Identifying Media Bias With Computer Vision

Yilang Peng
University of Pennsylvania, pengonelang@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations
Part of the Communication Commons, Political Science Commons, and the Psychology Commons

Recommended Citation
https://repository.upenn.edu/edissertations/3508

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/3508
For more information, please contact repository@pobox.upenn.edu.
Identifying Media Bias With Computer Vision

Abstract
Although visual content prevails in the digital media environment, previous scholarship that attempts to detect bias and stereotypes in media content has mostly focused on textual data. Meanwhile, recent advances in computer vision have made the analysis of visual data on a large scale possible. Drawing theoretical insights from media bias, social cognition, visual persuasion, and gender studies literature, this dissertation investigates how various computer vision techniques—such as facial recognition, emotion detection, and computational aesthetics—can help us better analyze media bias in visual representations of politicians. In particular, study 1 examines partisan bias in media coverage of presidential candidates in the 2016 presidential election. Study 2 analyzes gender bias in politicians’ self-presentations on Instagram. In addition, this dissertation also hopes to illuminate the promises and caveats of using computer vision algorithms as data analysis tools.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Communication

First Advisor
Sandra González-Bailón

Subject Categories
Communication | Political Science | Psychology
IDENTIFYING MEDIA BIAS WITH COMPUTER VISION

Yilang Peng

A DISSERTATION

in

Communication

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

________________________

Sandra González-Bailón, Associate Professor of Communication

Graduate Group Chairperson

________________________

Guobin Yang, Grace Lee Boggs Professor of Communication and Sociology

Dissertation Committee

Michael X. Delli Carpini, Professor of Communication

Jessa Lingel, Assistant Professor of Communication
IDENTIFYING MEDIA BIAS WITH COMPUTER VISION

COPYRIGHT

2019

Yilang Peng
ACKNOWLEDGMENT

First, I would like to express my gratitude to my advisor Dr. Sandra González-Bailón for her intellectual guidance, constant support, and valuable advice on my career. I also thank my dissertation committee members Dr. Michael X. Delli Carpini and Dr. Jessa Lingel for their thoughtful feedback on my scholarly works.

I thank my former committee member Dr. Paul Messaris for his insights and mentoring. His scholarship and pedagogy will always be in my heart and our memories. I also thank Dr. Guobin Yang, Dr. John B. Jemmott III, and other Annenberg faculty members whose mentoring and teaching have shaped my scholarship.

I also own my deepest thanks to my family, friends, and colleagues. I thank Sijia Yang and Tian Yang especially for their emotional support when I was on the job market. I would also like to thank my 2014 cohorts for their encouragement and company. I hope we all have bright futures ahead of us. I thank members of the DiMeNet for their comments on my various research projects. I am grateful to many Annenberg alumni who have helped me on various occasions. I thank Jiaying Liu in particular for her suggestions on my job search and offer negotiations.

I thank the amazing Annenberg staff for their support.
ABSTRACT
IDENTIFYING MEDIA BIAS WITH COMPUTER VISION

Yilang Peng
Sandra González-Bailón

Although visual content prevails in the digital media environment, previous scholarship that attempts to detect bias and stereotypes in media content has mostly focused on textual data. Meanwhile, recent advances in computer vision have made the analysis of visual data on a large scale possible. Drawing theoretical insights from media bias, social cognition, visual persuasion, and gender studies literature, this dissertation investigates how various computer vision techniques—such as facial recognition, emotion detection, and computational aesthetics—can help us better analyze media bias in visual representations of politicians. In particular, study 1 examines partisan bias in media coverage of presidential candidates in the 2016 presidential election. Study 2 analyzes gender bias in politicians’ self-presentations on Instagram. In addition, this dissertation also hopes to illuminate the promises and caveats of using computer vision algorithms as data analysis tools.

Keywords media bias, visual bias, gender, social cognition, computer vision, computational social science
Table of Contents

CHAPTER 1 INTRODUCTION 1
The many facets of bias 3
Visual bias 5
Applying computer vision in communication research 7
Theoretical and practical contributions 9

CHAPTER 2 MEASURES OF MEDIA BIAS 13
Partisan media bias 13
  Audience-based approach 17
  Content-based approach 18
Gender bias 19
  Objectification theory 20
  Gender stereotypes 21
  Media coverage of women politicians 23
Visual bias 26
  Subjective rating approach 27
  Objective feature checklist approach 28

CHAPTER 3 PROPOSED APPROACH 33
An overview of computer vision methods 34
  Supervised learning and convolutional neural network 34
  Unsupervised learning and image clustering 39
  Computer vision libraries and applications 41
  Computational aesthetics 49
Coding favorability 50
  Crowdsourcing 50
  Dimensions of favorability 52

CHAPTER 4 STUDY 1: PARTISAN MEDIA BIAS 56
Theoretical framework 57
  Measuring partisan media bias in visual content 57
  Selection of visual features 60
  Visual bias across liberal and conservative media outlets 64
  Favorability and its dimensions 66
Method 68
Results 75
Discussion 81

CHAPTER 5 GENDER BIAS IN POLITICIANS’ SELF-PRESENTATIONS 86
Theoretical framework 87
List of Tables

