Faces of Diplomacy:

Visual Framing of Global Leaders in the China Daily (2020-2025)

Julius Pfundstein (University of Leipzig)

Abstract

This study investigates the visual framing of global leaders in the *China Daily* from 2020 to 2025, focusing on how emotional portrayals of international political elites vary depending on their countries' diplomatic relationships with China. Drawing on a dataset of 28,881 images extracted from *China Daily* articles, automated facial recognition and facial expression recognition are employed to classify and evaluate the emotional expressions of 225 de-facto leaders. Statistical analysis indicates a weak but significant correlation between predominantly neutral portrayals and nations aligned with China, whereas emotions, especially negative emotions, are more prevalent among leaders of adversarial states. This research contributes to understanding the role of visual media in shaping public perceptions of global diplomacy and highlights the potential for systematic bias in news image selection.

1 Introduction

Angry Trump, happy Obama, and cold Putin - these are the impressions readers might form if they were to base their assessments solely on the images accompanying newspaper articles. In the past, only a select few front-page stories featured images. Today, however, in the fierce competition for fragmented digital attention, every news outlet is required to curate an image to stand out amidst the clutter of social media and news aggregators. Images serve as powerful primers and emotional agenda-setters whose influence persists throughout an article even when they are just scrolled past. [1][2]

But just as with textual depictions, small variations in the selected images can alter the presented narrative. The potential for deliberate manipulation in the selection of images is particularly pronounced when depicting people. Unlike static objects, human faces appear in a wide variety of expressions within a short period of time. This grants journalists and editors the freedom to choose from a broad spectrum of emotions for their portrayal. As multimodal data becomes more prevalent in political communication and more accessible for research, the investigation of visual framing is both necessary and rewarding. For example, were bipartisan news outlets found to be depicting

[1] Dahmen, Nicole et al.; 2019; The (in)disputable 'power' of images of outrage: public acknowledgement, emotional reaction, and image recognition.

[2] Isaacs, David; 2016; Power of visual images.

aligned politicians with more positive emotions in the 2016 US election. [3] A comparative study in 2012 found that aligned politicians were portrayed with higher saturation and closer camera perspectives. [4] And while Trump is proven to be depicted more angry than other US-politicians, variations across opposing bipartisan media companies were insignificant. [5] That result is surprising and intrigues further investigation. To add to the study of visual framing in newspaper images, this study seeks to test the underlying hypothesis of systematic emotional selection through a case study of the state-run newspaper *China Daily*.

RQ: How does the emotional portrayal of international political elites in the *China Daily* (2020-2025) vary depending on their countries' political alignment with China?

---

2 Data Curation and Cleansing

To address the research question, three independent datasets were compiled through a systematic process.

2.1 De-Facto Leader List

The first step involved creating a comprehensive list of individuals whose portrayals would be analyzed. Given that the highest political office does not always correspond to the formal presidential role and considering the impracticality of conducting exhaustive research for all 192 internationally recognized states, LLMs were employed. [6] First, ChatGPT Search was tasked with generating a list of individuals who held the highest political office in their respective countries between January 1, 2020, and January 1, 2025. This preliminary list was subsequently refined using DeepSeek R1 Search and Kimi K1.5 Search to ensure accuracy and comprehensiveness. Particularly complex cases, such as Burkina Faso, Myanmar, and the United Kingdom, were tested against the resulting list and confirmed the robustness of this approach. The final dataset comprises a curated list of 225 *de-facto leaders*, organized by country, ensuring a robust foundation for subsequent analysis.

2.2 Wikimedia

The task of identifying the 225 *de-facto leaders* within the *China Daily* dataset necessitated the automation of facial recognition due to the scale of the data. Given the legal restrictions surrounding the use of personal data, no publicly accessible model exists that is pre-trained to classify politicians or other public figures. However, general-purpose facial recognition models can be adapted to recognize specific individuals when provided with labeled training data. To address this requirement, the Wikimedia API was utilized as a source of freely available and appropriately licensed images of the investigated politicians.

---

[3] Boxell, Levi; 2021; Slanted Images: Measuring Nonverbal Media Bias During the 2016 Election.

[4] Hehman, Eric; 2012; Warmth and competence: A content analysis of photographs depicting American presidents.

[5] Li, Chenfeng; 2024; Examining the Interplay Between Politicians' Facial Expressions in Media Images and News Corporation Bias.

