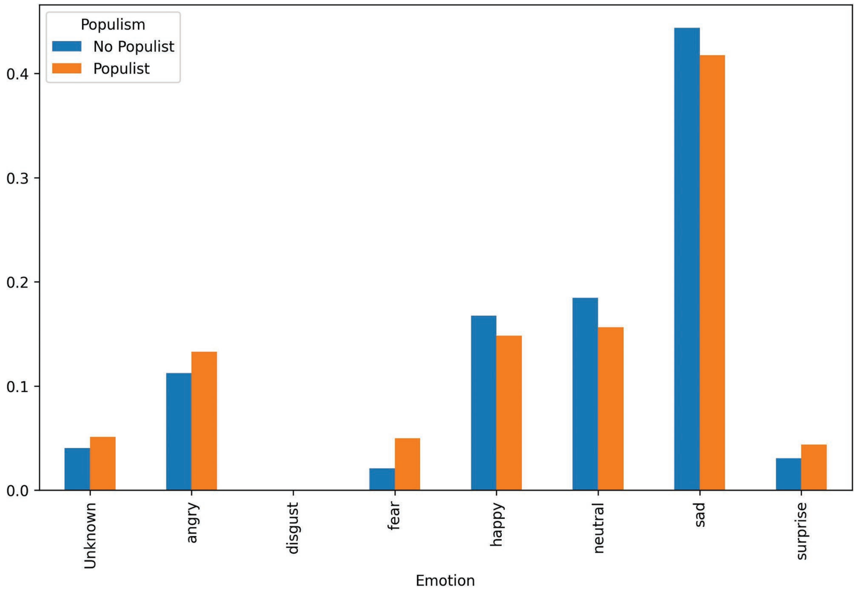

Figure 7.3 The result of such grouping in relative values.

Even so, we have endeavored to examine the relationships between the variables analyzed statistically. However, attempts to model the relationship between populism and emotions have proved to be unsuccessful. For example, logistic regression demonstrated a degree of success in predicting the presence of populism ($F_1 = .78$); nonetheless, it failed to predict the absence of populism in any instance. Consequently, we can conclude that, at least within the sample analyzed, this relationship does not appear to exist.

5. Discussion

The approaches outlined in this contribution primarily aim to expand the knowledge path paved by automated models. In this specific case, with the progress of language processing in disciplines such as humanities and social sciences, there is still a need for further research in using algorithmic models to analyze images. Conversely, this research also highlights the necessity for a broader scope, one that seamlessly integrates multidisciplinary variables into everyday practice. This starting point advocates for fields where human behavioral studies forge new drifts to extend knowledge further. Nevertheless, and in line with the preceding work, we find ourselves in the midst of a revolution that may not be all-encompassing. However, it contributes to the ongoing exploration of theories and reflections that, to date, lacked a sufficiently robust sample size to substantiate or challenge specific hypotheses. This stands as one of the central pillars for delving into this type of methodology. Reflections on topics such as the radicalization and polarization of society, as well as the rise and analysis of populist ideologies, serve as a basis for applying novel tools in an intricate social and media ecosystem. The prominence of visuals underscores the need for scientists to actively engage in the study of its massive impact (Russmann & Svensson, 2017). It is, therefore, imperative not only to continue developing these innovative computational models but also to reflect on how they can be advantageous, enhanced or effectively implemented.

Despite the systematization set out in this approach, it becomes evident that the data yield little significant value, especially when the objective is to identify distinctive patterns not merely among ideologies but also among leaders who employ different forms and techniques of expression. In such cases, certain traits might fit in with what we might anticipate from leaders who prioritize emotional appeal in their discourse, as is often seen in populist figures. Even so, as the results indicate, the differences are not as pronounced as initially expected. Thus, it is difficult, a priori, to pinpoint behaviors capable of unequivocally revealing an ideology or a particular style of politics.