Table 2.1 Definitions of partisan media bias in prior research .................................................. 16
Table 2.2 Summary and illustration of frequently used visual bias coding criteria........ 32
Table 3.1 Exemplar images in categories from Zhang and Peng (2019) ......................... 41
Table 3.2 A comparison of four object recognition services................................................. 43
Table 3.3 A comparison of facial attributes provided by different face detection APIs ... 46
Table 4.1 Correlations between human-perceived and computer vision services-detected facial expressions ........................................................................................................ 73
Table 4.2 Distribution of visual features across candidates and media outlets............... 75
Table 4.3 Principal component analysis of perceived traits................................................... 79
Table 5.1 Examples of Clarifai tags ......................................................................................... 98
Table 5.2 Topic modeling solution.......................................................................................... 103
Table 6.1 A comparison between manually coded results and facial recognition results 131
List of Figures

Figure 1.1 The front page of Daily Mail on March 28, 2017 .................................................. 7
Figure 3.1 An artificial neural network with an input layer, two hidden layers, and an output layer .......................................................................................................................... 35
Figure 3.2 Illustration of the VGG16 model .............................................................................. 36
Figure 3.3 Feature visualization of a CNN trained on ImageNet ............................................. 37
Figure 3.4 Two proposed approaches of unsupervised leaning for visual data .................... 40
Figure 3.5 Illustration of (a) face rectangle and (b) face landmarks .................................... 44
Figure 3.6 Confusion matrix tables between human-coded emotions (columns) and computer vision-calculated emotions (rows) for two computer vision APIs ............... 48
Figure 3.7 Social groups mapped on the stereotype content model .................................... 54
Figure 4.1 Facial orientation in pitch, roll, and yaw angles ...................................................... 61
Figure 4.2 Data preparation procedures .................................................................................. 70
Figure 4.3 Distribution of visual features across media outlets and candidates ................. 76
Figure 4.4 Effects of visual features on perceived slant, valence, and dominance .......... 81
Figure 5.1 Calculation of the face-ism index (a/b) ................................................................. 95
Figure 5.2 Effects of politicians’ characteristics on image features ........................................ 108
Figure 5.3 Effects of politicians’ characteristics on visual topic distributions ................. 110
Figure 5.4 Effects of post features on audience engagement in multilevel analyses .......... 113
Figure 5.5 Effects of post features on audience engagement by gender .............................. 114
Figure 6.1 Examples of non-frontal, blocked, and low-quality faces in study 1 ............... 127
Figure 6.2 Examples of pictures that have the politician’s face but the face is not complete .................................................................................................................................... 130
Figure 6.3 Examples of pictures featuring the politician’s complete face but the face was not recognized by the algorithm ................................................................................................. 130
Figure 6.4 Illustrations of facial action units provided in OpenFace ....................................... 133
Figure 6.5 Examples of images judged as high and low on valence and dominance ......... 134
Figure 6.6 Heat maps in a CNN that visualize important regions for predictions ............ 137
Chapter 1 Introduction

Communication research has long analyzed the forms and consequences of media bias (Hackett, 1984; D’Alessio & Allen, 2000; Groeling, 2013). In today’s media landscape, an increasing number of partisan media outlets are purposely producing ideologically slanted content to appeal to consumers of certain ideological orientations, discounting traditional journalistic norms of objectivity and balance (Iyengar & Hahn, 2009; Groeling, 2013). Media content might also incorporate biased and stereotyped portrayals of different social groups, further intensifying viewers’ prejudices and negative attitudes towards disadvantaged individuals (McClure et al., 2011; Kay et al., 2015). In response to the prevalence of bias in media content and its potential effects, scholars have made various attempts to automatically detect bias in media content in the hope of exposing media consumers to more balanced and diverse viewpoints (Garimella et al., 2017; Munson & Resnick, 2010; Park et al., 2009).

Previous studies that systematically quantify media bias have predominantly focused on textual data (Gentzkow & Shapiro, 2010; Groseclose & Milyo, 2005), paying less attention to an important part of media content—visuals (Groeling, 2013). This lack of attention to visual content is especially conspicuous today: visual content is ever more central in the digital media environment. It is estimated that over three billion images were uploaded daily to popular social media platforms like Snapchat, Facebook, Instagram, and Whatsapp in 2015 (Meeker, 2016). Visuals can have
substantial and distinct effects on media consumers (Messaris, 1997). For instance, compared with plain text, visual imagery can further attract audience selection (Knobloch et al., 2003; Zillmann et al., 2001), facilitate information diffusion (Guerini, Staiano, & Albanese, 2013), evoke strong emotion (Iyer et al., 2014; Powell et al., 2015), provide indexical evidence (Bock, 2016), and influence attitudes and behaviors (Houts, Doak, Doak, & Loscalzo, 2006).

This dissertation aims to contribute to prior research on media bias by analyzing visual content and, in particular, how politicians are represented. As previous studies have demonstrated, visual factors—such as facial appearance (Todorov et al., 2005; Banducci et al., 2008), emotional display (McHugo et al., 1985; Tiedens, 2001), skin tone (Alter et al., 2016; Caruso et al., 2009), and head position (Mandell & Shaw, 1973)—can significantly shape our impressions and evaluations of politicians. And visual representations are ubiquitous: they are quintessential in various modes of political communication, including newspapers (Barrett & Barrington, 2005a), newsmagazines (Moriarty & Garramone, 1986), television (Grabe & Bucy, 2009), websites (Hehman et al., 2012; Verser & Wicks, 2006), social media (Goodnow, 2013) and field campaigns (Horiuchi et al., 2012). And yet if the automated detection of bias in textual content on a large scale already presents some challenges, quantifying bias in visual content requires an ever more sophisticated approach.

