



Perceived masculinity from Facebook photographs of candidates predicts electoral success

Kunwoo Park^{1*}  and Jungseock Joo^{2*} 

*Correspondence:

kunwoo.park@ssu.ac.kr;
jjoo@comm.ucla.edu

¹School of AI Convergence,
Soongsil University, 369 Sangdo-ro,
Dongjak-gu, Seoul, 06978, Republic
of Korea

²Department of Communication,
University of California Los Angeles,
Los Angeles, 90095, California,
United States

Abstract

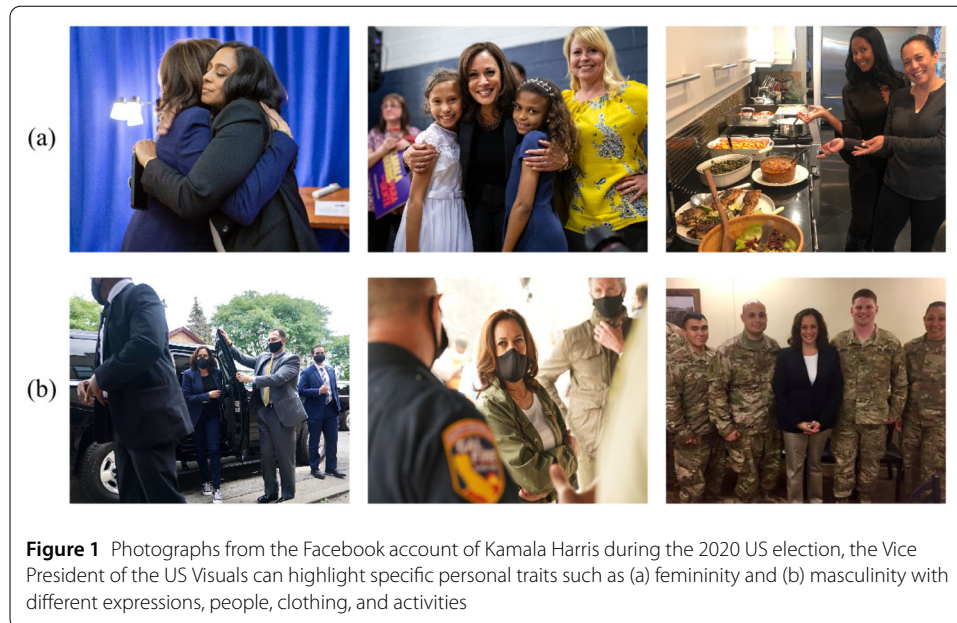
Politicians have used the web and social media to circumvent the gatekeeping behavior of traditional mass media by directly communicating with supporters in their accounts. This paper is aimed at understanding communication strategies used by politicians and campaigns, focusing on the role of gender cues in their visual self-presentation and their impact on election outcomes. Previous research has discussed the importance of visual portrayals of leaders in campaigns. These studies, however, have been mainly based on manual coding and are limited in scale and scope. This paper aims to fill the research gap by introducing a multi task method that infers perceived gender-stereotypical visual traits from social media images. We analyze 77,861 photographs collected from the Facebook accounts of 554 US politicians who ran in the 2018 elections. Regression analyses discover the positive association of the masculinity trait for electoral outcomes. We also identify an empirical evidence that the effect of gender stereotypes could vary according to the gender and party combinations of the candidates in a race. In the intersectional analysis, we found that the win of female democrats against the same gendered opponent was positively correlated with the femininity trait score. This study provides methodological foundations and empirical contributions to the understanding of politicians' campaign behaviors via photographs shared on social media and their relation to electoral success.

Keywords: Visual communication; Personal trait; Gender stereotype; Political campaign; Deep learning; Computer vision

1 Introduction

The advancement of internet and communication technology has allowed people to connect with others through online social platforms. More citizens are using social media as their primary news sources [1], and politicians are also quickly shifting to social media to reach out to their supporters by circumventing journalistic gatekeeping processes [2]. Figure 1 shows a few photographs posted by Kamala Harris on her Facebook account. As shown in the examples, politicians routinely rely on typical imagery and visual rhetoric to emphasize certain personal traits and leadership styles where they are self-assigned various roles such as a “good mother” or “invincible fighter”. These persuasive self-elected

© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.



characterizations are articulated by emotional expressions (smile or shout), performance activities (cooking or shooting), and interacting with people (children or veterans) in photographs [3]. As such, politicians have “professionalized” visual communication online by carefully and strategically selecting content to effectively appeal to their voters as an ideal leader [4, 5]. In contrast to text, politicians’ visuals can trigger voters’ heuristic routes of information processing [6]; it can thus facilitate prefigured perceptions of politicians’ traits in a certain way. Such accessible heuristics can play an important role in electoral success especially by affecting the decisions of voters who are not interested in politics [7].

A substantial body of literature has examined the use of gender traits in visual presentation in electoral campaigning. Previous research has shown that masculine visuals, such as wearing a formal suit, when perceived as a positive trait [8], can lead to voters’ perception of a politician as an ideal candidate of competency, leadership and eventually electoral success [9, 10]. In contrast, feminine stereotypes view women as emotional, warm, and nurturing, which contrasts with the expected characteristics of a political leader, such as being outspoken and aggressive [11]. While there has been empirical evidence of such adverse effects on voters through the use of feminine visuals [12, 13], several studies have argued that expressing femininity could be an effective presentation strategy for female candidates to induce more voters to support them [14, 15]. Others argued that masculine visuals expressed by female politicians could cause a backlash [16, 17], as suggested in Hillary Clinton’s defeat in the 2016 election.

The above observations in previous research suggest that expressing gender-stereotypical traits plays a role in political campaigning but only in a complicated fashion. The literature allows only for a limited understanding due to the restricted capability of manual coding and interview-based methods, which were generally employed. An analysis of selected photographs only allows us to have a partial understanding on the effects of gender cues displayed through a large number of photographs shared over an election campaign. Research that relies on only a few candidates in an election makes it difficult

to compare findings with other observations from a different election where contextual factors differ.

This study aims to fill this gap by analyzing a comprehensive dataset of campaign photographs shared on social media for a single election in the United States. For the automated inference of visual traits, we propose using techniques in deep learning and computer vision. In particular, we introduce a multitask learning method that predicts multiple personal traits simultaneously. Trained on crowdsourced annotations that could represent perceived visual traits, the proposed approach allows us to examine the nuanced effects of gender cues on electoral success in the analysis of 77,861 campaign images. Based on a research hypothesis that the effects of gender-stereotypical traits may vary according to the combination of gender and political party in a race, we investigate in which context masculinity and/or femininity are correlated to electoral success.

To summarize, this study asks three research questions:

- RQ1. Can we automatically infer gender-stereotypical traits portrayed in politicians' campaign images?*
- RQ2. How were gender traits in campaign photographs associated with electoral success in an election?*
- RQ3. How does the gender and party combination of politicians in a race interact with the association of gender traits?*

2 Related works

2.1 Deep learning for visual media analysis

Computational social scientists have recently been using computer vision and deep learning for the large-scale visual content analysis of massive amounts of data scrapped from the media. Computer vision is a subfield of computer science that deals with how computers can gain high-level understanding from digital images or videos. Deep learning is a branch of learning methods designed for artificial neural networks. In the last decade, deep learning has developed rapidly and boosted the performance of computer vision methods. The automated approach for analyzing visual content can significantly improve the efficiency of coding and provide new insights into human behaviors and social events such as emotional understanding [18], elections [19–22], collective actions and protests [23, 24], and inferences about personal traits and ideology [25–27]. Since a substantial portion of online communication is conducted in the form of visual data, image data offer unprecedented potential for social science research on the web [28]. Recent work has used deep learning to assess subjective psychological cues such as emotion [29] or personality [30, 31] from user photographs on social media. Other research examined visual arts and photography using computer vision that brings insight into the history analysis [32]. The increasing popularity of visual media platforms and advances in deep learning have enabled large-scale computational analyses to predict subtle cues from images. This paper takes a similar approach to infer perceived personal traits from images in the context of politics.