[6] San Marino was removed because its legislative term is only six months.

For each entry in the *de-facto leader* list, links to the 200 most relevant image files on Wikimedia were retrieved and subsequently downloaded into a structured directory already including the country of origin. This process yielded an initial dataset comprising 32,614 images (95.1 GB), which equals an average of 145 images per *de-facto leader*.

To optimize the dataset for training purposes, several preprocessing steps were undertaken:

[1] Filtering: Images that contained less or more than one face were removed. This reduced the dataset to 9,885 images. The magnitude of this reduction was primarily due to the frequent depiction of leaders from *small states* alongside representatives of *big states*.

[2] Cropping: Images were cropped to include only the detected faces, ensuring focus on the relevant facial features.

[3] Standardization: All images were resized to a uniform resolution of 300×300 pixels to maintain consistency across the dataset.

[4] Manual Filtering: A manual review was conducted to remove irrelevant or problematic images, including caricatures, images of images, misidentified individuals, non-human subjects, format errors, and childhood photographs. After this step, the dataset contained 6,504 images (87.8 MB).

[5] Data Augmentation: To ensure a minimum of 100 images per *de-facto leader*, new training data was generated by applying random variations to the existing images. These transformations included: (1) Rotation by up to 15 degrees, (2) horizontal mirroring, (3) grayscale conversion, and (4) partial occlusion using black rectangles. Following augmentation, the final Wikimedia dataset comprised 22,607 images (312 MB).

2.3 China Daily

To construct the dataset central to this study, it was initially planned to source data from the Instagram accounts of various Chinese state-led newspapers. However, the successive cutback of the GraphAPI renders it largely obsolete for quantitative analysis. Despite employing techniques such as rotating IP addresses, randomized sleep intervals, and multiple accounts, no more than a few thousand images could be downloaded per day. Beyond this limitation there is the fact that Instagram's content is less political, yielding only 4% relevance. This means that 96% of the downloaded images did not depict any *de-facto leaders*, reducing the daily count of relevant images to mere triple digits. [7]

In search of a more viable alternative, *China Daily* emerged as the preferred source over *Xinhua, CCTV*, and *CGTN* owing to its advanced search functionality. The website's search tool allows users to specify a time frame (01-01-2020 to 01-01-2025), include a list of required keywords within articles, and enable a duplication removal feature (as *China Daily* operates multiple outlets, and some articles are republished across different platforms). [8]

Since all links in the result page are dynamically created and a "next button" had to be located and clicked after downloading images of 10 articles, Selenium's headed browser automation was used. For initial queries, generic terms commonly

[7] This estimate is based on an initial scrape of 6115 images of CGTN_europe.

[8] China Daily advanced search.

associated with political leaders, such as "president," "prime minister," or "chancellor" were used. However, these terms proved problematic, as institutional functions are not always explicitly mentioned, and non-state entities may use similar terminology for their personnel. [9] To address this issue, a more precise approach was adopted: Assuming that any individual prominently depicted in an image would be mentioned at least once by their full name in the accompanying article, it was iterated through the *de-facto leader list* and searched for each entry independently. This method introduces duplications because one article can mention several *de-facto leaders*. But fortunately, the images are named with unique identifiers, and therefore the duplicates are easy to remove.

The 225 queries yielded a total of 28,881 images (11.04 GB). After removing duplicates, the dataset was reduced to 20,736 images. The faces were subsequently cropped and standardized to align with the format of the training data, ensuring compatibility for subsequent analyses.

3 Method

With all datasets prepared, the analysis proceeds in four stages: Training a facial recognition model to identify *de-facto leaders*, classifying their appearances in the *China Daily* dataset, analyzing the emotional expressions depicted in these images, and testing the aggregated emotions against political alignment with China.

3.1 Training Model

The facial recognition model was trained using Pythons face recognition library to generate 128-dimensional feature vectors from the Wikimedia dataset. [10] These encodings were used to train a Support Vector Machine (SVM) classifier capable of distinguishing between the 225 *de-facto leaders*. The training pipeline consists of the following:

[1] Dataset Scanning: The script iterates through the directory structure of the preprocessed Wikimedia dataset, which is organized by ISO country codes and politician names. Each image file is processed to detect faces and extract feature encodings.

