



The face of populism: examining differences in facial emotional expressions of political leaders using machine learning

Sara Major¹ · Aleksandar Tomašević¹

Received: 8 February 2025 / Accepted: 12 May 2025

© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2025

Abstract

Populist rhetoric employed in online media is characterized as deeply impassioned and most often imbued with strong emotions. This paper investigates the differences in affective non-verbal communication of political leaders. We use a deep-learning approach to process a sample of 220 YouTube videos depicting political leaders from 15 different countries, analyze their facial expressions of emotion, and then examine differences in average emotion scores representing the relative presence of six emotional states (anger, disgust, fear, happiness, sadness, and surprise) and a neutral expression for each frame of the processed YouTube video. Based on a sample of manually coded images, we find that this machine learning approach has 53–60% agreement with human annotation. We observe statistically significant differences in the average score of expressed negative emotions between groups of leaders with varying degrees of populist rhetoric. Overall, our contribution provides insight into the characteristics of non-verbal emotional expression among political leaders, as well as an open-source workflow for further computational studies of their affective communication.

Keywords Populism · Facial expressions · Emotion recognition · Machine learning · Political communication

✉ Aleksandar Tomašević
atomashevic@ff.uns.ac.rs
<https://scholar.google.com/citations?user=nPTApYAAAAJ>

Sara Major
sara.mejdzor@ff.uns.ac.rs
<https://scholar.google.com/citations?user=Cz21H5kAAAAJ>

¹ Department of Sociology, University of Novi Sad, Novi Sad, Serbia

1 Introduction

In the growing body of literature on populism, the category of emotion has most frequently been utilized in discrete ways to describe its emergence and tenacity [1]—by reference to the disillusioned silent majority, heightened emotions surrounding political leaders, and emotionally charged divisions between the people and the elite [2–4]. Compared to pluralist parties, who are commonly portrayed as more affectively neutral, populist rhetoric is most often characterized as deeply impassioned and imbued with strong (negative) emotions [5, 6]. Although the nature of the relationship between emotionality and populist discourse has rarely been subject to systematic empirical scrutiny [7], this link has become one of the main focal points of research on political communication in recent years. Comprehensive research on the matter is thus crucial for gaining insight into the extent to which populist narratives and performances are rooted in and shaped by emotion, and by extension, how emotion may be utilized as a tool for building and maintaining political support [8].

In the past, studies of populist communication have generally focused on its verbal elements, by analyzing the affective content of the words used in oral or written addresses of political leaders [7, 9–14]. While advantageous in its own way, this type of analysis can pose a challenge to researchers primarily due to language barriers and the fact that written text usually contains only a small amount of words that have an inherent affective connotation [7]. At the same time, currently available sentiment analysis tools are predominantly based on and developed for the English language, while the output for other languages is often of lower quality [15]. These circumstances therefore limit the possibility of result generalization and cross-cultural comparison.

Visual analyses, although insufficiently represented in the given field of research, may provide further insight into the characteristics of affective populist communication as they are not hindered by the same limitations as text-based studies [16] and allow for analysis of the visual rhetoric constructed and utilized by political leaders [17]. Within the scope of visual research on political emotion display, two main schemes of analysis have shown to be especially valuable: the ethological approach [18] and the Facial Action Coding System (FACS) [19].

Studies within the former largely focus on the analysis of political candidates' non-verbal representations in news coverage, their displays during political campaigns and debates, as well as the impact of different forms of representation on viewers and success in political elections [20–24]. The Facial Action Coding System, on the other hand, allows for a more in-depth analysis of the (micro) expressive behavior of political leaders during speeches and public performances. Due to its complexity, Ekman's classification system has been applied in a limited number of studies [25–27] but as the process of visual analysis gradually becomes automated [28] this research method is becoming more accessible to researchers from various fields of interest [29, 30].

Our study expands on the existing framework of FACS research on political emotion display and applies a computer vision approach for detecting emotions

from the facial expressions of political leaders. Because it is relatively new, this approach faces various methodological challenges which we aim to address by following the recommendations outlined by Barrett et al. [31], such as using big data techniques and automated detection in naturalistic settings, integrating traditional methods with machine learning, and sampling visual data across diverse contexts and cultures.

Since the introduction of the FER dataset [32], different machine learning algorithms have been developed for solving the challenge of facial expression recognition, which essentially requires performing two tasks: (1) detection of facial expression, and (2) emotion recognition from the detected facial expressions. We develop a reproducible and principled computational workflow in Python for the analysis of video content of political leaders' public performances. Our work relies on the existing Python library *fer* [33] which bundles a pre-trained convolutional neural network (CNN) Keras models for face detection [34] and emotion recognition [35, 36].

We find that populist leaders on average express more negative emotions and that the emotionally neutral facial expression is less frequently detected in their public appearances compared to non-populist leaders. Overall, our work points towards a new direction in computational studies of the affective aspects of populist political communication and provides a roadmap for overcoming the limitations of the current approach in future studies.

2 Populism as performance

Defining populism represents a challenging task for researchers as there is barely any consensus on what such a definition should (or should not) include, and the concept itself becomes increasingly muddled as the term is used to describe a growing number of phenomena [37]. Most conceptualizations of populism can, however, be classified [38] into one of three main approaches: (1) the ideational approach [39], (2) the discursive approach [4, 40] and (3) the political-strategic approach [41]. For the purposes of this study, we adopt the discursive approach, which conceives populism primarily as a rhetorical style based on first-order principles pertaining to *who* should rule—the people, as “the only legitimate source of political and moral authority in a democracy”—but which says very little about the *how* [4]. Whereas the ideational approach claims populism is a “thin-centered” ideology [39] and emphasizes the interpretation of politics as “a Manichean struggle between a reified will of the people and a conspiring elite” [42], the discursive approach highlights the ways in which this interpretation is framed and communicated as an anti-establishment discourse [43]. We find this to be the most comprehensible understanding of populism as it considers the operationalized elements of the populist communication logic [44] and examines the rhetorical components implicit in the ideological definition [45].

From this perspective, populist rhetoric has widely been characterized by the use of “dramatization, polarization, moralization, directness, ordinariness, colloquial and vulgar language” [44], and low-order appeals within which leaders “use a language that includes slang or folksy expressions and metaphors, are more

demonstrative in their bodily or facial expressions as well as in their demeanor, and display more raw, culturally popular tastes” [40]. Because it represents a form of symbolic resistance to the more formalized and composed self-representation of mainstream pluralist leaders, i.e., the “establishment” and “political elite”, the populist performance¹ is subversive to the normative communication structures of political dialogue. As such, it appeals to “common sense” [20] and claims to give a voice to “the ordinary people” [46].

This mode of communication has shown to be highly effective as a means of garnering media attention and cultivating a dedicated base of supporters—especially in the new media environment—and populist leaders seem to be overall more adept at capitalizing on it than other, non-populist political figures [9, 47]. Since digital media hinges on the attention economy [48] and thus tends to favor “dark participation” [49], populist leaders and political parties may fare well online due to the transgressive nature of their self-representation and the simplification, emotionalization, and negativity inherent in populist messaging [20, 44]. Social media is a particularly convenient tool for the proliferation of populist messages and imagery as it allows political actors to circumvent the mainstream media, reach large audiences, and communicate and interact directly with “the people” [50, 51]. At the same time, online platforms allow for an instant audience response, as the “feedback loop between producer and consumer is much more intense and rapid than with television and radio” [52]. This form of “mediatized populism” [53] allows for a specific type of political performance, one which is largely unmitigated and relies heavily on portraying political leaders favorably to (online) audiences by framing them as expressive, dominant, and disruptive figures who are in tune with “the will of the people”.