2.2 Gender stereotype and personal traits in political communication

Here, we review personal and visual traits covered in political communication research as a relevant trait for electoral success.

Research has highlighted the importance of stereotypical gender dimensions, “feminine” and “masculine,” by identifying the adaptation of dimensions in the campaign ads of female candidates. For example, Hillary Clinton aired a campaign ad that showed a mother

checking in on a sleeping child while the narration talked about protecting the country from national security threats [17]. The “qualified” dimension was covered in a study [33]. Even if a voter favors a male candidate, a female candidate stands a chance as long as her unique information makes it evident that she is more qualified than her male opponent. Similarly, the dimension “competent” has been identified as a key factor in the trait evaluation of politicians; once a voter, who initially holds gender stereotypes about female candidates, learns from relevant information that she is actually competent, a voter becomes motivated to be more readily engaged in information search [34]. The dimension “ordinary”, one of the subdimensions of populist narratives that are built on the idea that ordinary people stand in opposition to self-serving elites, has also been identified as a potential factor in electoral success for candidates whose campaign theme is coherent with populist framing. The dimension “elitist” was included in our study based on the same literature that has reported that engaging in expensive recreational activities reinforces the aristocratic image, causing the candidates to be perceived as elitist figures that are distant from the middle class [35]. The dimensions “attractive” and “threatening” have been discussed as negative traits in political communication; fleeting attractiveness and covertly threatening faces could backfire on politicians by making them look incompetent [36]. Along with the dimensions “aggressive”, the dimension “ambitious” was covered in previous studies [37]. While aggressive female candidates were perceived as more qualified, unambitious female candidates received a higher overall rating associated with candidate image. The “communal” dimension was discussed as a subdimension of the feminine trait. Visual cues or linkages to the dimensions “formal” and “patriotic” have been identified as subcategories that manifest statesmanship for the ideal candidate frame [35]. Other personal dimensions such as “energetic,” “trustworthy,” and “confident” are also frequently used in studies in the visual analysis of perceived personality and persuasive intent of politicians in media [3, 20].

Based on the literature, this study exploits 22 visual traits with a focus on masculinity and femininity. Due to their abstract nature, voters may perceive those traits to be interrelated. We aim to identify gender-stereotypical traits with correlation analysis and use them for further analysis.

3 Data and methods

In this section, we describe the dataset of social media images shared for election campaigns. We also present a deep learning method used for inferring personal traits related to gender stereotypes and electoral success (RQ1).

3.1 Data description

This study examines a comprehensive dataset of campaign images collected in our previous study [22]. It consists of the images shared by the Facebook accounts of US politicians who ran for the 2018 House, Senate, or Governor election. Using the list of political candidates collected from Ballotpedia¹ and manually identified Facebook accounts, we downloaded public photographs shared over the year 2018 until the election date (November 6). As shown in Fig. 2, politicians posted more images as the election day approached. To control for temporal differences between politicians, we focus on the last three months before

¹<https://ballotpedia.org>.

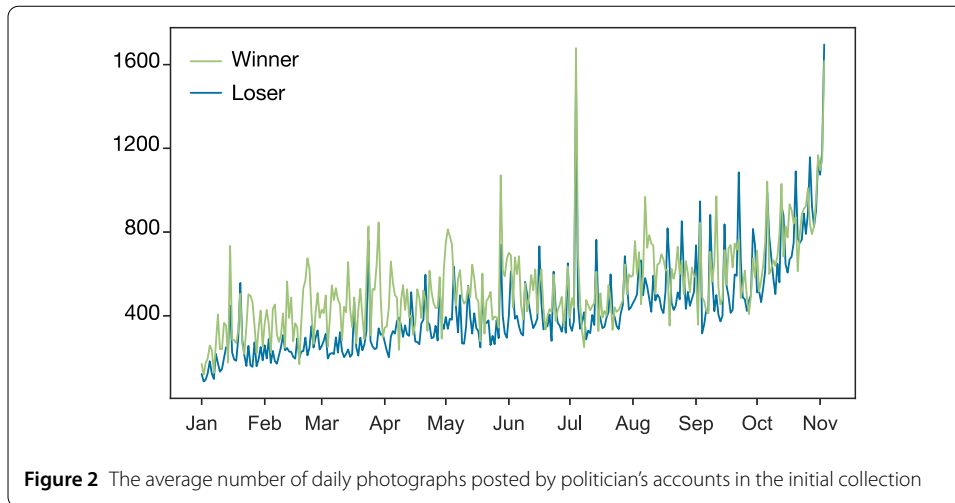


Table 1 Descriptive statistics of the target dataset

Party	Gender	Elected?	Number of politicians	Mean image count per politician	Median image count per politician
DEM	Male	Win	75	131.2	89
		Lose	71	154.9	91
	Female	Win	53	136.8	83
		Lose	81	141.8	121
GOP	Male	Win	137	126.9	89
		Lose	94	156.9	78
	Female	Win	14	192.3	132
		Lose	29	119.1	74

the election date. Table 1 shows the descriptive statistics of the target dataset, which consists of 77,861 images posted by 554 politicians. The dataset is well-balanced according to self-identified gender, party, and election outcomes.

Based on the literature on visual communication in the political context, we summarize 22 personal traits that were identified as factors associated with electoral success, as shown in Table 2. The traits provide a critical basis for understanding the characteristics that voters perceive an ideal leader to have. Note that the visual traits are abstract concepts that can be individually and distinctly interpreted by each viewer. Some people could consider smiling faces to be a feminine concept, while others can think it is gender-neutral. Therefore, it is crucial to capture collective perception of each trait because the *perceived* traits indeed affect the voting decisions of the electorate and election outcomes. To obtain collective perception on visual traits, we conducted crowdsourced annotations on Amazon Mechanical Turk. Using a sample of 8462 images balanced by gender and political party, we asked each annotator to evaluate to what extent a politician expresses a trait on a five-point scale (1 to 5). We instructed the annotators to make their assessment after giving them an objective definition of each trait. For example, the definition of femininity was given as “the quality or nature of the female sex and can be either explicitly (made obvious) or implicitly (indirectly stated) expressed.” We controlled the annotators’ characteristics by excluding responses from annotators who were not familiar with US politics or who could recognize the politicians in the given photographs. Ten annotators were assigned to each image.

Table 2 Split-half reliability for the crowdsourced annotations

Trait	Agreement	Trait	Agreement
Feminine	0.63	Ambitious	0.306
Formal	0.565	Qualified	0.304
Patriotic	0.564	Agreeable	0.303
Masculine	0.561	Attractive	0.3
Professional	0.525	Elitist	0.28
Friendly	0.446	Competent	0.261
Maternal	0.418	Threatening	0.261
Energetic	0.409	Reassuring	0.245
Caring	0.374	Trustworthy	0.245
Communal	0.344	Credible	0.235
Confident	0.32	Ordinary	0.228

Table 2 presents the annotation quality. Instead of conventional agreement measures such as Fleiss's kappa or Cronbach's alpha, we compute split-half reliability (SHR) values that are commonly used for measuring internal consistency of subjective opinions [38, 39]. In particular, the method splits the annotations into two groups and then evaluates the correlation of average scores between the two via Pearson's r . This method intuitively tells us how well the first half of the annotations predict the ratings made by the second half of the annotators. That is, a high correlation suggests that the annotators tend to have a high degree of internal consistency. This approach has been widely used in psychological studies [40, 41], and a recent study used SHR for annotations about the offensiveness of online text [39].