[2] Label Encoding: Politician names are encoded into numerical labels using *LabelEncoder* to facilitate classification.

[3] Model Training: A linear-kernel SVM classifier is trained on the feature encodings and corresponding labels. The classifier is configured to return probability estimation to support confidence scoring during inference.

[4] Model Saving: The trained model, along with the label encoder and class mappings, is serialized and saved using *joblib* for subsequent use in classifying images from the *China Daily* dataset.

3.2 De-Facto Leader Classification

The trained facial recognition model was applied to classify *de-facto leaders* in the *China Daily* dataset. In the 20,736 images, it found 20,217

[9] An experimental search for ["Donald Trump"] during his first term generates 6692 results while a search for ["Donald Trump" AND "president"] yielded only 6508 results.

[10] Ageitgey, Davis; 2024; face-recognition PyPl.

faces. However, due to the model's high sensitivity and a significant rate of false positives, manual verification was conducted to ensure accuracy, reducing the dataset to 5,202 reliably classified images. No images were found for 45 *de-facto leaders* from 36 countries, as these individuals are rarely depicted individually and often appear only in group photos at summits or panel discussions (where resolution for individual faces is low), making recognition challenging. [11] A cross-check against *China Daily's* advanced search confirmed the plausibility of these findings.

3.3 Facial Expression Recognition

After successfully identifying and verifying the faces of political figures, the next step involved analyzing the emotional expressions depicted in the images. This was achieved using the *vit-Facial-Expression-Recognition* model, which was trained on large amounts of data (FER2013, MMI Facial Expression, and AffectNet) to classify facial expressions into seven categories: ***Anger***, ***Disgust***, ***Fear***, ***Happiness***, ***Neutrality***, ***Sadness***, and ***Surprise***. [12] For each image, the model outputs a probability distribution across these emotions, enabling a detailed analysis of emotional portrayals. The results for each country-folder were aggregated and stored for further analysis.

4 Results

Before analyzing the emotional portrayals, it is essential to examine the distribution of the 5,202 images of the sorted *China Daily* dataset across the 225 *de-facto leaders*. This distribution not only provides valuable insights but also highlights significant caveats regarding data quality. A striking imbalance is evident: Xi Jinping alone is depicted 1,926 times (37.02%), while the top 10 most frequently depicted leaders account for 3,565 images (68.53%). In contrast, the least depicted 200 leaders are represented in only 1,079 images (21.09%), with 45 leaders having no images at all. **Graph 1** illustrates this uneven distribution across politicians, countries, and continents.