In recent years, much attention has been paid to the strategies of (affective) populist communication online, especially through analyses of verbal content shared by populist leaders and parties on Twitter and Facebook [e.g., 54, 55, 13, 14, 7]. Given the rise of visual communication in politics and the rapid developments of Web 2.0 practices [23, 49, 56], our primary aim was to expand the scope of research on the affective self-representation of populist leaders online by examining the characteristics of their non-verbal emotional communication. The facial expressions of political leaders are a particularly valuable area of research in this context, since digital media provide a breadth of material for analysis—well-lit, high-quality close-up shots of political leaders are now more prevalent than ever. We used this abundance of available data sources to our advantage and processed 220 YouTube videos uploaded to official accounts of political parties and their leaders, amounting to a total of 77 h of video material. Our analysis examines potential differences in facial expressions between populist and non-populist political leaders during public performances, and more specifically, the characteristics of their negative emotional expressions and neutrality in emotional stance (i.e., absence of emotional expression).

¹ By “the populist performance” we mean the populist communication style as laid out by Bucy et al. [20]—that is, the verbal, tonal, and non-verbal display of a political leader which features three main communicative dimensions: simplification, emotionalization, and negativity.

3 Data & method

3.1 Party selection

To create our sample, we first had to differentiate and categorize political parties (and, by extension, their leading figures) as either populist or pluralist. For this purpose, we used data from the Global Party Survey (GPS) [57], an expert survey from 2019 which offers an overview of key ideological and issue positions for 1043 parties from 163 countries worldwide. While expert surveys certainly have their limitations [58], we believe this database remains the most comprehensive comparative overview of political parties globally.

Among other data, such as ideological framework and party size, the survey also offers an estimate of the degree to which existing parties employ populist rhetoric. In the questionnaire design, populist rhetoric is defined as “language which typically challenges the legitimacy of established political institutions and emphasizes that the will of the people should prevail” [57], whereas pluralist rhetoric “rejects these ideas, believing that elected leaders should govern, constrained by minority rights, bargaining and compromise, as well as checks and balances on executive power” [57]. The populist rhetoric scale (V8 in the GPS dataset) is based on expert assessments of rhetorical style given these parameters. The variable *Type_Populism* categorizes this scale into four ordinal groups: strongly pluralist, moderately pluralist, moderately populist, and strongly populist. Although these categories may not be able to fully capture the nuances of populist rhetoric—its culturally specific manifestations, or what “the will of the people” may connote for people across different political contexts—they allow insight into the fundamental characteristics of party rhetoric and enable fair comparison between countries.

Given that a great concern in expert surveys is the possibility of experts’ backgrounds introducing cognitive bias into their assessments [59], it is also important to note that internal validity testing done within the GPS showed that the personal characteristics of experts (such as ideological stance, gender, and nationality) were not significant predictors when it comes to the placement of parties on the Pluralism-Populism scale [57].² Additionally, to establish that differences in expression of emotion are indeed more of a matter of populist leaning rather than ideological preference, we use the “Type_Values” variable in the supplementary analysis to examine ideological differences in emotion expression (see the Supplement section E – Ideological differences).

From the GPS dataset, fifteen countries with competitive popular elections were chosen: Argentina, Australia, Brazil, Croatia, France, Hungary, India, Italy, Japan, the Netherlands, Serbia, South Africa, Turkey, the United Kingdom, and the United States.³ This sample, although limited, provides somewhat broad coverage of

² For external validity measures, see robustness tests in Norris [57].

³ The average number of experts per country in the selected dataset was approximately 27, with the highest number being in the United States (50) and the lowest in Serbia (6). While this number is higher than the average for the GPS as a whole—19 [57]—the variance between countries could potentially entail variance in overall quality of assessment.

various cultures and political systems globally. Three parties with varying degrees of populist rhetoric—strongly populist, strongly pluralist, and either moderately populist or moderately pluralist—were selected from each country. We primarily focused on minor and major political parties as defined by the GPS variable categorizing parties based on their share of the vote in contests for the lower house of the national parliament. Fringe parties (less than 3% of the vote) were generally excluded from the sample since their audience reach and influence is considerably less than that of their larger counterparts competing at the same scale in their respective countries. Likewise, due to this criteria, only two parties from the US were included in the sample. The table from Supplement D – Political leaders contains information about the selected parties and their leaders.

3.2 Video selection

Once party selection was complete, we identified one leader per party who has held a prominent political position over the past decade (up to 2019, covered by the GPS), and found their official YouTube channel or their party's official YouTube channel. From these channels, five videos were selected for each leader based on the following criteria:

- (1) *View count.* Videos that reached a larger audience were given priority over videos that did not perform comparatively as well on the platform. View count was used as a filtering mechanism to single out videos with large and potentially diverse audiences, because the most viewed videos on the official YouTube channels of major political parties usually refer to topics of wider public interest. This is partially due to the “popularity bias” on YouTube, meaning it has a tendency to recommend political content that aligns with the interests of the majority [60].
- (2) *Unmediated representation.* Only direct and unmediated representations of the given political leaders were included, i.e., videos which were not in any way previously modified by third parties such as news outlets. These consisted mainly of speeches, press conferences, ads, and promotional videos. This was done to ensure that the given results would reflect the leaders' self-representation, rather than their representation in news coverage.
- (3) *Consistent framing.* A preference was given to videos where the face of the leader was in focus for an extended period of time and which did not frequently switch between disparate camera angles. This allowed for an overall consistent frame of reference for each given video.
- (4) *Video quality.* High-quality, well-lit videos were selected whenever possible so as to improve analysis accuracy.

This selection process resulted in a sample of 220 videos, representing 44 political leaders from 15 different countries. The sample is not gender-balanced as only five female leaders are present, which reflects the global underrepresentation of women in leadership positions within major political parties. As a result, our analysis does not account for gender, which we acknowledge as a limitation of our approach.

Additionally, because the sample is limited by the timeframe of the currently available GPS dataset, it hasn't captured more recent global developments on the populist-pluralist spectrum (e.g., Giorgia Meloni and the Brothers of Italy movement and the Servant of the People party in Ukraine are not present in the dataset). Given these limitations, our sample is conceived as a diverse snapshot of the state of populist and pluralist rhetoric in the second decade of the 21st century and does not capture the full scope or evolution of populism today.

3.3 Video processing and analysis

Each video in the sample was processed by extracting 300 frames uniformly distributed over the length of the video, which were the basis for our main analysis. To improve the robustness of the overall approach, we also included an alternative frame processing approach in the supplementary material, based on processing every 50th frame of the video. Given that the majority of the YouTube videos have 24–30 frames per second, this is equivalent to selecting a frame every 1.67–2.00 s, which is a reasonable time window if we expect to capture potential changes in facial expression. The average number of frames per video extracted using this approach was 273.03, which makes the resulting time-series of emotion scores comparable in length with the first approach.

For each selected frame of every video in the sample, the following procedure is applied. First, face detection is applied to the extracted frames using a three-stage cascaded CNN architecture with default parameters (scale factor: 0.709, minimum face size: 40 pixels, thresholds: [0.6, 0.7, 0.7] for the three detection stages) [34]. This approach is particularly suitable for detecting faces in videos because it performs simultaneous face detection and alignment, which is important as we cannot expect that political leaders will be looking straight at the camera in each extracted video frame. After face detection was applied to all frames, there were videos in which multiple faces were detected throughout the frame (when a second face appeared in >5% of frames). These videos were manually checked to confirm which of the detected faces corresponds to the selected political leader, and in some cases, specific frame ranges were specified to use only portions with consistent face detection.