The results show that the crowdsourced annotations achieved a moderate level of agreement of 0.561 and 0.63 for masculine and feminine traits, respectively. Other visual traits, such as Formal (0.565) and Professional (0.525), also have an acceptable level of agreement. On the other hand, there are visual traits of which the agreement rate is low, such as Ordinary (0.228) and Reassuring (0.245), suggesting that annotators see such traits in a distinct view. According to the rule-of-thumb interpretation of correlation coefficients, we exclude the traits with low agreement (≤ 0.3). For the target traits, we aggregate the five annotations on each image by transforming the responses into a value between 0 and 1 and averaging them.

3.2 Personal trait inference

We introduce a deep learning model that automatically infers personal traits from campaign photographs. Given an image I , the task aims to predict k -dimensional vectors, each of which corresponds to a visual trait value from 0 to 1. The task can be seen as a multi-output regression. As used in a recent study [22], a standard method is to train a neural network based on a backbone image encoder, which predicts a numeric trait from an image. If we apply the method to the target problem, k different models need to be trained to predict the corresponding k different values. Here, we propose a multitask regression model that predicts k traits simultaneously from a single convolutional neural network (CNN) backbone. Its underlying assumption is that the entire model would perform better when predicting k traits together because the backbone model could obtain generalized representation from the politician's photographs. The training objective is to minimize the sum of the differences in each predicted trait and truth value. Using the collective perception of visual traits obtained from the 8462 images, we trained a model that uses a CNN

Table 3 Cross-validated model performance

Trait	Standard	Proposed	Trait	Standard	Proposed
Agreeable	0.393	0.436	Formal	0.572	0.59
Ambitious	0.289	0.374	Friendly	0.45	0.496
Caring	0.364	0.442	Masculine	0.353	0.492
Communal	0.524	0.527	Maternal	0.26	0.436
Confident	0.353	0.422	Patriotic	0.448	0.432
Energetic	0.375	0.419	Professional	0.458	0.53
Feminine	0.361	0.535	Qualified	0.31	0.39

Table 4 Model performance measured on a separate test set

Trait	Test accuracy	Trait	Test accuracy
Agreeable	0.355	Formal	0.597
Ambitious	0.35	Friendly	0.442
Caring	0.372	Masculine	0.517
Communal	0.496	Maternal	0.448
Confident	0.406	Patriotic	0.409
Energetic	0.389	Professional	0.578
Feminine	0.55	Qualified	0.431

backbone to automatically annotate corresponding features for the remaining unlabeled images. Technical details of the method are available in the [Appendix](#).

Table 3 presents the prediction results of the standard and proposed multitask methods. The standard method is a CNN-based model that predicts each trait separately. Using 10-fold cross-validation of the 8462 annotated images, we calculated the average of Pearson's r across the 10 test sets. The results indicate that our model predicts the fourteen visual traits with reasonable performance; the maximum accuracy is 0.59 for Formal, and the minimum accuracy is 0.374 for Ambitious. Except for Communal and Patriotic, our method achieves higher accuracies than the standard CNN. The low accuracy in predicting ambitious and qualified traits could be explained by the highly abstract quality of the trait.

To evaluate the generalizability of the annotation method, we inferred the trait scores for a newly sampled (unlabeled) collection of 500 images by obtaining crowdsourced annotations. Table 4 presents the accuracy of our method evaluated on the new set. The results show that the model can achieve a similar level of accuracy for the unlabeled set as in the test performance (Table 3). While there are several differences, the test set performance is, on average, equivalent to the cross-validated performance, which suggests that the method can be used for the automatic annotation of politicians' visual traits. Accordingly, we inferred the fourteen trait scores on the entire set of 77,861 images and used them for the following analyses.

4 Correlation analysis on gender cues

In this section, we investigate how feminine and masculine traits are associated with other personal traits and visual features in politicians' photographs.

We first examine what the trait prediction model focuses on when inferring the stereotypical gender traits using gradient-weighted class activation mapping (Grad-CAM) [42]. In summary, it highlights important regions in an image for predicting the target concept (i.e., the gender stereotype). Figure 3 illustrates salient features identified for inferring the feminine and masculine traits by the CNN model. The color spectrum from blue (0) to

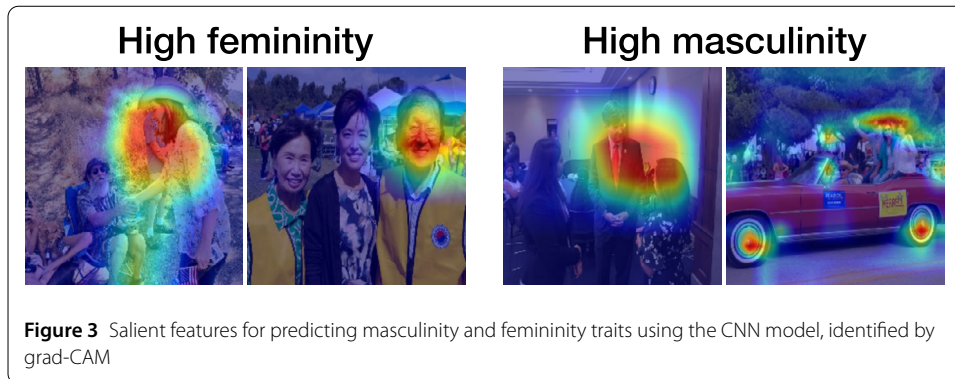


Figure 3 Salient features for predicting masculinity and femininity traits using the CNN model, identified by grad-CAM

Table 5 Correlation of the masculine- or feminine-related visual traits with the gender stereotype traits

Type	Trait	Correlation with masculinity	Correlation with femininity
Masculine-related	Formal	0.325	0.032
	Patriotic	0.325	0.073
	Professional	0.388	0.079
Feminine-related	Agreeable	0.088	0.571
	Caring	-0.03	0.617
	Communal	0.036	0.418
	Friendly	0.018	0.586
	Maternal	-0.466	0.89

red (1) indicates to what extent the model relies on image regions for making a prediction. The preliminary observation suggests that femininity may be formed around communal activities represented by handshakes and smiling faces, and masculinity may be conveyed through formal activities, which could be captured by politicians wearing a suit. The model attends to the smiling face of a man for inferring femininity in the second photograph with the presence of two women. To understand what constitutes gender stereotypes more systematically after the anecdotal observation by Grad-CAM, we examine correlations between visual traits and granular concepts.

4.1 Correlation with inferred visual traits

We first analyze whether each of the inferred visual traits is more related to masculinity or femininity. We consider a trait t to be masculine-related (or feminine-) if $r_{t \leftrightarrow \text{masculinity}}$ is statistically larger than $r_{t \leftrightarrow \text{femininity}}$ and vice versa, where $r_{x \leftrightarrow y}$ is Pearson's r between x and y . To measure the statistical significance of a difference, we convert each correlation coefficient into a z score using Fisher's r -to- z transformation and conduct an asymptotic z test for estimating statistical significance. In summary, the method measures the difference between $r_{t \leftrightarrow \text{masculinity}}$ and $r_{t \leftrightarrow \text{femininity}}$ considering $r_{\text{masculinity} \leftrightarrow \text{femininity}}$. We set the threshold of significance as 0.05. Refer to the textbook for more details [43].