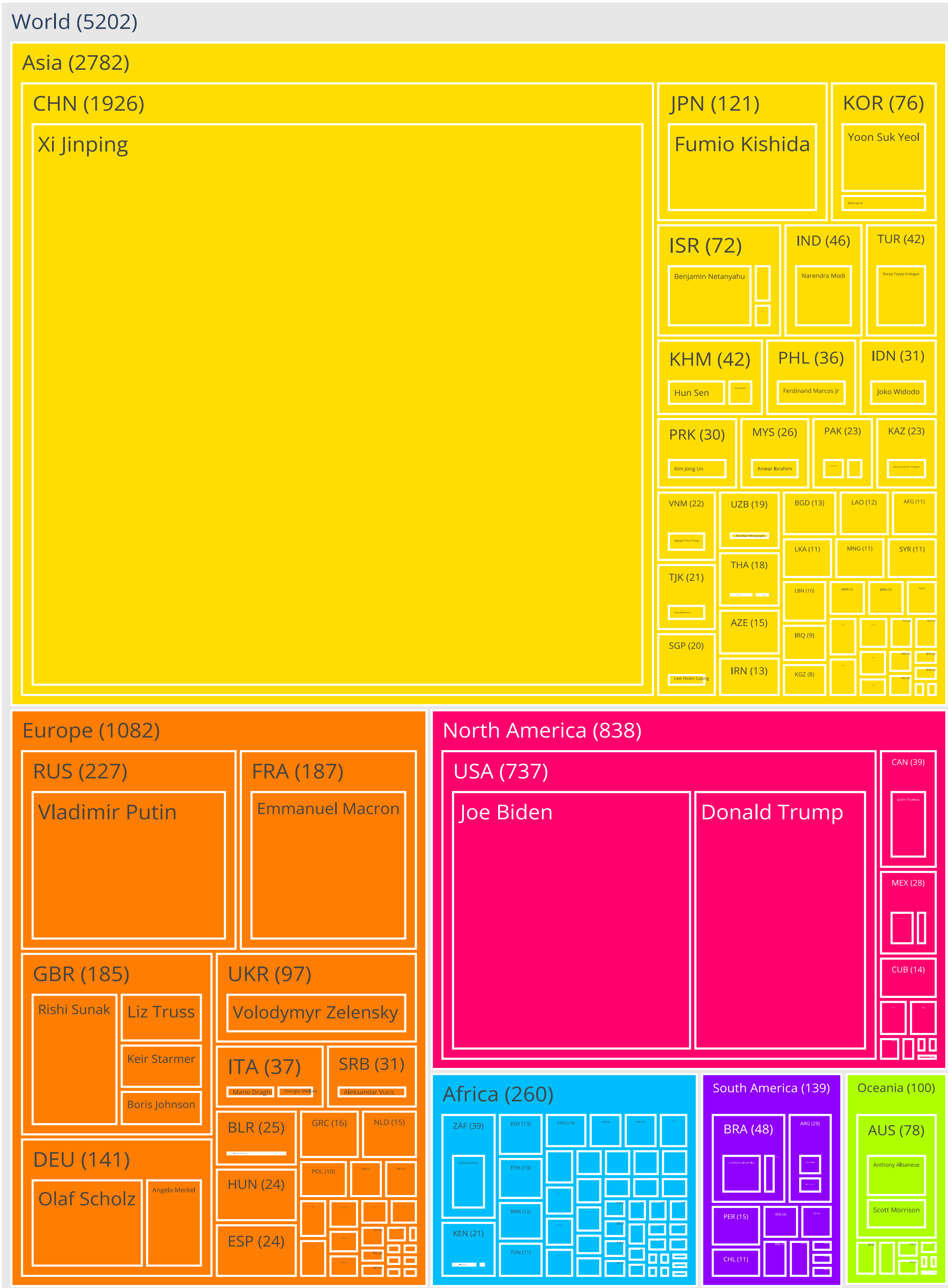
With this context in mind, we proceed to the analysis of emotional portrayals. The heatmap in **Graph 2** shows the results of the facial expression recognition analysis for each country and each emotion. Countries are sorted by their voting alignment with China on UNGA resolutions from January 1, 2020, to June 1, 2024. [13] At the top are China and its closest allies, while the United States and its allies occupy the lower end. The investigated hypothesis predicts that leaders aligned with China (at the top) are more likely to exhibit positive emotions (***Happiness***), while those at the bottom are more likely to display negative emotions (***Anger***, ***Fear***, ***Sadness***, ***Disgust***).

[11] ISO codes of the countries, not contributing any images to the *China Daily Dataset*: AND, BFA, BHS, BLZ, CHE, CPV, DOM, EST, GRD, GUY, KNA, LCA, LIE, LVA, MCO, MDV, MHL, MKD, MLT, MNE, MUS, NER, NIC, NRU, PAN, PLW, SLV, SMR, SSD, TCD, TTO, TUV, VAT, VCT, VUT, YEM

[12] Abdeldayem, Mohammed; 2024; vit-Facial-Expression-Recognition.

[13] The UNGA-Voting Viewer was used to aggregate the voting coincidences. Ordinal rankings were preferred to interval-scaled coincidences to reflect the non-equidistant nature of diplomatic alignment.

Graph 1 - Distribution of De-Facto Leader Images from *China Daily* by Continent and Country



At first glance, the hypothesis is not strongly supported. While the prevalence of ***Anger*** shows a slight concentration toward the lower end of the graph, and ***Happiness*** and ***Neutrality*** exhibit a marginal upward bias, the overall pattern is marked by outliers and cross-country similarities. ***Happiness*** and ***Neutrality*** dominate most country averages, while ***Fear***, ***Sadness*** and ***Disgust*** are negligible for nearly all countries. ***Surprise*** and ***Anger*** are dominant in only two states - Slovenia and Cyprus - but closer inspection reveals that these anomalies stem from single-image representations.