The second step is emotion detection within the verified detected face box. We relied on implementing the mini-Xception architecture of CNN [35] aimed at detecting six basic facial expressions of emotion, as well as a neutral expression. This model uses four residual depth-wise separable convolutions with input shape of 64x64 pixels (grayscale) and outputs seven emotion scores. We initialized the detector with MTCNN enabled using the `FER(mtcnn=True)` constructor from the *fer* library. Similar to the face detection approach outlined above, this implementation is aimed at real-time emotion detection from video input for robotic applications. This model achieves 66% accuracy [see 61, arriaga2017Realtimea] on the FER2013

dataset [32] which is a standard benchmark for emotion detection. This dataset consists of 30,000 small images (48x48 pixels) showing diverse facial expressions of different emotions.⁴ Each emotional expression has approximately 6000 images, except disgust, which is only present in 600 images. According to Goodfellow et al. [32], human classification accuracy is in the range of $65\% \pm 5\%$, but no additional information about the sample size, number of coders, or detection procedure is given.

In order to provide additional information and to validate the output of the machine learning approach, we conducted a small image annotation study where five coders annotated a sample of 514 randomly selected frames extracted from 220 videos. Coders received basic training in emotion recognition and, due to limited resources, answered only two questions about every image: (1) Can you clearly see the face of a politician in this image? (Yes/No), (2) Is this face expressing positive emotion, negative emotion, or is it a neutral expression? (“Positive emotion”, “Negative emotion”, “Neutral expression”). The majority vote label was set as the “human-coded label” and compared with the output of the machine learning algorithm, whose label is defined based on the highest score out of the sum of positive emotions, negative emotions, or the neutrality score. A detailed description and the results of the data annotation study can be found in Supplement F – Image annotation.

The result of the emotion detection procedure on a single video frame is an array of seven emotion scores, each representing the predicted proportion of an expressed emotion. Each score is matched with one of seven labels: “angry”, “disgust”, “fear”, “happy”, “sad”, “surprise” and “neutral”. The lowest possible score indicating the absence of a particular emotion is 0, while the highest score, which indicates a “clean” expression of a single emotion, is 1. In a statistical sense, the resulting data are compositional as they do not carry absolute, but rather relative information about the expressed emotions. This data is organized as a data frame consisting of seven columns and approximately 300 rows, storing emotional scores for each selected frame of the video.

Because we were primarily interested in neutral emotional expressions and expressions of negative emotions, we created an additional variable which sums the scores of negative emotions: anger, fear, disgust, and sadness. Figure 1 shows examples of faces with a high detected score of negative emotion and a high score of neutral expression.

The entire analysis was performed in Python and the code is publicly available in the GitHub repository: <https://github.com/atomashevic/face-of-populism>. The

⁴ FER2013 was constructed using Google Images API queries for queries for emotion-related keywords like “blissful” and “enraged”, combined with queries introducing diversity in gender, age and ethnicity represented in the images. Images were then labeled by human coders [32, pp. 3–4] Unfortunately, no information is provided by the authors of the dataset regarding the number of human coders nor their background, or the labeling methodology applied to extract training labels for every image included in the FER2013 dataset. Benchmarks based on the FER2013 dataset should be taken with caution, especially in social science applications, because the dataset construction and validation documentation lacks crucial information about inter-coder reliability, coding procedures, and potential biases in the image selection process.



Fig. 1 Example frames showing a high detected score of negative emotions (left: angry 0.78, disgust 0.00, fear 0.09, happy 0.00, sad 0.07, surprise 0.00, neutral 0.04) and a high detected score of neutral expression (right: angry 0.07, disgust 0.00, fear 0.04, happy 0.00, sad 0.02, surprise 0.08, neutral 0.79)

code is organized into three sections. The data processing part of the code can be used to reproduce emotion score data files from the list of YouTube URLs for the videos included in the sample. It can easily be modified to expand the sample to a larger number of videos. Face and emotion detection relies heavily on the *fer* Python library and that part of the code is presented in the Jupyter notebook. Finally, we have data analysis scripts for performing data wrangling and statistical tests as well as producing Figs. 2, 3, and 4.

3.4 Data processing pipeline

Our complete data processing pipeline consists of several sequential steps:

1. **Video acquisition:** Videos were downloaded from YouTube using the *pytube* library.
2. **Frame extraction:** Using the OpenCV (*cv2*) library, we extracted video meta-data (fps, frame count, duration) and selected frames according to our sampling method.
3. **Face detection:** The MTCNN model implemented in the *fer* library detected and aligned faces in each frame.
4. **Face verification:** Videos with multiple detected faces were manually reviewed to identify the political leader.
5. **Emotion recognition:** The mini-Xception CNN model processed each detected face to output seven emotion scores.
6. **Data aggregation:** For each video, we calculated mean scores for each emotion category and the composite negative emotion score.

7. **Statistical analysis:** We performed t-tests, ANOVA with post-hoc Tukey HSD tests, and calculated effect sizes to examine differences between populist and pluralist leaders.

All code for this pipeline is organized in a publicly available repository with a structured organization where raw data is processed through these steps, with intermediate and final results stored in appropriate directories. Videos requiring specific handling (e.g., cropping to specific time ranges) were processed using the *moviepy* library before entering the main pipeline.

4 Results

In the entire sample, there are 203 successfully processed videos using the approach of selecting 300 frames uniformly distributed across the entire length of the video clip. For 17 videos, multiple faces were detected without consistent classification of each person as a unique face, and these videos were discarded from the analysis. Pluralist leaders are present in 96 videos (31 strongly pluralist and 65 moderately pluralist) and populist leaders are present in the remaining 107 videos (69 strongly populist and 38 moderately populist). For every video, the emotion recognition procedure returns the score of the six detected emotions (anger, disgust, fear, happiness, sadness, surprise) as well as the score of the neutral expression for each of its frames. These scores sum up to 1 and they can be viewed as proportions of each emotion present in the compound facial expression detected in the frame.

Human image annotation was performed on a sample of 514 images extracted from 220 videos. Agreement between human coders ranged from 30.3% to 69.8%, while Cohen's Kappa coefficient varied from 0.15 to 0.462. Overall, Fleiss' Kappa is 0.231, which suggests fair agreement between coders but is also indicative of how difficult the task of recognizing the emotional expression in a single, randomly selected video frame is. When only selecting cases where a majority vote label exists (458 images), agreement between that label and the result of the machine learning classification is 55.2%, with a Cohen's Kappa of 0.311 and F1 score of 0.502. This suggests moderate agreement and performance of the machine learning model relative to human consensus. When only selecting cases where the majority of coders and the model agree that a face is clearly visible (present) in the image (356 cases), agreement is 59.8% (Kappa 0.320, F1 score 0.530).

For each video, we calculated the mean score of each emotional state for all frames, as well as the mean of neutral expression, and the mean of negative emotional expression for the entire video. Descriptive statistics of these values are presented in Table 1.

Results suggest that negative emotions are more frequently expressed in the entire sample, with anger (0.227) as the emotion having the highest average score, followed by sadness (0.206) and fear (0.127). As a consequence, the average proportion of negative emotions is greater than 0.5, which means that on average negative emotions are detected more frequently than positive emotions (e.g., happiness) and neutral expressions combined. There are clear issues with the detection of disgust

Table 1 Descriptive statistics for extracted emotions for the entire sample, $N = 203$ videos

	Mean	SD	Min	Median	Max
Anger	0.227	0.134	0.034	0.200	0.726
Disgust	0.003	0.008	0.000	0.001	0.066
Fear	0.127	0.091	0.014	0.099	0.548
Happiness	0.077	0.083	0.000	0.051	0.414
Sadness	0.206	0.110	0.023	0.181	0.568
Surprise	0.031	0.037	0.000	0.020	0.273
Neutral	0.328	0.169	0.020	0.300	0.807
Negative emotions	0.563	0.185	0.156	0.570	0.964

and surprise, where a low maximum level suggests low positive predictive value for these emotions. However, our focus is on differences in the expression of negative emotions and in scores of neutral states. Aggregating negative emotions enables us to surpass the false positive issues within the group of negative emotions [e.g., expression of fear being classified as sadness, disgust as anger, 61] and minimizes the impact of the low positive predictive value of disgust and surprise.