Table 5 presents the list of visual traits identified as related to masculine and feminine traits. The traits of formal, professional, and patriotic are more correlated with masculine traits than the feminine traits. On the other hand, the Agreeable, Communal, and Friendly traits are more correlated with the feminine trait, suggesting that such correlated features constitute masculinity and femininity accordingly.

Table 6 Correlation of visual concepts with gender stereotype traits

Type	Trait	Correlation with masculinity	Correlation with femininity
Masculine	Official	0.232	-0.151
	Businessperson	0.211	-0.152
	Suit	0.192	-0.108
Feminine	Smile	-0.122	0.234
	Fun	-0.133	0.239
	Youth	-0.094	0.231

4.2 Correlation with granular visual concepts

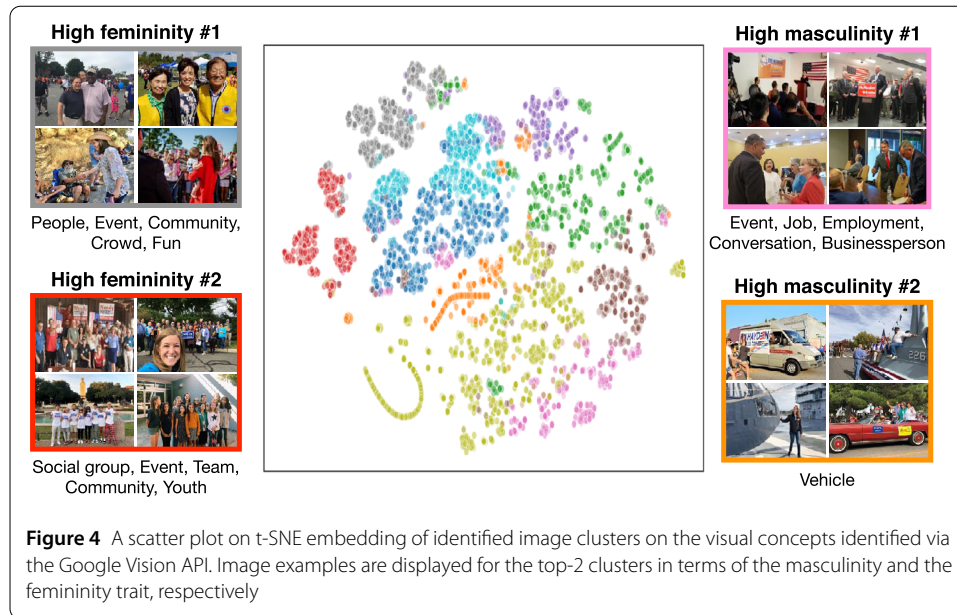
The visual traits estimated by the CNN model are abstract concepts, such that we do not know how the traits are composed from the visual details such as the presence of a particular object (e.g., Suit). We explore granular visual concepts that appear in images with high scores for masculinity and femininity.

We analyze the target photographs using the Google Vision API [44]. It helps understand what an image contains by automatically annotating the presence of previously identified image categories (e.g., crowd, tree) with a confidence score given by a pretrained machine learning model. After applying the API to the target dataset of 77,861 images, we examine masculine- and feminine-related concepts using the above method based on Fisher's r -to- z transformation. To distinguish the outcomes of the Google Vision API from the visual traits inferred from our CNN-based model, we refer to the vision API outputs as the *visual concept* for the rest of the paper.

Table 6 presents the correlation coefficients of masculinity- and femininity-related visual concepts among potential categories of the Google Vision API. The results show that the masculine trait is correlated with the visual concepts of Official (0.232), Businessperson (0.211), and Suit (0.192). The feminine traits are positively associated with the visual concepts of Smile, Fun, and Youth with correlations of 0.234, 0.239, and 0.231, respectively. Taken together with the results of Table 5, the above observation suggests that masculinity may be conveyed through politicians' formal and professional activities when they are wearing suits. In contrast, femininity may be formed around communal activities, where politicians may express themselves emotionally.

To further examine how femininity and masculinity are displayed differently in terms of visual concepts in the images, we conduct a clustering analysis using the visual concepts, which are the outcomes of the Google Vision API. The method aims at identifying image clusters from visual traits such as the presence of objects, and thus, it prevents an algorithm from focusing on prominent visual traits such as politician gender. Thus, we can better understand what kinds of visual concepts are more associated with each gender stereotype. We apply the k-means clustering algorithm to the 28 frequent traits. Method details are available in the [Appendix](#).

Figure 4 displays a scatter plot of two-dimensional t-stochastic neighbor embedding (t-SNE) embedding of V [45], which is used to visualize high-dimensional data in a low-dimensional space (usually 2D). For each cluster, we measure an average of the masculine and feminine traits inferred for each image; for the top-2 clusters in terms of masculinity and femininity, we display four sampled images balanced against the gender of the politicians in a bounding box. The color indicates the corresponding cluster displayed in the



scatter plot. We also present the names of the Google Vision concepts that appear in the centered image of each cluster below the corresponding box.

The scatter plot shows that overall, images are well clustered within visual concepts, implying that there may exist a shared set of visual concepts used for election campaigns. In the clusters of high gender-stereotypical traits, we observe that the corresponding visual concepts may contribute to each gender stereotype. In Masculinity#1, engaging in a formal event while wearing a formal suit appears as a prominent concept, which supports the high correlation of Formal and Suit with Masculinity in Table 5 and 6. We also discover an association of the visual concept of vehicle and masculinity in Masculinity#2, supported by other findings in the literature that cars are seen as masculine concepts [46]. In the clusters of high femininity, images tend to contain events in which people spend time together outside with positive sentiment, as visual concepts related to social groups, crowds, and fun appear prominently.

Overall, the results imply that crowdworkers (and our models) perceive masculinity as a formal and professional trait involving official events where people wearing formal suits are present. In contrast, the collective perception of femininity may be formed around a communal and friendly atmosphere involving people smiling. The results are congruent with the general perception of gender stereotypes found in the literature [8], which therefore supports the reliability of the annotations and the method for quantifying visual stereotypes.

5 Visual gender cues for electoral success

We now turn to the question of how election outcomes are correlated with perceived gender stereotypes in campaign photographs (RQ2). Previous research has analyzed gender cues by manual coding methods, but there have been inconsistent results, potentially due to small-sized samples. In this section, we tackle the question by analyzing the comprehensive dataset of the 2018 US election using the CNN deep learning model.

Table 7 Fitted OLS regression results using the inferred trait scores ($N = 554$)

	Model 1 (Masculinity)	Model 2 (Femininity)	Model 3 (Incumbency)
(Intercept)	0.019 (0.081)	0.566*** (0.053)	0.242*** (0.066)
Gender (Female = 1)	-0.107 (0.128)	-0.171 (0.092)	-0.14 (0.103)
Party (Democrats = 1)	0.038** (0.012)	0.018 (0.013)	0.053*** (0.01)
Incumbency			0.175*** (0.01)
Masculinity	0.976*** (0.166)		0.32* (0.138)
Femininity		-0.214 (0.16)	
Masculinity × Gender	0.38 (0.305)		0.405 (0.244)
Femininity × Gender		0.358 (0.227)	
Adjusted R-squared	0.102	0.009	0.426

Note: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

5.1 Regression analysis

To understand the potential role of visual stereotypes in election outcomes, we fit politician-level regression models using the ordinary least square method (OLS). Independent variables are politician-level features obtained by the average of image-level features for each politician, and the dependent variable is voting shares. On average, 140.54 photographs are aggregated to represent a politician's trait scores. We first set two models to test the role of masculinity (Model 1) and femininity (Model 2). We also add dummy variables for gender (Female = 1) and party membership (Democrats = 1) to control such effects. We have another model (Model 3) that includes incumbency as a control, which we will explain later in this section. All models have variation inflation factor (VIF) values lower than 5 for their independent variables, suggesting that the models have low risks of multicollinearity.