To develop a deeper understanding of the prevalence patterns of emotional expressions across countries, statistical metrics were employed to quantify and evaluate the relationships between emotional portrayals (interval-scaled emotion distribution averages) and diplomatic alignment rankings (ordinal-scaled voting coincidence ranking). Pearson's correlation was used to assess the linear relationships, while Kendall's T, a non-parametric measure, captured monotonic relationships while mitigating the influence of outliers. The following table shows the strength and significance of the relationships between emotional portrayals of *de-facto leaders* and their countries political alignment with China.

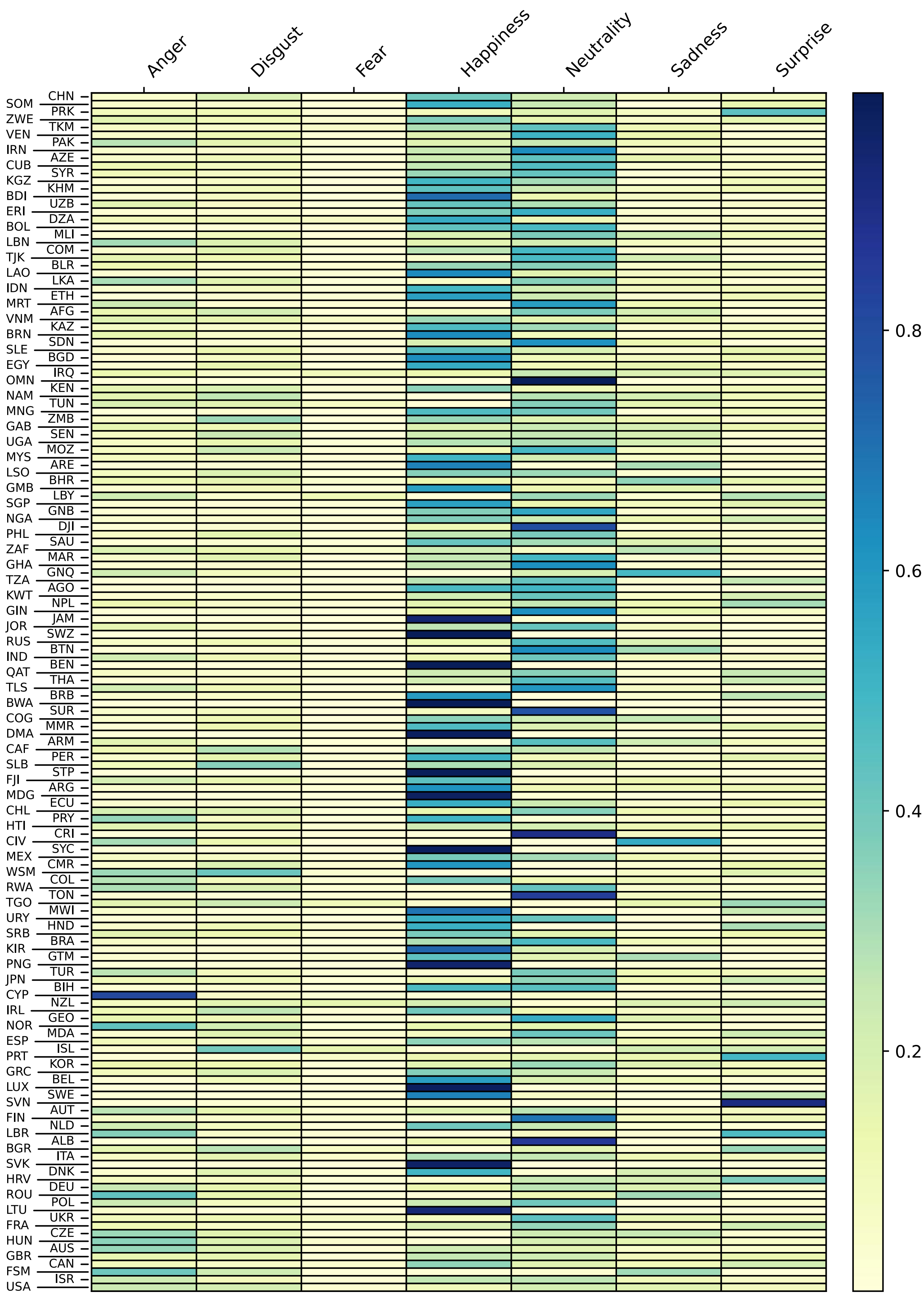
Emotion	Pearson's r (p-value)	Kendall's T (p-value)
Anger	0.251**(0.0015)	0.098 (0.0691)
Disgust	0.090 (0.2624)	0.031 (0.5645)
Fear	0.139 (0.0835)	0.112* (0.0372)
Happiness	-0.074 (0.3556)	-0.095 (0.0775)
Neutrality	-0.210* (0.0086)	-0.158** (0.0035)
Sadness	0.047 (0.5639)	0.009 (0.8661)
Surprise	0.168* (0.0364)	0.031 (0.5645)