However, the slight differences found elicit diverse explanations. On the one hand, we assume from the outset that both political leaders and political parties use their YouTube channels to disseminate specific, calculated information, constituting what can be termed "YouTubization of politics" (May, 2010). This approach also serves as an alternative medium for free riders or a mechanism to circumvent media authoritarianism in certain regimes (Litvinenko, 2021). While this assumption may hold true, the analyzed samples fail to show noticeable visual or non-verbal differences. Several factors may account for this, including the construction of the audiovisual narrative, which tends to be essential primarily during election campaigns, when electoral marketing takes center page, generating increased interest from supporters and the media. Hence, most of the published videos, with the exception of American leaders from both North and South America, predominantly consist of clips from rallies and parliamentary sessions. On another note, the composition of the sample itself also plays a significant role, impacting the results of data collection and analysis. Naturally, the present study only scratches the surface of possibilities for this type of analysis, revealing more questions than answers. The exploration of diverse communication styles in politics remains a viable avenue for further research, including an examination of whether different styles depend on the format (e.g. electoral debates, meetings, interviews, etc.). Furthermore, in order to target specific audiences, we could assign strategies, with election advertising serving as a control group aimed at the entire population.

This last factor reinforces the importance of embracing multidisciplinarity in our research. Such analysis affects variables that extend beyond traditional social science disciplines like sociology, psychology or political science. They necessitate the incorporation of insights and theories from the field of audiovisual communication, where a brief close-up shot can reveal distinct information compared to a long shot. Finally, we would like to highlight that automated models for image classification, like the ones discussed here, are beginning to become popular; and this, along with forthcoming contributions, should continue to raise thought-provoking questions that further our comprehension and expand our studies on topics such as populism or political polarization, and other complex social phenomena.

References

Altheide, D. L. (2020). Media logic and media psychology. In J. Bulck (Ed.), *The international encyclopedia of media psychology* (pp. 1–15). New York: Wiley.

Arias Maldonado, M. (2016). La digitalización de la conversación pública: Redes sociales, afectividad política y democracia. *Revista de Estudios Políticos*, *173*, 27–54.

Barbieri, F., Ballesteros, M., & Saggion, H. (2017). *Are emojis predictable?* arXiv. https://doi.org/10.48550/arXiv.1702.07285

Berners-Lee, T. (2000). *Tejiendo la red*. Madrid: Siglo XXI de España.

Cichocka, A., & Dhont, K. (2023). *The SAGE handbook of personality and individual differences: Applications of personality and individual differences* (pp. 323–351). Los Angeles, CA: SAGE Publications Ltd.

Debnath, A., Pinnaparaju, N., Shrivastava, M., Varma, V., & Augenstein, I. (2020). Semantic textual similarity of sentences with emojis. In *Companion Proceedings of the Web Conference 2020* (pp. 426–430). ACM, New York.

Durántez-Stolle, P., & Martínez-Sanz, R. (2019). Politainment in the transmedia construction of the image of politicians. *Communication & Society*, *32*(2), Article 2. https://doi.org/10.15581/003.32.37854

García-Marín, J., & Serrano-Contreras, I.-J. (2023). (Un)founded fear towards the algorithm: YouTube recommendations and polarisation. *Comunicar*, *31*(74). https://doi.org/10.3916/C74-2023-05

Gentzkow, M., Kelly, B., & Taddy, M. (2017). *Text as data*. Cambridge: National Bureau of Economic Research. https://doi.org/10.3386/w23276

Gerbaudo, P. (2018). Social media and populism: An elective affinity? *Media, Culture & Society*, *40*(5), 745–753.

Gómez-García, S., Zamora, R., & Berrocal, S. (2023). New frontiers for political communication in times of spectacularization. *Media and Communication*, *11*(2), 109–112.

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2022). *Text as data: A new framework for machine learning and the social sciences*. Princeton, NJ: Princeton University Press.

Ibrahim, H., AlDahoul, N., Lee, S., Rahwan, T., & Zaki, Y. (2023). YouTube's recommendation algorithm is left-leaning in the United States. *PNAS Nexus*, *2*(8). https://doi.org/10.1093/pnasnexus/pgad264

Joo, J., Bucy, E. P. (TTU), & Seidel, C. (2019). *Automated coding of televised leader displays: Detecting nonverbal political behavior with computer vision and deep learning*. https://ttu-ir.tdl.org/handle/2346/90282

Khaireddin, Y., & Chen, Z. (2021). *Facial emotion recognition: State of the art performance on FER2013*. arXiv. https://doi.org/10.48550/arXiv.2105.03588

Khoma, N., & Kozma, V. (2022). Spectacularization of political activism: Subject matter and social effect. *The Greek Review of Social Research*, 3–27.