Table 7 presents the regression models' estimated coefficients with standard errors in parentheses. In Model 1, we observe the statistical significance of the masculine trait with a positive coefficient of 0.976. In contrast, in Model 2, feminine traits are not statistically significant. Combined with the low adjusted R-squared value of 0.009, the femininity variable's insignificance in Model 2 suggests that expressing feminine visual traits may be less likely to affect outcomes in the target election. To further evaluate visual masculinity's role, we set Model 3 by adding another control variable indicating whether a politician runs as an incumbent. Incumbency has been considered one of the key determinants for election outcomes in the literature [47, 48]. Thus, the variable can function as a strong control for testing the effects of visual masculinity. While incumbency is the most significant variable in the model, the masculinity variable is positively associated with voting shares with significance ($p < 0.05$).

As a robustness check, we conduct regression analyses using the annotation dataset. Table 8 presents the results of three regression models that take the trait scores of 544 politicians. Using the 8964 images with the annotated gender trait scores by crowdworkers, we constructed the politician-level data by averaging the scores of the corresponding

Table 8 Fitted OLS regression results using the annotation data ($N = 544$)

	Model 1 (Masculinity)	Model 2 (Femininity)	Model 3 (Incumbency)
(Intercept)	0.3*** (0.055)	0.467*** (0.031)	0.377*** (0.043)
Gender (Female = 1)	-0.02 (0.075)	-0.003 (0.071)	-0.073 (0.059)
Party (Democrats = 1)	0.027 (0.013)	0.007 (0.013)	0.048*** (0.01)
Incumbency			0.183*** (0.01)
Masculinity	0.357*** (0.098)		0.032 (0.079)
Femininity		0.118 (0.108)	
Masculinity × Gender	0.142 (0.183)		0.211 (0.144)
Femininity × Gender		-0.108 (0.163)	
Adjusted R-squared	0.05	0.013	0.414

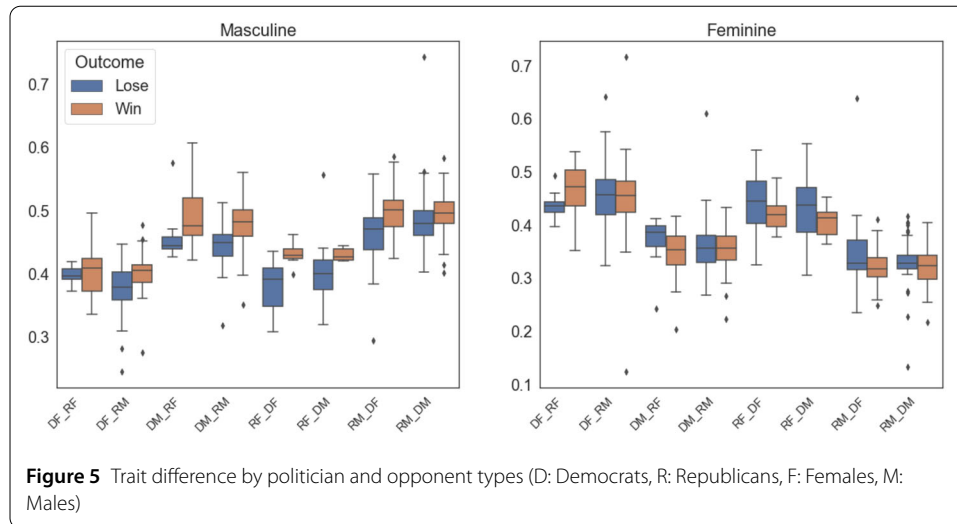
Note: * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$.

images of each politician. On average, 16.48 images were aggregated to represent the perceived gender traits in campaign photographs of a politician. In Model 1 and Model 2, we observe the patterns congruent with the findings based on the inferred scores by the proposed model (Table 7). The masculinity variable was associated with electoral success with a strong significance ($p < 0.001$), but the femininity variable was not correlated with electoral success. The statistical significance of the masculinity trait observed in the analysis supports the generalizability of the proposed method of automatic inference. The significance of the masculinity variable disappears with incumbency as an additional control (Model 3). Note that the masculinity variable was significant with incumbency in Table 7. The adjusted R^2 of the model with the human annotation was also smaller than that of the model with the model prediction. We suspect that the limited number of images per politician in the annotation set might have led to a higher variance in measurements aggregated per politician, whereas the CV model based prediction was obtained from all the images that each politician posted. This again highlights the effectiveness of the proposed inference method.

5.2 Varying association by gender and party

The regression analysis found a positive correlation of masculinity with electoral success even after controlling for the effects of gender, party, and incumbency. We further examine the trend by dissecting the data according to gender and party of the two politicians who run a race (RQ3). We assume that the role of visual gender stereotypes can be different according to the combination.

Figure 5 demonstrates the varying patterns of gender-stereotypical visual traits. The x-axis presents eight election types according to the party and gender combination of target politicians (who expressed such visuals) and their opponents. In the axis label, the first two characters indicate the target politician type, and the last two present the opponent type. The y-axis indicates the distribution of visual features of target politicians who belong to each race type. We also compare the distribution of stereotypical features of winners



and losers to understand how electoral success is associated with stereotypical gender expressions on social media.

Here, we make three main observations. First, expressing visual masculinity in campaign photographs is positively correlated with winning the election in most cases. Highly significant differences between winner and loser groups are observed for Democrat females against Republican males ($p < 0.001$), Democrat males against Republican males ($p < 0.001$), Republican females against Democrat females ($p < 0.01$), and Republican males against Democrat females ($p < .001$). The positive effects of masculinity on electoral success are prominent for election races featuring Republican males or Democrat females. Second, the visual femininity feature is negatively associated with electoral success in several cases, and the lowest p value is observed for Republican males against Democrat females ($p < 0.01$). The gender and party combination is where election outcomes are associated with masculinity (positive) and femininity (negative) in the most stereotypical way. Third, we observe an exception in that the visual traits may operate differently. In the intragender race of Democrat females against Republican females, the visual femininity trait is positively associated with electoral success with a weak significance ($p = 0.09$).

To summarize, we observe the positive association of visual masculinity with electoral success for different combinations of gender and party membership of politicians in an election. The findings support the positive role of stereotypical presentations of masculinity in election campaigns, which is aligned with previous research [9, 10]. We did not observe negative effects from female politicians expressing masculinity, which contradicts the observations in previous studies [16, 17]. Femininity is negatively associated with success in general, but a flipped correlation is observed for Democrat females running against Republican females. The finding implies that the effects of visual gender stereotypes can be contingent on the gender and party of the politician and their opponent.

6 Discussion and conclusion

This study investigated the effects of gender cues displayed through social media images for political campaigns. Politicians have intentionally employed gender-stereotypical traits in professional social media images to appeal to voters. However, previous studies have mainly relied on manual methods for analyzing the effects of visual cues, leading to

conflicting observations. To address the weakness, using a total of 77,861 photographs shared by the 554 political candidates in campaigns for the 2018 US general election, we presented a multitask deep learning method that learns visual gender-stereotypical traits from a set of crowdsourced perception ratings (RQ1). Annotation quality and performance evaluation results suggest that the participants have internal consistency on assessing visual traits and that the deep learning model can infer collective perception with reasonable accuracy. Accordingly, we inferred the traits for unlabeled data and employed the whole sample and labels for the subsequent analyses, thereby allowing us to draw a bigger picture while overcoming the limited scale of analyses relying only on manual annotation. The analysis suggests what the constituents of visual gender stereotype are; masculinity may be formed around formal activities, such as wearing a formal suit or giving a formal speech. On the other hand, femininity may be expressed through engaging in outdoor social activities and expressing emotions through smiles. The correlations, which are congruent with general perceptions of gender stereotypes [8], suggest that our method captures gender cues reasonably well.