Looking at the maximum values of average detected scores, the largest average scores of specific negative emotions were found in videos of populist leaders. The highest proportion of anger was detected in a video of Andrej Plenković (HDZ, Croatia), disgust in a video of Narendra Modi (BJP, India), fear in a video of Saša Radulović (DJB, Serbia) and sadness in a video of Donald Trump (GOP, USA). The video with the highest aggregate proportion of negative emotions, where one of the negative emotions is dominant in almost every analyzed frame, is the video titled “Special Message from President Trump”, published on December 23, 2020. On the other hand, the highest average score of neutral expression is found in a video of Yukio Edano (CDP, Japan), whose party belongs, in contrast with previous examples, to the moderately pluralist group.⁵

To further examine differences between pluralist and populist leaders, we used the GPS measure of populist rhetoric in a binary way to split the sample into groups of pluralist (*Party Populism* 1-2) and populist (*Party Populism* 3-4) leaders. The distribution of the average scores of negative emotion and neutral expression is shown in Fig. 2.

Overall, all average values have a large spread in both groups, with pluralist leaders having several videos with less than a 0.2 average negative emotion score and more than a 0.75 average score of neutral expression. In terms of central tendency, we see denser grouping of pluralist leaders between the 0.4 and 0.6 negative emotions score, while populists tend to cluster in areas between 0.6 and 0.8. In terms

⁵ This does not imply that, on average, there is a tendency for Japanese leaders to be more neutral. In fact, in Japan, we observe the same tendency as seen in the entire sample—the speeches of Shinzo Abe (LDP) from the populist group have the highest average score of negative emotions (0.58), while Natsuo Yamaguchi (NKP) from the moderately pluralist group has the lowest average score (0.30). Conversely, the average scores of neutrality show the opposite trend. These results are not extreme compared to the sample distributions shown in Fig. 2.

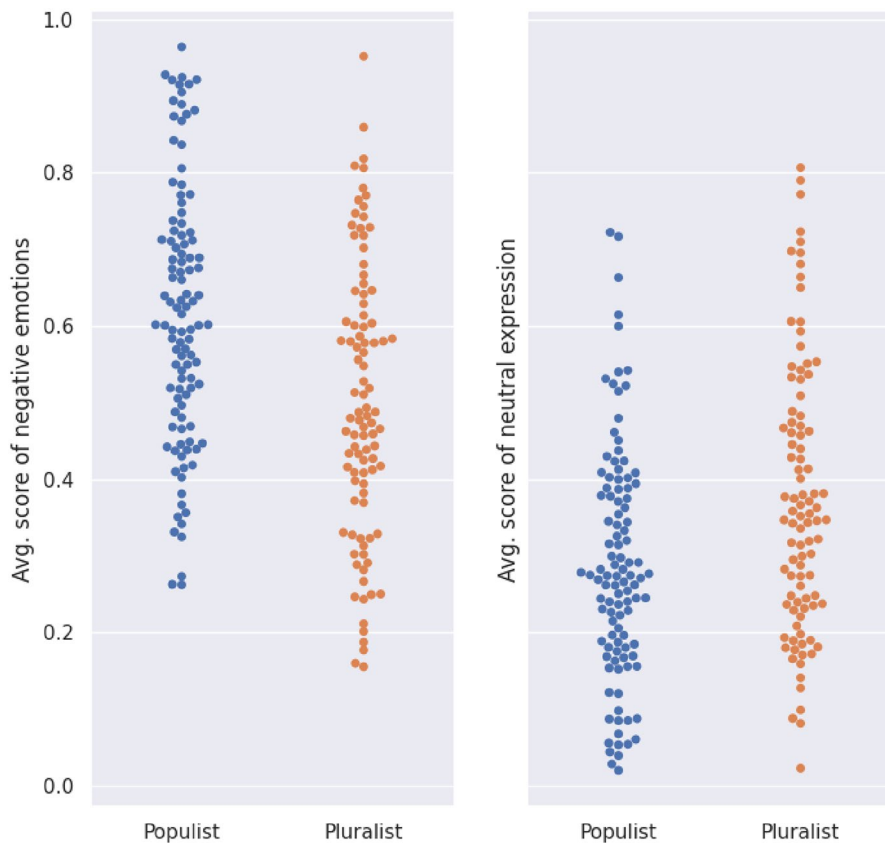


Fig. 2 Distribution of average scores of negative emotions (left) and neutral expression (right) in videos featuring populist and pluralist leaders

of neutral expression, we have a thicker cluster of points for populists in the range between 0.2 and 0.4. Figure 1 suggests that, for populist leaders, we are more likely to observe lower average scores of neutral expression and higher average scores of negative emotion. The reverse tendency can be observed in the case of pluralist leaders. Negative emotions are dominant in 73.8% of populist videos, compared to 45.8% of pluralist videos.

The mean average score of negative emotions in the populist group is 0.616 (95% CI 0.584–0.649), compared to the 0.500 mean average score for pluralists (95% CI 0.464–0.537). The difference between the two groups is statistically significant ($t = 4.691$, $p = 0.000005$, $d = 0.66$ moderate effect size). This finding suggests that pluralists generally tend to have an average score of negative emotions around 0.5, which roughly translates to the situation where negative emotion is the prevailing expression in one half of the video (in fact, that percentage is 46.1% compared to 66.7% for populists; see table in Supplement B – Dominant emotions per frame). On

the other hand, for populists, more than one half of each video where they appear is dominated by negative emotions.

The mean average score of neutral expression for populists is 0.289 (95% CI 0.260–0.318), compared to the 0.374 mean for pluralists (95% CI 0.338–0.409). While this difference in means is statistically significant ($t = -3.636$, $p = 0.00035$, $d = -0.515$ moderate effect size), it is smaller than the one observed in the case of negative emotions.

To investigate the robustness of these differences, we took into account the differences between moderate and strong pluralists/populists provided by the GPS. The distributions of average scores in these four groups are compared and shown in Figs. 3 and 4.

In the case of negative emotions (Fig. 3), we observe four flat distributions with thick tails. However, the median point of all distributions progressively moves towards more negative emotions as we move from strong pluralists to strong populists. What is apparent in these differences is the low frequency of videos featuring populist leaders where the average score of negative emotions is below 0.5. It is worth pointing out that examples of extremely high scores of negative emotions are present also in pluralist groups.

Similarly, in the case of neutral expression (Fig. 4), we have flat distributions (except in the case of moderate populists) with long right tails containing videos of leaders who maintain a neutral expression throughout most of their recorded appearance. Again, median values move progressively towards lower neutrality as we move from pluralists towards populists, except that no large difference can be seen when comparing strong and moderate populists. In the case of strong populists, however,

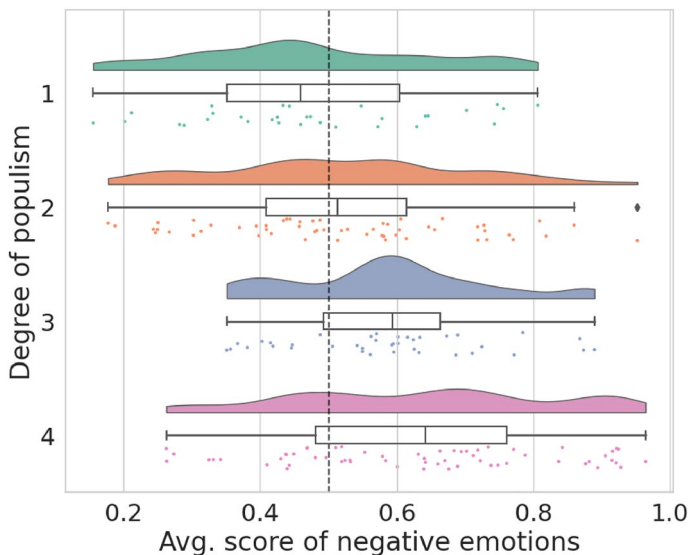


Fig. 3 Raincloud plot showing the distribution of average scores of negative emotions across four degrees of populism. Every data point is presented as a colored dot and above them boxplots and density plots are displayed. The dashed vertical bar indicates the score of 0.5