To advance our analysis of visual data, this dissertation draws inspiration from the emerging field of computer vision. Computer vision aims to imitate the human
vision’s ability to perceive and process images, and it is already being applied in a variety of contexts, such as medical imaging, facial recognition, and automated driving (Szeliski, 2016). Compared with traditional image analysis methods that often require a significant amount of human labor (e.g., Hu, Manikonda, & Kambhampati, 2014; Lazard, Dudo, Dennis, Ewald, & Love, 2017), computer vision techniques provide communication researchers with efficient, convenient, and standardized tools of analyzing visual data.

The many facets of bias

How does media bias manifest itself in the new media landscape? One of the most examined types of bias is partisan media bias—whether media outlets favor one political ideology, party or candidate (Groeling, 2013). As people tend to selectively consume news stories that resonate with their pre-existing opinions (Iyengar & Hahn, 2009), media outlets also adjust their content to respond to consumer preferences (Gentzkow & Shapiro, 2010). Scholars have argued that we are entering an age when media outlets explicitly label their political affiliations to market themselves to consumers in an increasingly competitive media environment (Groeling, 2013).

Media bias also manifests itself in representations of different social groups with regards to gender (Fredrickson & Roberts, 1997), age (Vasil & Wass, 1993), race (Alter et al., 2016), obesity (Heuer et al., 2011; Pearl et al., 2012; McClure et al., 2011), mental illness (Wahl, 1992), immigration status (Branton & Dunaway, 2009;
Matthes & Schmuck, 2017), etc. Biased depictions of different social groups have real-world consequences—it may exaggerate pre-existing stereotypes (Kay et al., 2015), elicit negative attitudes towards minority and disadvantaged groups (Caruso et al., 2009; McClure et al., 2011; Messing et al., 2016; Pearl et al., 2012), intensify group tensions and conflicts (Schmuck et al., 2017), and reduce physical and psychological well-being of groups that receive biased coverage (Fredrickson & Roberts, 1997).

One type of media bias that is particularly relevant to this dissertation’s context—visual representations of politicians—is the bias in covering politicians of different genders. Gender plays an important role in shaping voters’ evaluations of political candidates (Barrett & Barrington, 2005b; Dolan, 2014; Kahn, 1992; Kahn, 1994). Though women politicians overall have gained increasing prominence on the global stage, they are still substantially underrepresented in the political world and they often receive more biased coverage than their male counterparts (Aday & Devitt, 2001; Devitt, 2002; Center for American Women and Politics, 2017). For example, scholars have demonstrated women politicians often received more coverage on their personal information—for instance, personality, attire, appearance, and family—while men politicians received more coverage on their issue positions (Aday & Devitt, 2001; Devitt, 1999; Devitt, 2002; Dunaway et al., 2013; Kahn, 1994).

To summarize, I plan to examine two types of bias in visual representations of politicians. The first refers to partisan media bias—which has been extensively
examined in previous research: the extent to which media outlets favor one political party, candidate or ideology over the other (Groeling, 2013; Waldman & Devitt, 1998). The second bias this dissertation pays attention to is gender bias, that is, the extent to which women and men politicians are presented differently in a systematically biased and stereotyped way. By looking into both partisan media bias and gender bias, this dissertation hopes to advance our understanding of how bias is perpetuated visually.

**Visual bias**

Visual content plays an increasingly important role in today’s digital media environment. Camera phones and image-based social media, such as Instagram and Snapchat, have gained widespread popularity, particularly among the younger generation (Lenhart, 2015). In addition, empirical evidence has frequently demonstrated that visual features like facial appearances (Todorov et al., 2005) and emotional expressions (McHugo et al., 1985; Sullivan & Masters, 1988; Stewart et al., 2009) can shape how we perceive political candidates. For example, a simple judgment of competence from a candidate’s face alone can predict real-world election results (Todorov et al., 2005), and darkening the skin tone of a candidate’s face can make people less likely to vote for the candidate (Caruso et al., 2009).

Figure 1.1 provides a vivid example of how visual media may biasedly portray politicians. In the front page of *Daily Mail* published on March 28, 2017, Nicola
Sturgeon, Scotland’s First Minister, and Theresa May, U.K.’s Prime Minister, were sitting together with their legs shown prominently, accompanied by a headline “Never mind Brexit, who won Legs-it!” This emphasis on their legs is consistent with previous studies showing that media coverage of women politicians disproportionately focuses on their appearance compared with men politicians (Archer, Iritani, Kimes, and Barrios, 1983; Aday & Devitt, 2001; Kahn, 1994) and women’s bodies are more likely to be subject to sexual objectification (Fredrickson & Roberts, 1997) than men’s.