Next, we examined how visual gender traits are associated with electoral success (RQ2). From regression analyses, we made an observation that supports the importance of masculinity for electoral success. The masculinity variable is positively associated with voting shares, even after controlling for strong control variables such as gender, party, and incumbency. This observation is congruent with previous studies on the positive role of visual masculinity in election campaigns [9, 10]. We further examined how the correlation of visual gender stereotypes for election outcomes varies according to the gender and party combinations of the two politicians in an election race (RQ3). The analysis not only supports the positive role of masculinity but also provides a novel observation for the role of visual gender stereotypes. Visual femininity played a positive role in the intragender race of Democrat female candidates against Republican female politicians. The complicated effects of gender cues for females support the reflection on the challenges of women candidates in image managements during elections [49, 50].

This study could make a contribution to the research community by providing the deep learning method used for capturing the crowdsourced perception of visual gender-stereotypical traits. The method could serve as a methodological reference for future research on visual communication. We are releasing the inference code alongside the model checkpoint to facilitate broader usage.² The findings in the analyses add an empirical understanding of the potential role of gender traits in election campaigns discussed in the literature. Furthermore, we believe this study has general implications for research on computational social science that aims to estimate personal traits, human perception, and bias from online photographs. The method used for annotation and deep learning approach could be tested in a broader context.

This study bears several limitations with future directions. First, the deep learning approach learns visual patterns based on the perceptions of crowdworkers, and hence, the model can also capture their underlying biases. This study is aimed at capturing the “perceived” gender stereotypes and thus learning hidden biases is intended. Second, this study only focuses on a single election year in the US, and hence the findings should be carefully interpreted. Unlike the studies aiming to build a prediction model [51], our analysis seeks

²<https://github.com/ssu-humane/visual-gender-trait-politician>

to understand the role of visual gender stereotypes in successful election campaigns. The methodological foundation we built in this study could contribute to future studies on the politics in the US and other countries. It would be exciting to see how gender stereotypes are formed across different cultures from online visual data, as a recent study found a cultural effect on the perception of politicians' traits [52]. Third, the analysis is based on observational data, and thus, the correlations in the analysis do not imply causality. Future studies could examine its causal relationship using difference-in-difference estimation or propensity score matching methods [53]. Fourth, while the proposed method showed reasonable performance in inferring key traits such as masculinity and femininity, its prediction could be inaccurate for some traits such as Ambitious and Qualified. Users should be aware of the prediction errors in a downstream analysis, which could misrepresent the real patterns of perceived gender traits displayed through campaign photographs. Manual validation on a small set of samples might be necessary for a reliable analysis. Future studies could boost the performance by constructing a more extensive set of annotations or adopting more recent deep learning and computer vision technologies.

Appendix

A.1 Technical details

A.1.1 Automatic annotation

Using a CNN model f_{cnn} , our method extracts the d_{img} -dimensional features of I and projects the values into a k -dimensional output through a fully-connected layer. Specifically, the model is expressed as

$$\begin{aligned} \mathbf{h}_{\text{img}} &= f_{\text{cnn}}(\mathbf{I}; \theta_{\text{cnn}}) \in \mathcal{R}^{d_{\text{img}}}, \\ \hat{\mathbf{y}} &= \sigma(f_{\text{fc}}(\mathbf{h}_{\text{img}}; \theta_{\text{fc}})) \in \mathcal{R}^k, \end{aligned} \quad (1)$$

where $\hat{\mathbf{y}}$ is the k -dimensional vector of predicted trait values, θ_{cnn} and θ_{fc} are learnable parameters, and $\sigma(\cdot)$ is sigmoid nonlinearity that transforms the model output into values between 0 and 1.

The training objective is to minimize \mathcal{L} , which is defined as

$$\mathcal{L} = \sum_{i=1}^N \sum_{j=1}^k \text{MSE}(\hat{y}_{ij}, y_{ij}), \quad (2)$$

where $\text{MSE}(\text{predicted}, \text{truth})$ is the mean-squared error function and N is the number of training instances in a minibatch.

To train the model on the annotated visual traits, we employ Res-Net with 34 layers [54], which has shown robust performance across various image classification tasks. In particular, we set f_{cnn} to the ResNet model until the last hidden layer, which was pretrained on the ImageNet data. We resize \mathbf{I} to the 224×224 resolution before feeding it into the model; hence, D_{img} is 512. We set k to 14 as the number of target traits, batch size to 40, and the adam optimizer is used to minimize \mathcal{L} for every batch. We optimize the performance of the proposed model by varying the backbone architecture. Table A1 presents the 10-fold cross-validated model performance, measured by the Pearson's r . On average,

Table A1 Cross-validated model performance by varying the backbone architecture

Trait	ResNet-34	ResNet-50	Vit-B-16	Trait	ResNet-34	ResNet-50	Vit-B-16
Agreeable	0.436	0.431	0.261	Formal	0.59	0.581	0.382
Ambitious	0.374	0.37	0.232	Friendly	0.496	0.489	0.298
Caring	0.442	0.433	0.256	Masculine	0.492	0.492	0.18
Communal	0.527	0.531	0.282	Maternal	0.436	0.442	0.195
Confident	0.422	0.41	0.269	Patriotic	0.432	0.442	0.171
Energetic	0.419	0.42	0.27	Professional	0.53	0.519	0.328
Feminine	0.535	0.537	0.226	Qualified	0.39	0.383	0.233

Res-Net with 34 layers (0.466) is on par with Res-Net with 50 layers (0.463) and outperforms the vision transformer (0.256) [55]. We suspect that the limited size of the labeled dataset makes it challenging to exploit the full capacity of the larger models.

The standard approach uses k different CNN models where each model predicts a single trait value. Concretely, for each model, k in Equation (1) becomes 1, and \mathcal{L} is a single MSE for a corresponding trait. We compare this standard approach against the proposed method that uses the multitask learning objective in Table 3.