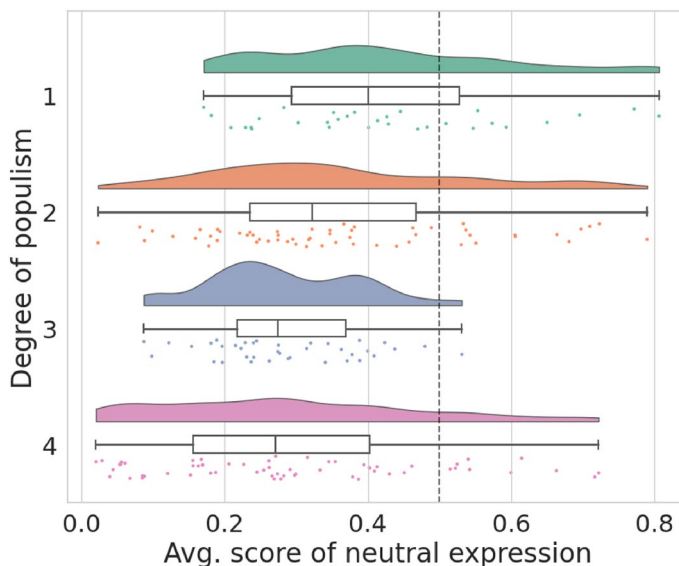


Fig. 4 Raincloud plot showing the distribution of average scores of neutral expression across four degrees of populism. Every data point is presented as a colored dot and above them boxplots and density plots are displayed. The dashed vertical bar indicates the score of 0.5

we have a larger number of videos where the neutral expression is almost completely absent.

Overall, the results of one-way ANOVA suggest that the differences between the group means are statistically significant both in the case of negative emotions ($F = 7.109$ $p = 0.000152$), with strong effect size measured by eta-squared (0.104), and neutral expression ($F = 5.625$ $p = 0.001038$), with moderate effect size (0.084). Results from Tukey HSD post-hoc tests reveal significant differences between leaders coming from parties employing strong populist rhetoric and leaders coming from pluralist parties (mean difference 0.1559 $p = 0.0005$ in comparison with strong pluralists, and mean difference 0.1148 $p = 0.0029$ in comparison with moderate pluralists). For neutral expression we find the opposite case, as strong pluralists differ from both groups of populists (mean difference 0.135, $p = 0.0056$ for moderate populists, and mean difference 0.129 $p = 0.0032$ for strong populists).

Overall, these results reveal differences in the average scores of negative emotion and neutrality between populist and pluralist leaders on a modest sample. These differences are more pronounced when we focus on candidates coming from parties who strongly rely on either pluralist or populist rhetoric, with strong pluralists on average favoring more neutral expressions, while strong populists more frequently express negative emotions—meaning that, for the majority of their public appearance, we can observe facial expressions of negative emotions.

5 Discussion

The present study examines facial emotional expressions of populist and non-populist political leaders across fifteen countries using a deep-learning-based computer-vision algorithm to investigate differences in their expressions of negative emotions (anger, disgust, fear, and sadness) and neutral stances, i.e., facial expressions which lack emotional disposition. Our findings indicate that populist leaders tend to express negative emotions more frequently than non-populist leaders, and conversely, that neutral expressions are less common in their public performances than they are in those of more conventional, pluralist political figures. The significance and the effect size of these differences are higher compared to differences between ideological groups (see Supplement E – Ideological differences). These findings align with the theoretical framework on the characteristics of affective populist communication [5, 6, 44], as well as with previous, mostly text-based research which indicates that populist leaders employ language which is associated with negative emotion at a higher rate than mainstream, pluralist politicians [7, 10, 14, 62, 63].

Overall, the results provide new evidence of a style of affective, non-verbal communication more common to political leaders whose rhetoric is populist-leaning, and thus further substantiate the idea that populism is performative in nature [16, 20]. As such, because it is subversive in relation to established political conventions, the populist performance is especially well-suited for the environment of social media where the principle of “any press is good press” often applies—that is, the quantity of audience attention bears more weight than its quality [48]. On one hand, the populist use of negative emotional displays in videos may therefore be a mechanism through which they maintain relevance, considering that controversial content online ordinarily generates more engagement among consumers [49]. On the other, populists’ facial expressions can provide affective heuristics for audiences, influencing their evaluation of candidates’ political competence [64] and allowing them to quickly reduce complex political issues into simple, emotional interpretations by visually framing them as negative.

Since this emerging dynamic has the potential to influence the behavior of voters—and perhaps even politicians themselves⁶—it holds several implications for populism, political communication, and polarization research. The Manichaen mentality of “Us versus Them” is fundamental to the populist worldview, and it further entrenches ideological positions and social divides. It is also largely dependent on fostering contentious discourse practices and negatively portraying political opponents and their beliefs. Thus, understanding how and when populist leaders bring negative emotional appeals to the forefront is vital for developing a more refined model of populist communication that would consider both the cognitive and affective sides of their political behavior. Studying non-verbal cues, in particular, can

⁶ Due to the unmediated relationship between political figure and voter, it may very well be possible for online audiences to directly influence politicians in return. For example, once videos prominently featuring negative emotions start performing well, they might create a feedback loop which, in the long run, incentivizes political figures to “go negative” when they may have otherwise taken a different communication approach.

provide a breadth of information for this line of research since facial expressions of emotion add a layer of meaning to political messaging, establishing credibility and trust (or lack thereof) in political figures and appealing to audiences on a more intuitive perceptive level.

Apart from these findings, the present study introduces an open-source workflow for the analysis of political videos and provides a new open dataset of time-series emotion scores, as well as tools for the annotation of images in R and a new dataset of human labelled images which can be used to benchmark different machine learning approaches to facial expression recognition. Overall, these contributions aim to make large-scale computational studies of facial expressions of emotion more accessible and rigorous in the future.

Given the exploratory nature of the study and its reliance on machine learning algorithms for emotional detection, there are several notable limitations with significant impact on the results presented in this paper. There is a group of limitations inherent to the way we constructed the sample, and another related to the machine learning approach to detecting emotions from facial expressions.

Firstly, the sample of YouTube videos is not globally representative and does not provide ample coverage of the global political spectrum in regard to the use of populist or pluralistic rhetoric. As we have shown in the supplementary analysis (see Supplement C – Country-level differences), the differences between populist and pluralist leaders are not consistent in all investigated countries. Larger research infrastructure is required to process a more substantial number of videos, leaders, and countries, and the current computational toolset needs to be extended to enable seamless automatic processing of videos. This also relates back to the gender imbalance of the current sample. While it is reasonable to hypothesize that gender influences emotional expression among political leaders [65], in this case, including more women without significantly increasing the overall size of the sample and the number of leaders would mean focusing only on countries where there is at least one female leader of a non-fringe political party. This creates the risk of introducing bias towards Western countries in the sample and disregarding the global scope.