** p < 0.0071 (Bonferroni adjustment for 7 tests)
* p < 0.05

As already assumed, the prevalence of ***Disgust***, ***Fear***, and ***Sadness*** are not systematically associated with political alignment to China. Their correlation strengths are weak (< 0.3) and insignificant. The differences between Person's r and Kendall's T for ***Surprise***, for which only five countries exceed a prevalence of 0.4, show the ladder's robustness against outliers. ***Happiness***, which most of the field's literature suggested to be the most dominant correlate of visual framing does show a small negative correlation but high p-values. ***Anger*** shows the strongest linear correlation but falls flat on both strength and significance of Kendall's T. Finally, ***Neutrality*** is the most robust correlate of the seven emotions featuring the highest correlation values and the lowest p-values.

Five of the seven measured emotions correlate positively with their country's opposition to China. While each of those five can be considered a set of outliers with a slight tendency individually, understanding them as a collective unit of "negative emotions" allows for a more coherent theory. That states with closer alignment to China feature more ***Neutrality*** could therefore be understood as residual value to the collective lack of "negative emotions". Also, in light of the weak correlation of ***Happiness***, this hypothesis could expand ***Neutrality*** to be the residual of emotionality per se and thereby indicating that allies are shown with more composure while rivals are shown with a broader range of emotions.

Graph 2 – Emotion Prevalence of De-Facto Leaders across all Countries



5 Conclusion

Before addressing the research question, several important caveats must be acknowledged to contextualize the findings and highlight areas for further investigation.

Interpretative Challenges: the curation of emotional expressions can serve divergent narrative goals depending on the ideological lens of the media outlet. For instance, as noted by Li, conservative outlets may frame anger in political figures like Trump as assertiveness or determination, while liberal outlets might frame the same images in the context of impulsiveness or demagoguery. [14] Research even suggests that such portrayals can be politically advantageous, with studies showing that Trump and Kasich benefited from depictions of anger. [15] This underscores the complexity of interpreting emotional framing and its potential alignment with specific political agendas.

Individual Emotional Variability: The research design carries the premise that all politicians exhibit similar distributions of emotional expressions, with variations in their portrayal attributed solely to editorial curation. However, this assumption holds only to a limited extent. Individual leaders inherently display unique emotional patterns, and it is possible that these patterns themselves correlate with their countries' diplomatic alignment with China.

Gender Bias: Gender dynamics in media representation warrant consideration. Research shows that women are often depicted as less composed than men, potentially introducing a bias in cross-country comparisons. [16] One might speculate that countries with lower voting alignment with China have more female leaders, and therefore exhibit greater emotional variability. However, this concern is mitigated by the distribution of the 15 women featured in the *de-facto leader list*, who are not disproportionately concentrated at either end of the heatmap. [17]

Pre-Trained Algorithmic Bias: Algorithmic limitations in facial recognition technology introduce additional layers of potential bias. Studies indicate that many classifiers perform better on Caucasian males due to their overrepresentation in training datasets. [18] [19] This limitation may manifest in two stages of the pipeline: The reduced training dataset from Wikimedia and the precision of the classifier. Although we lack detailed information about the

[14] Li, Chenfeng; 2024; Examining the Interplay Between Politicians' Facial Expressions in Media Images and News Corporation Bias.

[15] Facing the Electorate: Computational Approaches to the Study of Nonverbal Communication and Voter Impression Formation

[16] Facing the Competition: Gender Differences in Facial Emotion and Prominence in Visual News Coverage of Democratic Presidential Primary Candidates

[17] Women of the *de-facto leader list:* Sheikh Hasina (BGD), Mia Mottley (BRB), Angela Merkel (DEU), Mette Frederiksen (DNK), Liz Truss (GBR), Xiomara Castro (HND), Katrin Jakobsdottir (ISL), Giorgia Meloni (ITA), Maia Sandu (MDA), Claudia Sheinbaum (MEX), Dina Boluarte (PER), Zuzana Caputova (SVK), Natasa Pirc Musar (SVN), Samia Suluhu Hassan (TZA), Fiame Naomi Mataafa (WSM).