A.1.2 Clustering

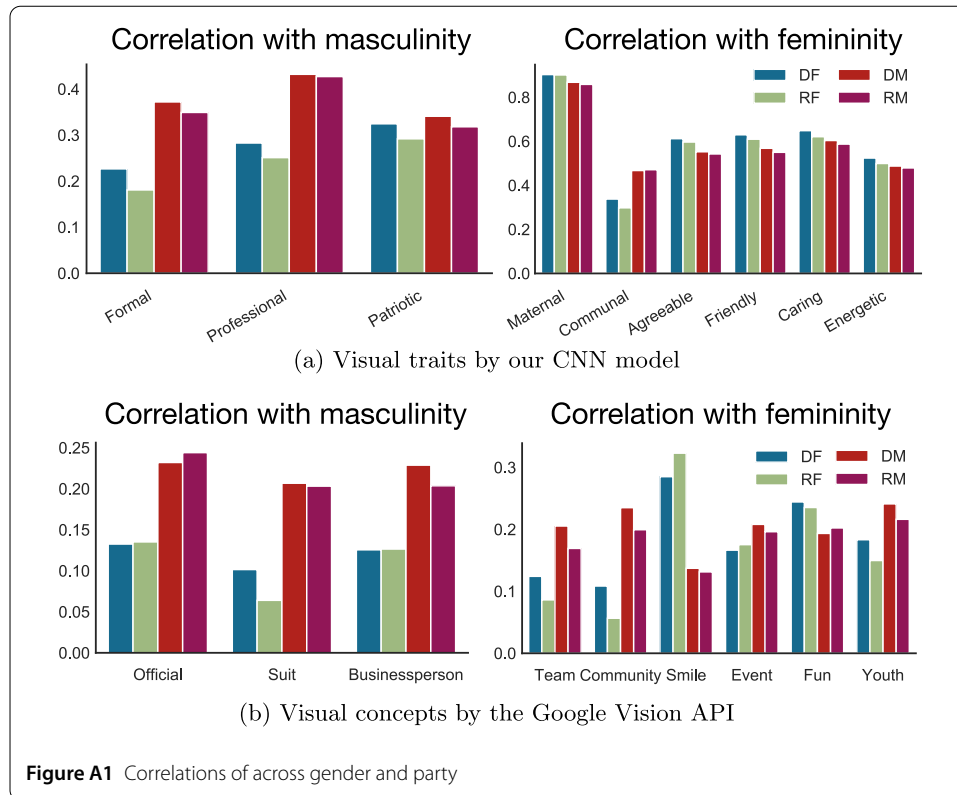
For each image I , we construct a d_{vis} -dimensional vector \mathbf{v}_I only considering popular concepts of the API that appear in more than 10% of the target dataset. In our dataset, d_{vis} is empirically determined to be 28. Using $V \in \mathcal{R}^{77,861 \times 28}$, we run the k-means++ clustering algorithm, which improves the standard k-means by assigning initial centroids based on the underlying data distribution. We set the optimal number of clusters ($k = 10$) by the elbow method.

A.2 Gender and party differences

Masculinity and femininity are concepts that are perceived to have characteristics related to each gender. Then, would ways of expressing gender stereotypes remain the same across different gender and party memberships?

To answer these questions, we split the data according to the gender and party membership of politicians and measure the correlation of inferred visual traits with each gender stereotype. Figure A1 demonstrates Pearson's r of visual traits and concepts for masculinity and femininity. For brevity, we only present the correlations that are identified as masculine- or feminine-related features in the aggregated analyses in Tables 5 and 6.

Here, we can observe that the list of correlated features is similar across the different gender and party groups, while the degree of correlation is different. From the correlations with the masculinity trait, we identify that images shared by male politicians tend to exhibit a higher value of correlation for each of the visual concepts of formal, professional, official, suit, and businessperson than those shared by female politicians. The high correlation suggests that the masculinity-related visual concepts may be perceived more keenly when expressed by male politicians. On the other hand, femininity-related visual concepts are more correlated with feminine traits when expressed by females. However, there are exceptions for traits in images of male politicians that are more correlated with femininity, such as communal, team, and community. We do not observe any significant difference across party memberships.



A.3 Evaluation resembling election settings

We test whether the gender-stereotyped self-presentation of politicians predicts election outcomes. To this end, we conduct prediction experiments that resemble the election settings. In particular, we split the set of 251 target positions (e.g., governor in New York) into five and use 5-fold cross-validation based on the position-level split. This step ensures that every politician has an opponent in the election in the same fold. For predicting outcomes, we compare the output scores of a model for each politician in the pair. If a model assigns a higher score for the winner than for the loser, the prediction is considered correct. Prediction accuracy is measured by dividing the number of correctly predicted positions by the total number of positions in each test fold and averaging them.

We compare the performance of the two different approaches. First, we employ a trait-based model that predicts the probability of winning an election given a politician with the averaged 14 visual trait scores. We use a logistic regression classifier, which is trained by a pair of trait score vector vectors and binary election outcomes. Second, we utilize a CNN-based model that predicts the winning probability from raw images. In particular, we train an image-level CNN network to minimize the cross-entropy loss of a binary election outcome and to set the sigmoid output from 0 to 1. For making a prediction for a politician, we measure an average of output scores of the images of the politician and compare that with the corresponding value of the race's counterpart. We employ ResNet-34 pretrained on the ImageNet data by replacing the last layer with a binary classifier and fine-tuning the whole network. To prevent the model from training on only the patterns of politicians who shared many images, we train the model by using up to 300 images for each politician. For inference, all images are employed for each target politician.

Table A2 Election outcome prediction accuracy (paired binary classification)

	Test accuracy
Direct model	0.724
Trait-based model	0.739

Table A2 presents the classification accuracy of the image-based models. The results show that the image-based direct model achieves a fair accuracy of 0.724. The trait-based model exhibits a slightly higher accuracy of 0.739 than the direct model. The above observation suggests that social media images shared by politicians have a certain degree of predictive power for election outcomes, and gender stereotypical traits embedded in images may play a role.

Acknowledgements

We are grateful for Danni Chen for the assistance of data collection and preliminary analysis.

Funding

This research was supported by the National Research Foundation of Korea (2021R1F1A1062691), the Institute of Information & Communications Technology Planning & Evaluation (IITP-2023-RS-2022-00156360), and the National Science Foundation SMA-1831848.

Abbreviations

SHR, split-half reliability; CNN, convolutional neural network; Grad-CAM, gradient-weighted class activation mapping; API, application programming interface; t-SNE, t-stochastic neighbor embedding; OLS, ordinary least square; VIF, variation inflation factors; ResNet, residual neural network.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available due to the copyright and privacy issues but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author contributions

JJ conceptualized this research and collected the data. KP developed the method and analyzed the data. All authors contributed to the manuscript's writing and editing. All authors read and approved the final manuscript.

Received: 27 March 2023 Accepted: 6 July 2023 Published online: 23 August 2023

References

1. Newman N, Fletcher R, Schulz A, Andi S, Robertson CT, Nielsen RK (2021) Reuters institute digital news report 2021. Reuters Institute for the. Study Journal
2. Liebhart K, Bernhardt P (2017) Political storytelling on Instagram: key aspects of Alexander van der Bellen's successful 2016 presidential election campaign. *Media Commun* 5(4):15–25
3. Joo J, Li W, Steen FF, Zhu S-C (2014) Visual persuasion: inferring communicative intents of images. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 216–223
4. Grabe ME, Bucy EP (2009) *Image bite politics: news and the visual framing of elections*. Oxford University Press, London
5. Haim M, Jungblut M (2021) Politicians' self-depiction and their news portrayal: evidence from 28 countries using visual computational analysis. *Polit Commun* 38(1–2):55–74
6. Schmuck D, Matthes J (2017) Effects of economic and symbolic threat appeals in right-wing populist advertising on anti-immigrant attitudes: the impact of textual and visual appeals. *Polit Commun* 34(4):607–626
7. Johns R, Shephard M (2011) Facing the voters: the potential impact of ballot paper photographs in British elections. *Polit Stud* 59(3):636–658
8. Flicker E (2013) Fashionable (dis-) order in politics: gender, power and the dilemma of the suit. *Int J Media Cultur Polit* 9(2):183–201
9. Koenig AM, Eagly AH, Mitchell AA, Ristikari T (2011) Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychol Bull* 137(4):616
10. Rosenwasser SM, Dean NG (1989) Gender role and political office: effects of perceived masculinity/femininity of candidate and political office. *Psychol Women Q* 13(1):77–85