Secondly, a key component of our analysis is the measure of populist rhetoric supplied by the Global Party Survey database [57] which significantly limits the timeframe of the analysis (no leaders and parties after 2019 were taken into consideration), but also imposes the inherent limits of an expert survey. While the GPS is the only available source of multi-dimensional measurements of populist rhetoric on a global scale, the experts surveyed may operate within different understandings of what populist rhetoric is. This is particularly challenging in developing countries, where fewer experts answered the survey, as well as in autocratic regimes where experts may not have been able to provide a fair assessment of the (ruling) political parties (or refused to do so). By focusing on large countries with free elections, we tried to mitigate some of the risks, but there is the possibility that some measurements used in this study reflect expert biases, especially since we used a single, aggregate measure of populist rhetoric. Future studies should aim for more granularity and robustness in the measurement of populist rhetoric, which may be obtained by following a text-as-data approach and obtaining the measures using

computational and machine learning tools, where expert survey scores provide external validation [66].

Furthermore, since we lacked the resources necessary to process every political speech from each party's official YouTube channel, the first criterion of selection is the view count of the videos. Our expectation is that this metric serves as a proxy for the relevance of the speech content to a wider audience over a longer period of time. However, since we have no information about the audience, it is possible that we are observing a demand-side effect, where audiences who consume populist videos might show a stronger preference for those that prominently feature negative emotions, introducing a specific bias in the sample.

This consideration may apply even more broadly, as our analysis does not fundamentally account for systemic bias in media consumption, coverage, and narrative construction, and consequently the differences in emotional expression may not be a result of varying political styles but rather of bias in media selection. Future research should address this issue by comparing emotional expressions across a wider range of media formats—including other social media platforms, television broadcasts of speeches and debates, mediated interviews, etc.—to determine whether the observed patterns persist beyond YouTube's algorithm and environment.

The next major limitation is a result of the machine learning model deployed for emotion recognition. We conducted a small validation study to compare the performance of our model with the consensus of five human coders. Although results show evidence of moderate agreement between humans and the machine learning model (53–60%), there is high inconsistency between human coders even for the simple task of assessing whether a given facial expression shows positive emotion, negative emotion, or a neutral expression. This level of agreement, while adequate for our exploratory study, indicates substantial uncertainty in the classification of emotional expressions that could potentially affect our conclusions. The higher detection rate for negative emotions compared to positive ones could potentially amplify differences between groups if the model is more sensitive to certain types of expressions.

There is an urgent need for large-scale image annotation studies with expert coders if we are to assess the performance and accuracy of FER models in the future, especially on difficult “in the wild” image datasets such as ours. However, it is worth noting that, in our main analysis, we focus on averages across 300 frames of each video, making our analysis robust to misclassification of a smaller number of individual frames. This methodological choice helps mitigate the impact of individual frame misclassifications on our overall findings. That being said, we believe that fair agreement between the human majority vote and machine learning model does not compromise the validity of our findings.

For future research, several technical improvements could enhance the accuracy of facial emotion recognition in political contexts. The most promising direction appears to be zero-shot classification based on open-source transformer models [67] such as OpenAI's CLIP (Contrastive Language-Image Pre-training) [68], which may offer more nuanced performance with the ability to process facial expressions in the specific context of the image along with other visual cues. These models can leverage large-scale pre-training on diverse datasets and potentially provide more culturally sensitive interpretations of facial expressions [69]. Additionally, fine-tuning

pre-trained models specifically on datasets of political figures' expressions could improve domain-specific accuracy, as could ensemble methods that combine multiple classification approaches to enhance robustness.

Overall confidence in the presented results would surely be higher if we based our analysis on a computer-vision approach which was shown to have higher accuracy, at least on FER2013 data. One of the possible alternatives would be to follow recent work [30, 70] and use proprietary API such as Amazon's Rekognition. Nonetheless, the aim of this paper was to introduce an open-source approach focusing on free computing resources and the deployment of machine learning algorithms on systems that are accessible to individual researchers in social sciences, including graduate students without significant institutional support for computational work.

This study was focused on exploring basic mean-level differences between the given groups, but it is worth noting that the full output of our emotion detection workflow includes time series data suitable for more advanced analyses. For instance, multivariate analysis could explore affective dynamics [71], identify differences related to various factors (such as type of public address or proximity of elections), and cluster political leaders according to temporal patterns of emotional expression. Furthermore, multivariate change point detection methods [72] could identify key moments where emotional expressions significantly change. Future research could extend this approach in other promising directions as well. Network psychometric methods could model dynamic interrelationships between emotional expressions, potentially revealing different temporal patterns between populist and pluralist leaders [73]. Recent speech-to-text models like Whisper [74] could enable integrated analyses of verbal and non-verbal emotional expressions, providing insights into the alignment between spoken content and facial displays. Additionally, simulation approaches could help test theoretical predictions about emotion dynamics in political communication [75]. Collectively, these approaches would move beyond static comparisons toward analyses of temporal dynamics, potentially revealing how leaders modulate emotional displays in response to audience feedback or political contexts.

As computational studies of political communication rapidly embrace machine learning tools and algorithms, this study offers an extension of an open-source toolbox for the analysis of affective non-verbal political communication. Analyses presented in this paper are based on a novel dataset which is only a starting point for further studies, and which will hopefully expand in the future to encompass a more diverse sample of political leaders in videos obtained from different platforms. Gaining this kind of insight into the visual aspects of emotional political communication, coupled with the existing analyses of its verbal elements, holds the potential for a more comprehensive understanding of how leaders convey political messages, persuade audiences, and present themselves to the public using digital media platforms.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s42001-025-00392-w>.

Acknowledgements The authors did not receive any financial support for this research. The annotation of images was made possible thanks to the dedicated work and effort of our graduate students, A. Nađ, M. Milanko, and A. Mišić, who volunteered as coders.

Data availability Data files are stored as CSV files. Processed data used for both the main and supplementary analyses can be found at: <https://github.com/atomashevic/face-of-populism>. Repository contains a CSV file with a list of URLs of all YouTube videos used in the analysis as well as Python code that can be used to reproduce the CSV files with processed data. Code used for image annotation and the resulting dataset can be found at: <https://github.com/atomashevic/fop-annotation>.