Given the difficulty in analyzing images on a large scale, however, studies that have systematically examined partisan media bias have mostly looked at textual data (e.g., Gentzkow & Shapiro, 2010; Groseclose & Milyo, 2005; for a review, see Groeling, 2013). Similarly, the majority of studies on media representations of women politicians have often focused on the textual part of media coverage (Aday & Devitt, 2001; Bystrom et al., 2001; Devitt, 1999; Devitt, 2002; Dunaway et al., 2013; Kahn, 1994). The bias in visual representations of politicians remains an under-examined but important field. Considering the substantial effects of politicians’ visual images on voters’ political attitudes and behaviors (Caruso et al., 2009; Todorov et al., 2005), and the increasing prevalence of visual content in today’s digital media environment (Lenhart, 2015; Meeker, 2016), it is important for us to understand how visual reporting of political candidates differ in media outlets across the political spectrum.
Applying computer vision in communication research

To advance our analysis of large-scale visual data, this dissertation connects communication research to an emerging field, computer vision. Some studies have already started to use computer vision techniques to investigate research questions that are relevant to social sciences, for example, predicting images’ aesthetical appeal (Dhar, Ordonez, & Berg, 2011; Ke, Tang, & Jing, 2006), online virality (Deza & Parikh, 2015; Guerini et al., 2013; Gelli, Uricchio, Bertini, Del Bimbo, & Chang, 2015; Totti, Costa, Avila, Valle, Meira Jr, & Almeida, 2014), users’ personality (Guntuku, Qiu, Roy, Lin, & Jakhetiya, 2015; Liu, Preotiuc-Pietro, Samani, Moghaddam, & Ungar, 2016), intelligence (Wei & Stillwell, 2016), and mental health status (Manikonda & De Choudhury, 2017; Reece & Danforth, 2017).

In this dissertation, I apply a variety of computer vision techniques, such as face
detection (Bakhshi et al., 2014), emotion detection (Liu et al., 2016), object recognition (Garimella et al., 2016), and computational aesthetics (Ke et al., 2006). Compared with traditional image analysis methods that often require human coding (Hu et al., 2014; Lazard et al., 2017), computer vision techniques may provide communication scholars with efficient and standardized tools to deal with the large amount of visual data in today’s digital media environment.

Nevertheless, questions still remain regarding how this approach can complement more traditional ways of image analysis and help us generate theoretically meaningful insights. Previous studies using computer vision methods—most in the field of computer science—often include a long list of features to predict certain outcomes and focus on the prediction accuracy of their models. For example, Totti et al. (2014) used 39 computer vision features to predict the popularity of images on Pinterest. Yet, the theoretical links between visual features and outcomes are rarely discussed (Totti et al., 2014; Wei & Stillwell, 2016). In addition, the accuracy of different computer vision applications has been rarely tested in the context of social science research. We need more research to determine how to best link computer vision features to concepts that are theoretically meaningful in visual communication as well as how to reliably apply computer vision methods to the analysis of media content. One of the methodological goals of this dissertation is to map out how we can leverage emerging computational methods to shed light on visual bias and improve our theoretical understanding of its expression.
Theoretical and practical contributions

To summarize, the main goals of this dissertation are (1) to identify the visual features that best discriminate positive and negative depictions of politicians, (2) to apply these results to investigate two types of media bias: partisan media bias, or how liberal and conservative media outlets portray the same politicians, and gender bias, or women and men politicians are presented differently in the pictorial domain, and (3) examine the methodological potential of computer vision techniques in communication research.

This dissertation aims to provide a pioneering large-scale investigation into the prevalence, extent, and patterns of bias in visual representations of politicians. First, traditional ways of image analysis often require a significant amount of human labor, and the majority of studies on visual bias often examine a limited number of media outlets—for instance, Moriarty and Popovich (1991) looked at three newsmagazines and Hehman et al. (2012) included five online media outlets in the analysis. Additionally, previous research on visual bias has often found limited evidence for substantial media bias in visual forms (Moriarty & Garramone, 1986; Moriarty & Popovich, 1991; Banning & Coleman, 2009; Grabe & Bucy, 2009). Yet, it is questionable whether this observation still holds true in a media environment where partisan media outlets are increasingly prevalent. By employing computer vision tools, this dissertation offers a more comprehensive look at the current media landscape in
the visual domain.

Theoretically, this dissertation bridges several broad fields of literature: research on media bias (Banning & Coleman, 2009; Moriarty & Garramone, 1986; Moriarty & Popovich, 1991; Waldman & Devitt, 1988), theories of social cognition and perception, which deal with how people form impressions and judgments related to others (Fiske, 1993), and politicians’ use of social media (Bene, 2017; Meeks, 2016; O’Connell, 2018). Previous research has demonstrated that various visual features—such as face size (Archer et al., 1983), skin tone (Caruso et al., 2009; Kemmelmeier & Chavez, 2014; West et al., 2014; Alter et al., 2016), and facial expressions (Stewart et al., 2009; Tiedens, 2001)—can indeed shape voters’ perceptions of other people as well as politicians. Yet, we know little about what kinds of visual content news consumers are actually exposed to and how frequently they see those features. This dissertation aims to fill in that empirical gap using theories of social cognition to guide the selection of visual features and the measurement of media bias.

In addition, this dissertation also aims to advance our understanding of how social norms around gender are constructed and mediated. Even though an increasing number of women are running for office on the global stage, women politicians are still underrepresented in the political world (Powell & Graves, 2003). Although extensive research has looked at how women and men are portrayed differently in the media (e.g., Martins et al., 2009), some evidence has also demonstrated that women politicians do not share the same stereotypes with women in general (Schneider &
Bos, 2014). This dissertation, therefore, provides a valuable look into the gender norms regarding women politicians specifically. Additionally, previous research on gender stereotypes have mostly used self-reported measures (Haines et al., 2016; Martin et al., 1990; Prentice & Carranza, 2002) and research on media bias in covering women politicians has predominately focused on textual content (Aday & Devitt, 2001; Bystrom et al., 2001). This dissertation probes into the visual representations of gender to surface some important aspects of gender norms that might not be captured by self-reported measures or textual forms, for example, nonverbal behavior.