[18] Buolamwini, Joy; 2018; Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.

[19] Grother, Patrick et al.; 2019; Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects.

characteristics of missing data, the proportional distribution of *de-facto leaders*, countries, and continents among retrieved images aligns closely with their representation in articles containing their names, suggesting minimal systematic bias in image retrieval. [20]

Finally, instead of target narratives, the nature of diplomatic formalism may influence the findings. Closer alignment with China likely correlates with a higher proportion of images captured in formal diplomatic settings, which tend to favor neutral expressions. Conversely, countries with limited diplomatic exchange may rely on publicly available photos, which are more likely to reflect emotionally charged contexts. This introduces a structural bias tied to the availability and context of images.

Despite these caveats, the hypothesis that there is a systematic relationship between the depiction of global leaders and their countries' geopolitical alignment with China is robustly supported. However, contrary to the expectations of most of the literature, the distinction does not primarily lie in the dichotomy of positive versus negative emotions but rather in the balance between emotional expressiveness and composed neutrality.

While this study provides valuable insights into the visual framing of global leaders, further academic scrutiny is essential to validate and expand upon these findings. Future research could explore how portrayals change over time, find the nexus of visual and textual framing, or expand on the functioning of narratives of emotion depictions.

[20] Distribution across continents (28881 articles in total): Asia (12600), North America (8177), Europe (5237), Africa (1329), South America (632), Oceania (450)

6 References

Boussalis, Constantine, and Travis G. Coan. “Facing the Electorate: Computational Approaches to the Study of Nonverbal Communication and Voter Impression Formation.” *Political Communication* 38, no. 1–2 (March 15, 2021): 75–97. https://doi.org/10.1080/10584609.2020.1784327.

Boxell, Levi. “Slanted Images: Measuring Nonverbal Media Bias During the 2016 Election,” April 15, 2021. http://dx.doi.org/10.2139/ssrn.3837521.

Buolamwini, Joy, and Timnit Gebru. “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.” In *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, 77–91. PMLR, 2018. https://proceedings.mlr.press/v81/buolamwini18a.html.

Dahmen, Nicole Smith, Natalia Mielczarek, and Daniel D Morrison. “The (in)Disputable ‘Power’ of Images of Outrage: Public Acknowledgement, Emotional Reaction, and Image Recognition.” *Visual Communication* 18, no. 4 (November 2019): 453–74. https://doi.org/10.1177/1470357217749999.

Grother, Patrick, Mei Ngan, and Kayee Hanaoka. “Face Recognition Vendor Test Part 3: Demographic Effects.” Gaithersburg, MD: National Institute of Standards and Technology, December 2019. https://doi.org/10.6028/NIST.IR.8280.

Gruszczynski, Mike, Danielle K. Brown, Haley Pierce, and Maria E. Grabe. “Facing the Competition: Gender Differences in Facial Emotion and Prominence in Visual News Coverage of Democratic Presidential Primary Candidates.” *Journalism & Mass Communication Quarterly* 100, no. 3 (September 1, 2023): 498–528. https://doi.org/10.1177/10776990221124944.

Hehman, Eric, Elana C. Graber, Lindsay H. Hoffman, and Samuel L. Gaertner. “Warmth and Competence: A Content Analysis of Photographs Depicting American Presidents.” *Psychology of Popular Media Culture* 1, no. 1 (January 2012): 46–52. https://doi.org/10.1037/a0026513.

Isaacs, David. “Power of Visual Images.” *Journal of Paediatrics and Child Health* 52, no. 9 (September 2016): 859–60. https://doi.org/10.1111/jpc.13330.

Li, Chenfeng. “Examining the Interplay Between Politicians’ Facial Expressions in Media Images and News Corporation Bias.” Master Thesis, University of Chicago, 2024. https://www.chenfengli.com/research.

Thomas, Christopher, and Adriana Kovashka. “Predicting the Politics of an Image Using Webly Supervised Data.” arXiv, October 31, 2019. https://doi.org/10.48550/arXiv.1911.00147.

Click here to find the GitHub repository (featuring code & datasets)