References

1. Demertzis, N. (2006). Emotions and Populism. In: Clarke, S., Hoggett, P., Thompson, S. (eds.) *Emotion, Politics and Society*, pp. 103–122. Palgrave Macmillan UK, London. https://doi.org/10.1057/9780230627895_7
2. Moffitt, B., & Tormey, S. (2014). Rethinking populism: politics, mediatisation and political style. *Political Studies*, 62(2), 381–397. <https://doi.org/10.1111/1467-9248.12032>
3. Mudde, C. (2017). Populism: An Ideational Approach. In: Rovira Kaltwasser, C., Taggart, P.A., Ochoa Espejo, P., Ostiguy, P. (eds.) *The Oxford Handbook of Populism*. Oxford Handbooks, pp. 27–47. Oxford University Press, Oxford, United Kingdom ; New York
4. Norris, P., & Inglehart, R. (2019). *Cultural Backlash: Trump, Brexit, and Authoritarian Populism*. Cambridge University Press. <https://doi.org/10.1017/9781108595841>
5. Bonansinga, D. (2020). Who thinks, feels, the relationship between emotions, politics and populism. *PArtecipazione e Conflitto*, 13(1), 83–106. <https://doi.org/10.1285/120356609V13I1P83>
6. Salmela, M., & von Scheve, C. (2018). Emotional dynamics of right- and left-wing political populism. *Humanity & Society*, 42(4), 434–454. <https://doi.org/10.1177/0160597618802521>
7. Widmann, T. (2021). How emotional are populists really? factors explaining emotional appeals in the communication of political parties. *Political Psychology*, 42(1), 163–181. <https://doi.org/10.1111/pops.12693>
8. Nguyen, C. (2019). Emotions and populist support. SocArXiv. <https://doi.org/10.31235/osf.io/e2wm6>. Accessed 2023-06-22
9. Bobba, G. (2019). Social media populism: Features and ‘likeability’ of Lega Nord communication on Facebook. *European Political Science*, 18(1), 11–23. <https://doi.org/10.1057/s41304-017-0141-8>
10. Breeze, R. (2019). Emotion in politics: Affective-discursive practices in UKIP and Labour. *Discourse & Society*, 30(1), 24–43. <https://doi.org/10.1177/0957926518801074>
11. Breeze, R. (2020). Angry tweets: A corpus-assisted study of anger in populist political discourse. *Journal of Language Aggression and Conflict*, 8(1), 118–145. <https://doi.org/10.1075/jlac.00033.bre>
12. Cochrane, C., Rheault, L., Godbout, J.-F., Whyte, T., Wong, M.W.-C., & Borwein, S. (2022). The automatic analysis of emotion in political speech based on transcripts. *Political Communication*, 39(1), 98–121. <https://doi.org/10.1080/10584609.2021.1952497>
13. Jacobs, K., Sandberg, L., & Spierings, N. (2020). Twitter and Facebook: Populists’ double-barreled gun? *New Media & Society*, 22(4), 611–633. <https://doi.org/10.1177/1461444819893991>
14. Martella, A., & Bracciale, R. (2022). Populism and emotions: Italian political leaders’ communicative strategies to engage Facebook users. *Innovation: The European Journal of Social Science Research*, 35(1), 65–85. <https://doi.org/10.1080/13511610.2021.1951681>
15. Mohammad, S. M. (2016). Sentiment analysis: detecting valence, emotions, and other affectual states from text. In H. L. Meiselman (Ed.), *Emotion Measurement*. Amsterdam: Elsevier; Woodhead Publishing.
16. Moffitt, B. (2022). Taking account of the visual politics of populism. *Polity*, 54(3), 557–564. <https://doi.org/10.1086/719829>
17. Mendonça, R. F., & Caetano, R. D. (2021). Populism as parody: The visual self-presentation of Jair Bolsonaro on Instagram. *The International Journal of Press/Politics*, 26(1), 210–235. <https://doi.org/10.1177/1940161220970118>
18. Masters, R., Sullivan, D., Lanzetta, J., Mchugo, G., & Englis, B. (1986). The facial displays of leaders: Toward an ethology of human politics. *Journal of Social and Biological Systems*, 9(4), 319–343. [https://doi.org/10.1016/S0140-1750\(86\)90190-9](https://doi.org/10.1016/S0140-1750(86)90190-9)
19. Ekman, P., & Friesen, W. V. (1976). Measuring facial movement. *Environmental Psychology and Nonverbal Behavior*, 1(1), 56–75. <https://doi.org/10.1007/BF01115465>

20. Bucy, E. P., Foley, J. M., Lukito, J., Doroshenko, L., Shah, D. V., Pevehouse, J. C., & Wells, C. (2020). Performing populism: Trump's transgressive debate style and the dynamics of Twitter response. *New Media & Society*, 22(4), 634–658. <https://doi.org/10.1177/1461444819893984>
21. Bucy, E. P., & Grabe, M. E. (2007). Taking television seriously: A sound and image bite analysis of presidential campaign coverage, 1992–2004. *Journal of Communication*, 57(4), 652–675. <https://doi.org/10.1111/j.1460-2466.2007.00362.x>
22. Bucy, E. P., & Grabe, M. E. (2008). Happy warriors Revisited: Hedonic and agonic display repertoires of presidential candidates on the evening news. *Politics and the Life Sciences*, 27(1), 78–98. https://doi.org/10.2990/27_1_78
23. Grabe, M.E., & Bucy, E. P. (2009). Image Bite Politics: News and the Visual Framing of Elections. Series in Political Psychology. Oxford University Press, Oxford ; New York
24. Sullivan, D. G., & Masters, R. D. (1988). "Happy warriors": Leaders' facial displays, viewers' emotions, and political support. *American Journal of Political Science*, 32(2), 345–368. <https://doi.org/10.2307/2111127>
25. Stewart, P. A., Salter, F. K., & Mehu, M. (2009). Taking leaders at face value: Ethology and the analysis of televised leader displays. *Politics and the Life Sciences*, 28(1), 48–74. https://doi.org/10.2990/28_1_48
26. Stewart, P. A., Bucy, E. P., & Mehu, M. (2015). Strengthening bonds and connecting with followers: A biobehavioral inventory of political smiles. *Politics and the Life Sciences*, 34(1), 73–92. <https://doi.org/10.1017/pls.2015.5>
27. Stewart, P. A., & Ford Dowe, P. K. (2013). Interpreting president barack obama's facial displays of emotion: revisiting the dartmouth group: Interpreting Obama's facial displays of emotion. *Political Psychology*, 34(3), 369–385. <https://doi.org/10.1111/pops.12004>
28. Joo, J., Bucy, E. P., & Seidel, C. (2019). Automated coding of televised leader displays: Detecting nonverbal political behavior with computer vision and deep learning. *International Journal of Communication*, 13, 23. <https://ijoc.org/index.php/ijoc/article/view/10725>
29. Ash, E., & Caliskan, C. (2022). Fight or Flight on Fox?: Partisan Fear Responses on U.S. Cable News Shows, Rochester, NY. <https://doi.org/10.2139/ssrn.4204973>
30. Bossetta, M., & Schmökel, R. (2023). Cross-platform emotions and audience engagement in social media political campaigning: Comparing candidates' Facebook and Instagram images in the 2020 US Election. *Political Communication*, 40(1), 48–68. <https://doi.org/10.1080/10584609.2022.2128949>
31. Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest: A Journal of the American Psychological Society*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>
32. Goodfellow, I.J., Erhan, D., Carrier, P.L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.-H., Zhou, Y., Ramaiah, C., Feng, F., Li, R., Wang, X., Athanasakis, D., Shawe-Taylor, J., Milakov, M., Park, J., Ionescu, R., Popescu, M., Grozea, C., Bergstra, J., Xie, J., Romaszko, L., Xu, B., Chuang, Z., & Bengio, Y. (2013). Challenges in Representation Learning: A Report on Three Machine Learning Contests. arXiv. <https://doi.org/10.48550/arXiv.1307.0414>
33. Shenk, J., CG, A., & Arriaga, O. (2021). Owlwasrowk: Justinshenk/Fer: Zenodo. Zenodo. <https://doi.org/10.5281/zenodo.5362356>
34. Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499–1503. <https://doi.org/10.1109/lsp.2016.2603342>
35. Arriaga, O., Valdenegro-Toro, M., & Plöger, P. (2017). Real-time convolutional neural networks for emotion and gender classification. arXiv. <https://doi.org/10.48550/arXiv.1710.07557>
36. Arriaga, O., Valdenegro-Toro, M., Muthuraja, M., Devaramani, S., & Kirchner, F. (2020). Perception for Autonomous Systems (PAZ). arXiv. <https://doi.org/10.48550/arXiv.2010.14541>
37. Gidron, N., & Bonikowski, B. (2013). Varieties of populism: Literature review and research agenda. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2459387>
38. Rovira Kaltwasser, C., Taggart, P. A., Ochoa Espejo, P., & Ostiguy, P. (2017) (eds.) The oxford handbook of populism. Oxford Handbooks. Oxford University Press, Oxford, United Kingdom ; New York
39. Mudde, C. (2004). The Populist Zeitgeist. *Government and Opposition*, 39(4), 541–563. <https://doi.org/10.1111/j.1477-7053.2004.00135.x>