11. Eagly AH, Karau SJ (2002) Role congruity theory of prejudice toward female leaders. *Psychol Rev* 109(3):573
12. Bystrom DG, Robertson TA, Banwart MC (2001) Framing the fight: an analysis of media coverage of female and male candidates in primary races for governor and us senate in 2000. *Am Behav Sci* 44(12):1999–2013
13. Bauer NM (2015) Emotional, sensitive, and unfit for office? Gender stereotype activation and support female candidates. *Polit Psychol* 36(6):691–708
14. Dolan K (2010) The impact of gender stereotyped evaluations on support for women candidates. *Polit Behav* 32(1):69–88
15. McGregor SC, Lawrence RG, Cardona A (2017) Personalization, gender, and social media: gubernatorial candidates' social media strategies. *Inf Commun Soc* 20(2):264–283
16. Bauer NM, Carpinella C (2018) Visual information and candidate evaluations: the influence of feminine and masculine images on support for female candidates. *Polit Res Q* 71(2):395–407
17. Carpinella C, Bauer NM (2019) A visual analysis of gender stereotypes in campaign advertising. *Polit Groups Ident* 1–18
18. Chatterjee A, Gupta U, Chinnakotla MK, Srikanth R, Galley M, Agrawal P (2019) Understanding emotions in text using deep learning and big data. *Comput Hum Behav* 93:309–317. <https://doi.org/10.1016/j.chb.2018.12.029>
19. Zhu J, Luo J, You Q, Smith JR (2013) Towards understanding the effectiveness of election related images in social media. In: 2013 IEEE 13th international conference on data mining workshops. IEEE Press, New York, pp 421–425
20. Joo J, Steen FF, Zhu S-C (2015) Automated facial trait judgment and election outcome prediction: social dimensions of face. In: Proceedings of the IEEE international conference on computer vision, pp 3712–3720
21. Wang Y, Li Y, Luo J (2016) Deciphering the 2016 us presidential campaign in the Twitter sphere: a comparison of the trumpists and clintonists. In: Tenth international AACL conference on web and social media
22. Chen D, Park K, Joo J (2020) Understanding gender stereotypes and electoral success from visual self-presentations of politicians in social media. In: Joint workshop on aesthetic and technical quality assessment of multimedia and media analytics for societal trends, pp 21–25
23. Won D, Steinert-Threlkeld ZC, Joo J (2017) Protest activity detection and perceived violence estimation from social media images. In: Proceedings of the 25th ACM international conference on multimedia. ACM, New York, pp 786–794
24. Zhang H, Pan J (2019) Casm: a deep-learning approach for identifying collective action events with text and image data from social media. *Sociol Method* 49(1):1–57
25. Huang X, Kovashka A (2016) Inferring visual persuasion via body language, setting, and deep features. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp 73–79
26. Kim Y, Kim JH (2018) Using computer vision techniques on Instagram to link users' personalities and genders to the features of their photos: an exploratory study. *Inf Process Manag* 54(6):1101–1114. <https://doi.org/10.1016/j.ipm.2018.07.005>
27. Xi N, Ma D, Liou M, Steinert-Threlkeld ZC, Anastasopoulos J, Joo J (2020) Understanding the political ideology of legislators from social media images. In: Proceedings of the international AACL conference on web and social media, vol 14, pp 726–737
28. Steinert-Threlkeld ZC (2019) The future of event data is images. *Sociol Method* 49(1):68–75
29. Zhao S, Yao H, Gao Y, Ji R, Xie W, Jiang X, Chua T-S (2016) Predicting personalized emotion perceptions of social images. In: Proceedings of the 24th ACM international conference on multimedia, pp 1385–1394
30. Liu L, Preotiuc-Pietro D, Samani ZR, Moghaddam ME, Ungar LH (2016) Analyzing personality through social media profile picture choice. In: ICWSM, pp 211–220
31. Skowron M, Tkalcic M, Ferwerda B, Schedl M (2016) Fusing social media cues: personality prediction from Twitter and Instagram. In: Proceedings of the 25th international conference companion on world wide web, pp 107–108
32. Sigaki HY, Perc M, Ribeiro HV (2018) History of art paintings through the lens of entropy and complexity. *Proc Natl Acad Sci* 115(37):8585–8594
33. Mo CH (2015) The consequences of explicit and implicit gender attitudes and candidate quality in the calculations of voters. *Polit Behav* 37(2):357–395
34. Ditonto TM, Hamilton AJ, Redlawsk DP (2014) Gender stereotypes, information search, and voting behavior in political campaigns. *Polit Behav* 36(2):335–358
35. Grabe ME, Bucy EP (2010) Image bite analysis of political visuals. *Sourceb Polit Commun Res: Methods Meas Anal Techniq*, 209–237
36. Mattes K, Spezio M, Kim H, Todorov A, Adolphs R, Alvarez RM (2010) Predicting election outcomes from positive and negative trait assessments of candidate images. *Polit Psychol* 31(1):41–58
37. Bauer NM (2017) The effects of counterstereotypic gender strategies on candidate evaluations. *Polit Psychol* 38(2):279–295
38. Isola P, Xiao J, Torralba A, Oliva A (2011) What makes an image memorable? In: CVPR. IEEE Press, New York, pp 145–152
39. Hada R, Sudhir S, Mishra P, Yannakoudakis H, Mohammad SM, Shutova E (2021) Ruddit: Norms of offensiveness for English Reddit comments. In: Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing. Association for Computational Linguistics, pp 2700–2717. <https://doi.org/10.18653/v1/2021.acl-long.210>
40. Liaw S-S (2002) An Internet survey for perceptions of computers and the world wide web: relationship, prediction, and difference. *Comput Hum Behav* 18(1):17–35. [https://doi.org/10.1016/S0747-5632\(01\)00032-2](https://doi.org/10.1016/S0747-5632(01)00032-2)
41. Constantine MG, Ponterotto JG (2006) Evaluating and selecting psychological measures for research purposes. *Psychol Res Handb: Guide Grad Stud Res Assist* 2:104–113
42. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D (2017) Grad-cam: visual explanations from deep networks via gradient-based localization. In: Proceedings of the IEEE international conference on computer vision, pp 618–626
43. Steiger JH (1980) Tests for comparing elements of a correlation matrix. *Psychol Bull* 87(2):245
44. Google: Google Cloud Vision API. <https://cloud.google.com/vision>. [Online; accessed 11-Jan-2022]
45. Maaten L, Hinton G (2008) Visualizing data using t-sne. *J Mach Learn Res* 9:2579–2605
46. Walker L, Butland D, Connell RW (2000) Boys on the road: masculinities, car culture, and road safety education. *J Men's Stud* 8(2):153–169

47. Gerber A (1998) Estimating the effect of campaign spending on senate election outcomes using instrumental variables. *Am Polit Sci Rev* 401–411
48. Abramowitz AI (1991) Incumbency, campaign spending, and the decline of competition in us house elections. *J Polit* 53(1):34–56
49. Lawless JL (2009) Sexism and gender bias in election 2008: a more complex path for women in politics. *Polit Gend* 5(1):70–80
50. Haraldsson A, Wängnerud L (2019) The effect of media sexism on women's political ambition: evidence from a worldwide study. *Femin Media Stud* 19(4):525–541
51. You Q, Cao L, Cong Y, Zhang X, Luo J (2015) A multifaceted approach to social multimedia-based prediction of elections. *IEEE Trans Multimed* 17(12):2271–2280
52. Lin C, Adolphs R, Alvarez RM (2017) Cultural effects on the association between election outcomes and face-based trait inferences. *PLoS ONE* 12(7):0180837
53. Pearl J, Mackenzie D (2018) *The book of why: the new science of cause and effect*. Basic Books, New York
54. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
55. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S et al (2021) An image is worth 16x16 words: transformers for image recognition at scale. In: *International conference on learning representations*

Publisher's Note

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

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