Ultimately, this dissertation also aims to have practical implications. Bias exists not in only textual but also visual forms. It is prevalent in traditional media sources (Kahn, 1994; Devitt, 2002) as well as in online information providers like Google Image Search (Kay et al., 2015) and Wikipedia (Graells-Garrido et al., 2015; Wagner et al., 2015). Previous scholars have made various attempts to expose media consumers to more ideologically diverse and balanced news content (Park et al., 2009; Munson & Resnick, 2010). Future practitioners, ranging from developers to journalists, can build tools to detect bias in visual content. For example, one implication of this research is that it can help social media designers, reporters, and editors to automatically label the potential biases in images in newsfeeds and present readers with more diverse visual content. Likewise, search engine engineers can factor known biases into the design of their algorithms to present results that do not
intensify pre-existing stereotypes about certain social groups.
Chapter 2 Measures of Media Bias

Partisan media bias

With the advent of digital media and cable news, we have witnessed an expansion of partisan media outlets that explicitly favor one political ideology over the other to appeal to like-minded consumers (Bennett & Iyengar, 2008; Groeling, 2013). Today’s news consumers increasingly resort to information outlets that echo their pre-existing ideological affiliations, which is a well-documented phenomenon called “selective exposure” (Garrett, 2009; Hart et al., 2009). For example, in an experimental study, liberals and Democrats were more likely to select news stories attributed to CNN and NPR, avoiding Fox News, whereas conservatives and Republicans tended to read stories attributed to Fox News (Iyengar & Hahn, 2009). A survey from the Pew Research Center showed that consistent liberals resorted to a variety of left-leaning media outlets like CNN, MSNBC, NPR, and the New York Times for news, whereas consistent conservatives predominantly turned to Fox News for information (Mitchell, Gottfried, Kiley, & Matsa, 2014).

News outlets are also adjusting their content to attract readers who prefer partisan information (Bennett & Iyengar, 2008; Prior, 2013). For example, Baum and Groeling (2008) showed online partisan websites (DailyKos.com, FreeRepublic.com, FoxNews.com) were more likely to select stories in accord with their partisan alignment, compared with newswire sites like Reuters and the Associated Press. Scholars have also started to question if the combination of audience selectivity and
partisan media outlets intensifies attitudinal polarization among citizens and creates online “echo chambers” where news consumers are constantly exposed to information that affirms rather than challenges their views (Fletcher & Nielsen, 2017; Garrett, 2009; Prior, 2013).

Despite the extensive research on media bias, scholars divide on its definition (see Table 2.1). In this dissertation, I define partisan media bias as systematical content patterns in media outlets that favor one political party, candidate, or ideology over another. Several aspects of this definition need to be clarified first:

(1) First, media bias should be distinguished from news slant in a particular piece of media content (Entman, 2007)—it should be “systematic, rather than anecdotal, episodic, or fleeting” (Groeling, 2013) (p.133).

(2) In addition, my definition abandons the notion that bias should be established as a deviance from the truth, a distortion of reality, or a violation of journalistic balance. First, many partisan media outlets today are intentionally producing ideologically slanted content that favors one party or ideology over another and do not adhere to traditional journalistic standards of objectivity or balance (Groeling, 2013). Furthermore, the goals of balance and accuracy are not always consistent (Hackett, 1984). For instance, if a political candidate has more scandals than his or her competitor, an accurate approach is to report more negative stories on this candidate, yet a balanced approach is to produce an equally number of negative stories for both sides. Finally, the benchmarks for establishing where the truth lies are
often unobservable to researchers, making it difficult for researchers to empirically operationalize media bias as the difference between media coverage and the reality (Hackett, 1984; Groeling, 2013).

(3) The operationalization of bias should incorporate the potential effects on the audiences. For example, if a piece of media content is determined by researchers as favoring one political party, it should produce more favorable impressions of that party among the audience—or at least, viewers exposed to it can perceive the intended slant.

(4) Unlike previous inquiries (Groseclose & Milyo, 2005; D’Alessio & Allen, 2000), my dissertation does not examine whether media have an overall liberal or conservative bias. Given the rise of partisan media outlets and the fragmentation of the media environment (Groeling, 2013), instead of treating media as a monolithic subject of inquiry, examining the patterns of ideological slant within each individual media outlet is more relevant to the current media landscape.

(5) Previous research has identified three major kinds of partisan media bias: coverage bias (also called as visibility bias, “which considers the relative amounts of coverage each party receives”); gatekeeping bias (also called as selection bias or agenda bias, “which is the preference for selecting stories from one party or the other”); and statement bias (also called as tonality bias or presentation bias, “which focuses on the favorability of coverage toward one party or the other”) (D’Alessio & Allen, 2000, p.133; Groeling, 2013; Eberl, Boomgaard, & Wagner, 2015). This
dissertation will focus on presentation or tonality bias—specifically, whether visual media content portrays political candidates favorably or unfavorably.