40. Ostiguy, P. (2017). Populism: A socio-cultural approach. In: Rovira Kaltwasser, C., Taggart, P.A., Ochoa Espejo, P., Ostiguy, P. (eds.) *The Oxford Handbook of Populism*. Oxford Handbooks, pp. 73–95. Oxford University Press, Oxford, United Kingdom ; New York
41. Weyland, K. (2017). Populism: A political-strategic approach. In C. Rovira Kaltwasser, P. A. Taggart, P. Ochoa Espejo, & P. Ostiguy (Eds.), *The Oxford Handbook of Populism*. Oxford Handbooks (pp. 48–72). Oxford University Press, Oxford, United Kingdom, New York.
42. Hawkins, K.A., & Kaltwasser, C. R. (2018). Introduction: The ideational approach. In: *The Ideational Approach to Populism*, pp. 1–24. Routledge
43. Aslanidis, P. (2017). Avoiding bias in the study of populism. *Chinese Political Science Review*, 2(3), 266–287. <https://doi.org/10.1007/s41111-017-0064-0>
44. Engesser, S., Fawzi, N., & Larsson, A. O. (2017). Populist online communication: introduction to the special issue. *Communication & Society*, 20(9), 1279–1292. <https://doi.org/10.1080/1369118X.2017.1328525>
45. Aslanidis, P. (2016). Is populism an ideology? a refutation and a new perspective. *Political Studies*, 64, 88–104. <https://doi.org/10.1111/1467-9248.12224>
46. Inglehart, R., & Norris, P. (2016). Trump, brexit, and the rise of populism: Economic have-nots and cultural backlash. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2818659>
47. Davidson, T., & Enos, J. (2022). Are populist political parties more popular on social media? Evidence from Europe, 2010–2020. SocArXiv. <https://doi.org/10.31235/osf.io/kdy6c> . Accessed 2023-07-01
48. Wells, C., Shah, D., Lukito, J., Pelled, A., & Pevehouse, J. C. (2020). Trump, twitter, and news media responsiveness: A media systems approach. *New Media & Society*, 22(4), 659–682. <https://doi.org/10.1177/1461444819893987>
49. Larsson, A. O. (2022). Picture-perfect populism: Tracing the rise of european populist parties on facebook. *New Media & Society*, 24(1), 227–245. <https://doi.org/10.1177/1461444820963777>
50. Lalancette, M., & Raynauld, V. (2019). The power of political image: Justin Trudeau, Instagram, and celebrity politics. *American Behavioral Scientist*, 63(7), 888–924. <https://doi.org/10.1177/0002764217744838>
51. Moffitt, B. (2022). How do populists visually represent ‘The People’? a systematic comparative visual content analysis of donald trump and bernie sanders’ Instagram Accounts. *The International Journal of Press/Politics*, 19401612221100418. <https://doi.org/10.1177/19401612221100418>
52. Finlayson, A. (2022). YouTube and political ideologies: Technology, populism and rhetorical form. *Political Studies*, 70(1), 62–80. <https://doi.org/10.1177/0032321720934630>. Accessed 2023-06-23.
53. Mazzoleni, G. (2014). Mediatization and Political Populism. In F. Esser & J. Strömbäck (Eds.), *Mediatization of Politics*. London: Palgrave Macmillan.
54. Bracciale, R., & Martella, A. (2017). Define the populist political communication style: The case of Italian political leaders on Twitter. *Information, Communication & Society*, 20(9), 1310–1329. <https://doi.org/10.1080/1369118X.2017.1328522>
55. Gründl, J. (2022). Populist ideas on social media: A dictionary-based measurement of populist communication. *New Media & Society*, 24(6), 1481–1499. <https://doi.org/10.1177/1461444820976970>. Accessed 2023-05-12.
56. Messaris, P. (2019). The digital transformation of visual politics. In A. Veneti, D. Jackson, & D. G. Lilleker (Eds.), *Visual Political Communication* (pp. 17–36). Springer.
57. Norris, P. (2020). Measuring populism worldwide. *Party Politics*, 26(6), 697–717. <https://doi.org/10.1177/1354068820927686>
58. Wiesehomeier, N. (2018). Expert surveys. In K. A. Hawkins, R. E. Carlin, L. Littvay, & C. Rovira Kaltwasser (Eds.), *The Ideational Approach to Populism* (pp. 90–111). New York: Routledge.
59. Curini, L. (2010). Experts’ political preferences and their impact on ideological bias: An unfolding analysis based on a Benoit-Laver expert survey. *Party Politics*, 16(3), 299–321. <https://doi.org/10.1177/1354068809341051>
60. Heuer, H., Hoch, H., Breiter, A., & Theocharis, Y. (2021). Auditing the biases enacted by YouTube for political topics in germany. In *Proceedings of Mensch und Computer 2021*. MuC ’21, (pp. 456–468). Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3473856.3473864>
61. Khairuddin, Y., & Chen, Z. (2021). Facial emotion recognition: State of the art performance on FER2013. arXiv. <https://doi.org/10.48550/arXiv.2105.03588>

62. Caiani, M., & Di Cocco, J. Italian political science review/Rivista Italiana di Scienza Politica. *Populism and emotions: a comparative study using machine learning*, 53(3), 1–16. <https://doi.org/10.1017/ipo.2023.8>
63. Macagno, F. (2022). Argumentation profiles and the manipulation of common ground. The arguments of populist leaders on Twitter. *Journal of Pragmatics*, 191, 67–82. <https://doi.org/10.1016/j.pragma.2022.01.022>
64. Boussalis, C., & Coan, T. G. (2021). Facing the electorate: Computational approaches to the study of nonverbal communication and voter impression formation. *Political Communication*, 38(1–2), 75–97. <https://doi.org/10.1080/10584609.2020.1784327>
65. Boussalis, C., Coan, T. G., Holman, M. R., & Müller, S. (2021). Gender, candidate emotional expression, and voter reactions during televised debates. *The American Political Science Review*, 115(4), 1242–1257. <https://doi.org/10.1017/S0003055421000666>
66. Jankowski, M., & Huber, R. A. (2023). When correlation is not enough: Validating populism scores from supervised Machine-Learning models. *Political Analysis: An Annual Publication of the Methodology Section of the American Political Science Association*, 31(4), 591–605. <https://doi.org/10.1017/pan.2022.32>
67. Min, S., Yang, J., & Lim, S. (2024). Emotion recognition using transformers with random masking. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (pp. 4860–4865). IEEE. <https://doi.org/10.1109/cvprw63382.2024.00489>
68. Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. arXiv. <https://doi.org/10.48550/arXiv.2103.00020>
69. Waligora, P., Aslam, H., Zeeshan, O., Belharbi, S., Koerich, A.L., Pedersoli, M., Bacon, S., & Granger, E. (2024). Joint Multimodal transformer for emotion recognition in the wild. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, (4625–4635). <https://doi.org/10.1109/CVPRW63382.2024.00465>
70. 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. <https://doi.org/10.1080/19331681.2021.1928579>
71. Ryan, O., Dablander, F., & Haslbeck, J. M. B. (2024). Toward a generative model for emotion dynamics. *Psychological Review*. <https://doi.org/10.1037/rev0000513>
72. James, N. A., & Matteson, D. S. (2015). Ecp: An R package for nonparametric multiple change point analysis of multivariate data. *Journal of Statistical Software*, 62, 1–25. <https://doi.org/10.18637/jss.v062.i07>
73. Tomašević, A., & Major, S. (2024). Dynamic exploratory graph analysis of emotions in politics. *Advances in Psychology*, 2(1), Article 312144. <https://doi.org/10.56296/aip00021>
74. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision. ArXiv. <https://doi.org/10.48550/arXiv.2212.04356>
75. Tomašević, A., Golino, H., & Christensen, A. P. (2024). Decoding emotion dynamics in videos using dynamic Exploratory Graph Analysis and zero-shot image classification: A simulation and tutorial using the transforEmotion R package. PsyArXiv. <https://doi.org/10.31234/osf.io/hf3g7>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.