Litvinenko, A. (2021). YouTube as alternative television in Russia: Political videos during the presidential election campaign 2018. *Social Media + Society*, *7*(1). https://doi.org/10.1177/2056305120984455

Lorenzo-Rodríguez, J., & Torcal, M. (2022). Twitter and affective polarisation: Following political leaders in Spain. *South European Society and Politics*, *27*(1), 97–123.

Luengo, Ó. G. (2014). Twitter vs medios tradicionales ¿Ha implicado Twitter un espacio ciudadano real de intercambio de información? In R. Cotarelo & J.-A. Olmeda (Eds.), *La democracia del siglo XXI. Política, medios de comunicación, internet y redes sociales* (pp. 409–428). Madrid: Estudios Políticos.

May, A. L. (2010). Who Tube? How YouTube's news and politics space is going mainstream. *The International Journal of Press/Politics*, *15*(4), 499–511.

Pelinka, A. (2013). Right-wing populism: Concept and typology. In R. Wodak, M. Khosravinik, & B. Mral (Eds.), *Right-wing populism in Europe* (pp. 3–22). London: Bloomsbury Academic.

Peng, Y., & Lu, Y. (2023). Computational visual analysis in political communication. In D. Lilleker & A. Veneti (Eds.), *Research handbook on visual politics* (pp. 42–54). Cheltenham: Edward Elgar Publishing.

Plazas-Olmedo, M., & López-Rabadán, P. (2023). Selfies and Speeches of a President at War: Volodymyr Zelensky's Strategy of Spectacularization on Instagram. *Media and Communication*, *11*(2), 188–202.

Rahyadi, I., Syahraihan, A. M., Agustina, D., & Irawan, D. A. (2023). Comparison of different political figures and evaluation of Twitter utilization over time in the political communication media. In *2023 17th international conference on ubiquitous information management and communication (IMCOM)* (pp. 1–7). https://doi.org/10.1109/IMCOM56909.2023.10035586

Russmann, U., & Svensson, J. (2017). Introduction to visual communication in the age of social media: Conceptual, theoretical and methodological challenges. *Media and Communication*, *5*(4), 1–5.

Schmøkel, R., & Bossetta, M. (2022). FBAdLibrarian and Pykognition: Open science tools for the collection and emotion detection of images in Facebook political ads with computer vision. *Journal of Information Technology & Politics*, *19*(1), 118–128.

Segerberg, A., & Bennett, W. L. (2011). Social media and the organization of collective action: Using Twitter to explore the ecologies of two climate change protests. *The Communication Review*, *14*(3), 197–215.

Serrano-Contreras, I.-J. (2021). *Medios de comunicación y polarización; un análisis del feminismo en España mediante procesamiento del lenguaje natural* [Universidad de Granada]. https://digibug.ugr.es/handle/10481/71409

Sterrett, D., Malato, D., Benz, J., Kantor, L., Tompson, T., Rosenstiel, T., Sonderman, J., & Loker, K. (2019). Who shared it?: Deciding what news to trust on social media. *Digital Journalism*, *7*(6), 783–801.

Tay, J., & Turner, G. (2008). What is television? Comparing media systems in the post-broadcast era. *Media International Australia*, *126*(1), 71–81.

Weismueller, J., Harrigan, P., Coussement, K., & Tessitore, T. (2022). What makes people share political content on social media? The role of emotion, authority and ideology. *Computers in Human Behavior*, *129*, 107150. https://doi.org/10.1016/j.chb.2021.107150

Zeng, J., & Abidin, C. (2021). '#OkBoomer, time to meet the Zoomers': Studying the memefication of intergenerational politics on TikTok. *Information, Communication & Society*, *24*(16), 2459–2481.