Although media bias frequently appears in the public and academic discourse, it is not an easy task to systematically quantify whether and to what extent the political bias exists in different media outlets (for a review, see Groeling, 2013). It is well-known in communication research that individuals’ interpretations of media content are influenced by pre-existing attitudes and cognitive biases like hostile media effects—among readers with two opposing opinions, even an identical article could both be seen as biased against their sides (Vallone, Ross, & Lepper, 1985). Therefore, it is important to come up with objective measures of media bias instead of relying on news consumers’ subjective interpretations. Here, I summarize previous endeavors to measure media bias into two broad approaches: audience-based and content-based approach (Budak, Goel, & Rao, 2016).

<table>
<thead>
<tr>
<th>Table 2.1 Definitions of partisan media bias in prior research</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hackett (1984)</strong></td>
</tr>
<tr>
<td>“the concept of news bias has two moments which are not entirely consistent. One is a lack of ‘balance’ between competing viewpoints; the other is a tendentious, partisan ‘distortion’ of ‘reality.’” (p. 230)</td>
</tr>
<tr>
<td><strong>Waldman &amp; Devitt (1998)</strong></td>
</tr>
<tr>
<td>“Bias can be defined as any systematic slant favoring one candidate or ideology over another.” (p. 302)</td>
</tr>
<tr>
<td><strong>Entman (2007)</strong></td>
</tr>
<tr>
<td>Content bias is defined as “consistent patterns in the framing of mediated communication that promote the influence of one side in conflicts over the use of government power.” (p. 166)</td>
</tr>
<tr>
<td><strong>Zeldes et al. (2008)</strong></td>
</tr>
<tr>
<td>Bias is defined “operationally as the extent to which partisan opponents are given equal prominence and scope to make their cases to the public.” (p. 565)</td>
</tr>
<tr>
<td>Source</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Groeling (2013)</td>
</tr>
<tr>
<td>Galvis et al. (2016)</td>
</tr>
<tr>
<td>Eberl et al. (2016)</td>
</tr>
</tbody>
</table>

**Audience-based approach**

Based on the assumption that audience members self-select themselves to news content that matches their political affiliations (Bennett & Iyengar, 2008; Hart, Albarracin, Eagly, Brechan, Lindberg, & Merrill, 2009), one way to quantify media bias is to examine the ideological breakdown of each media source’s audience (Flaxman et al., 2016). For example, the Pew Research Center measured the ideological alignment of media sources by aggregating the political affiliations of survey respondents who reported using each media outlet (Mitchell, Gottfried, Kiley, & Matsa, 2014). Zhou, Resnick, and Mei (2011) calculated the political slant of news articles on a news aggregator site Digg based on the assumption that users would vote for articles that shared the same political affiliation with them. Bakshy, Messing, and Adamic (2015) also adopted a similar approach and measured the ideological slant of an article by averaging the political orientations of users who had shared that article.
on Facebook.

Some studies have also integrated network relationships among audience members to quantify the extent of media bias. For instance, An, Cha, Gummadi, Crowcroft, and Quercia (2012) calculated the similarity between two media sources’ ideological affiliations by considering the proportion of Twitter users co-subscribing to both accounts. The authors then arranged various media outlets on a single dimension of ideological slant based on the similarity measures among news sources.

**Content-based approach**

Another way to measure media bias is to systematically compare media content from different sources. Groseclose and Milyo (2005) measured media bias by comparing the number of times that a media outlet cited certain think tanks and policy groups to the number of times that congresspeople of the two parties cited the same groups. Gentzkow and Shapiro (2010) used the similarity between the phrases used by news outlets and those used by congresspeople as an index of media slant. Ho and Quinn (2008) resorted to editorials published by newspapers and identified the explicit political positions held by these newspapers. Budak, Goel, and Rao (2016) measured media bias by aggregating crowdsourced workers’ perceived slant in articles produced by various media outlets.

These studies sometimes result in very similar results regarding the ideological rankings of some popular media outlets (e.g., Bakshy et al., 2015), implying that
media bias seems to be a well-established phenomenon. Still, studies do differ in terms of the magnitude of media bias and ideological positions of a few news outlets (Groseclose & Milyo, 2005; Gentzkow & Shapiro, 2010; Budak et al., 2016). As Groeling (2013) wrote: “It now seems clear that we are entering a new, more partisan era in American journalism. Ironically, it appears that the future will more closely resemble the nineteenth century’s press than the twentieth century’s, with some news organizations choosing to distinguish themselves in a crowded marketplace by delivering a reliably ideological product” (p. 148). A most interesting question, therefore, is not whether media bias exists, but in what ways media coverage can be biased, which is one focus of this dissertation.

**Gender bias**

Although women are increasingly represented in the political world (Powell & Graves, 2003), many studies have pointed out that women politicians are still receiving biased media coverage (Aday & Devitt, 2001; Devitt, 2002; Kahn, 1994). One way to quantify gender bias is to compare the amount and the tonality of news coverage given to women and men candidates. For example, a content analysis of news coverage of statewide campaigns between 1982 and 1988 found that in senatorial races, women candidates received less coverage than their men competitors, and the coverage was also more negative for them (Kahn, 1994). However, based on more recent studies, media coverage of women candidates has improved over the
decades, at least in terms of the amount and the tonality (Aday & Devitt, 2001; Banwart et al., 2003; Smith, 1997). For example, a study regarding four women political candidates (Elizabeth Dole, Claire McCaskill, Hillary Clinton, and Sarah Palin) found that women candidates actually received more coverage than their male opponents (Meeks, 2012).

Nonetheless, quantifying media bias on a single dimension ranging from negative to positive might be an over-simplification of a complex phenomenon. In addition to the amount and the positivity of media coverage, research has also examined various ways of how women and men candidates are covered in qualitatively different ways. Previous research about gender bias can be summarized into two theoretical traditions: objectification and gender stereotype (e.g., Bauer, 2013).

**Objectification theory**

Objectification theory posits that women’s bodies are constantly looked at, evaluated, and objectified in our society. Objectification occurs when women’s bodies or body parts are detached from them and become physical objects to be appraised and sexually desired by men (Fredrickson & Roberts, 1997).

Scholars have argued that various forms of media frequently adopt “the male gaze,” which portrays the world from a heterosexual male’s perspective and presents women as objects of men’s pleasure (Mulvey, 1975). For example, in visual media that depict interpersonal interactions between the two genders, women appear more
frequently as subjects to be looked at by men than the reverse (Goffman, 1979; Fredrickson & Roberts, 1997). What’s more, female bodies are more likely to be objectified, sexualized, and portrayed prominently than male bodies in the media (Mulvey, 1975). Empirical studies have also shown that visual media tend to focus on men’s faces but women’s bodies, a phenomenon termed as “face-ism” in prior scholarship (Archer et al., 1983).

Empirical studies have shown that focusing on a woman’s appearance leads to objectification and negative evaluation (Heflick & Goldenberg, 2009; Heflick, Goldenberg, Cooper, & Puvia, 2011). In an experiment about the 2008 vice presidential candidate, Sarah Palin, participants were less likely to rate her as competent, see her as fully human, and vote for her partnering presidential candidate, John McCain, when being instructed to focus on her appearance instead of her person (Heflick & Goldenberg, 2009). In a series of experiments, focusing on a woman’s appearance led participants to perceive her as less warm, competent, and moral, but this appearance focus did not effectively influence participants’ evaluation of a man (Heflick et al., 2011).

**Gender stereotypes**

Stereotypes are “shared beliefs about the attributes and behaviors of individuals based on their membership in groups defined by a singular characteristic such as race, gender, or age” (Hamilton & Sherman, 1994, cited in Bauer, 2013, p.24). Previous
research has identified four specific components of gender stereotypes, namely, traits, role behaviors, physical characteristics, and occupation (Deaux & Lewis, 1984; Jackson & Cash 1985; Haines et al., 2016). Some personality traits are stereotypically viewed as feminine, such as being compassionate, trustworthy, emotional, and family-oriented, and some traits are traditionally associated with masculinity, such as being ambitious, aggressive, intelligent, rational, and tough (Alexander & Andersen, 1993; Deaux & Lewis, 1984; Huddy & Terkildsen, 1993; Schneider & Bos, 2014).

Additionally, women and men are expected to perform different social roles. Women are traditionally expected to be caregivers, for example, taking care of children, tending the house, providing emotional support, doing the laundry and grocery shopping. In contrast, men were expected to be leaders and financial providers, and to be responsible for chores related to mechanics (Deaux & Lewis, 1984; Haines et al., 2016).

Previous research, however, has also shown that stereotypes about women politicians do not entirely overlap with typical stereotypes about women. Instead, women politicians might constitute a “subtype” of women with their own stereotypical characteristics (Schneider & Bos, 2014). The process of subtyping occurs when some members of a certain group disconfirm the group stereotypes, thus being placed in a subcategory that differs from other group members (Richards & Hewstone, 2001). Specifically, being a politician can make a woman more “masculine” (Meeks, 2012; Schneider & Bos, 2014). In one survey, women politicians were seen
as less feminine, emotional, caring, beautiful but more well-educated, confident, and assertive than typical women (Schneider & Bos, 2014). But they were still viewed as less likely to possess masculine traits like being competitive and ambitious than their male counterparts (Schneider & Bos, 2014). Similarly, in another study, survey respondents tended to associate women politicians with traditionally feminine traits compared with men politicians (Dolan, 2010). Therefore, women politicians seem to have disadvantages from both sides—they are not “man” enough, yet they also lose some positive traits typically associated with women like being compassionate and warm (Meeks, 2012).

Media coverage of women politicians

Frames. One way of measuring gender bias deals with how candidates are framed in the media (Aday & Devitt, 2001; Meeks, 2012; Kahn, 1994). Studies have frequently shown that women politicians often receive more coverage on their personal information (personal frame)—for instance, personality, attire, appearance, and family—while men politicians receive more coverage on their issue positions (issue-frame) (Kahn, 1994; Aday & Devitt, 2001). The overemphasis on women candidates’ appearance seems to resonate with the objectification scholarship showing that women tend to be sexually objectified in the media, portrayed as objects to be evaluated by a hypothetical male viewer (Fredrickson & Roberts, 1997; Heflick & Goldenberg, 2009; Mulvey, 1975), whereas the emphasis on their family seems to
echo with the stereotype that women are presumed to be caregivers or dependents of men (Deaux & Lewis, 1984; Haines et al., 2016).

Previous research has detected this type of framing bias for women candidates running for all levels of offices. Aday and Devitt (2001) compared five newspapers’ coverage of Elizabeth Dole and her male opponents during her run for the Republican presidential nomination in 1999. Compared with other male contenders, Dole received more coverage on her personality and background/qualification but less on her issue stances. Heldman, Carroll, and Olson (2005) also revealed a similar pattern in the media coverage of Elizabeth Dole, that media outlets disproportionately mentioned her dress/appearance, husband, and personality, and frequently emphasized her status as the first woman to be a serious presidential candidate.

Similar patterns have also surfaced for candidates running for U.S. governors and senators (Kahn, 1994; Devitt, 2002; Bystrom, Robertson, & Banwart, 2001; Dunaway, Lawrence, Rose, & Weber, 2013).

**Traits.** Given that studies have shown that women candidates consistently received more coverage on their traits than men candidates, the next question would be what kinds of traits are mentioned in the media. One possibility is that media coverage of politicians is in line with stereotypical traits assigned to woman. In an analysis of statewide campaigns between 1982 and 1988, women and men gubernatorial candidates did not received biased media coverage with regards to traditionally “masculine” traits (e.g., strong leader, intelligent) and “feminine” traits...
(e.g., compassionate, honest), though there was more discussion of “masculine traits” for men senatorial candidates.

Interestingly, there is growing evidence suggesting women politicians do not share the stereotypes that are commonly ascribed to women but instead, women professionals are a “subtype” of women (Schneider & Bos, 2014). Media coverage of women politicians also emphasize their masculine, gender-incongruent traits. For example, when covering the 1993 and 1997 Canadian election campaigns, reporters actually applied more neutral verbs (e.g., say, tell) to refer to men candidates’ speech while using more affectively negative and aggressive verbs (e.g., attack, blast, fire) to describe women candidate’s speech, highlighting their aggressiveness and combativeness (Gidengil & Everitt, 2003). It is unclear, however, whether this pattern is due to women leaders’ own strategies to appear tough or media outlets’ practices that emphasize women leaders’ gender-incongruent traits.

**Issues.** Different political issues are stereotypically associated with the two genders as well as masculine or feminine traits (Banwart et al., 2003; Bystrom et al., 2001; Kahn, 1994). Issues like healthcare and education that emphasize nurturing and caring are traditionally seen as feminine issues, whereas issues like economy, military, defense, crime, and foreign policy are traditionally regarded as masculine issues. For example, in one study, participants viewed a candidate with a female name as more capable of handling traditionally feminine issues like “improving education” or “assisting the poor” than one with a male name (Leeper, 1991). Likewise, participants
perceived a candidate portrayed with typically feminine traits (e.g., trustworthy, family-oriented) as more capable of handling welfare issues and a candidate portrayed with traditionally masculine traits (e.g., tough, ambitious) as better at tackling military and economic issues (Huddy & Terkildsen, 1993).

The evidence for gender bias in issue coverage is mixed. Kahn (1994) found that in senatorial races, women candidates received more media coverage on their stances on stereotyped feminine issues but less on stereotyped masculine issues than men candidates, although this pattern did not emerge among gubernatorial races. Yet, in another study on senatorial and gubernatorial races between 1992 and 2000, there were virtually no differences regarding media coverage of traditionally masculine and feminine issues assigned to candidates of the two genders (Jalalzai, 2006). In summary, although there are traits and issues stereotypically assigned to women and men politicians in people’s mind, it is unclear that media coverage actually mirrors these patterns based on analyses of textual data.

**Visual bias**

Previous studies examining media bias have predominantly focused on textual materials (Groseclose and Milyo, 2005; Gentzkow & Shapiro, 2010; Devitt, 2002; Kahn, 1994). Yet, one line of research has also investigated media bias in visual representations of politicians (Banning & Coleman, 2009; Barrett & Barrington, 2005a; Hehman, Graber, Hoffman, & Gaertner, 2012; Moriarty & Garramone, 1986; Moriarty & Popovich, 1991; Waldman & Devitt, 1988). Scholars have come up with
various ways of determining whether a piece of visual content portrays a politician favorably or unfavorably. Here I categorize them into two approaches.

**Subjective rating approach**

Some studies have used more subjective measures of media bias by directly instructing some coders to rate visual content’s tonality. For example, in Barrett and Barrington’s (2005a) study, three coders rated the favorability of a photo presenting a candidate on a five-point scale ranging from “highly unfavorable” to “highly favorable.” Their results indicated that newspapers indeed used photographs that were more favorable to the candidates they endorsed than the candidates’ opponents. In another study that examined photographs of candidates from five online media outlets, six coders rated the candidate in each photograph on two dimensions—warmth and competence. This study also revealed that media outlets portrayed candidates they supported as warmer and more competent (Hehman et al., 2012).

Yet, this approach suffers from a methodological weakness that affects many media bias studies—the issue of subjectivity (Groeling, 2013; Zeldes, Fico, Carpenter, & Diddi, 2008). First, individuals’ interpretations of media content are often subjective to prior attitudes and cognitive biases such as hostile media effects (Hansen & Kim, 2011; Vallone, Ross, & Lepper, 1985). In a classical experiment, readers with two opposing opinions both interpreted the same article as biased against their sides (Vallone et al., 1985). Additionally, even coders in one study can agree
upon the tonality in some media content and achieve high inter-coder reliability, it is questionable how well these coders’ evaluation of content bias in one study is transferrable to other researchers’ judgment of the same content or an average news consumer’s interpretation (Zeldes et al., 2008).

**Objective feature checklist approach**

Other scholars have come up with more objective measures of media bias by using a checklist of features that may reflect the favorable/unfavorable treatment of political candidates in visuals. The criteria they use in determining visual slant usually include facial expression, physical position, interaction with the audience, camera angle, setting, and so on (Kepplinger, 1982; Moriarty & Garramone, 1986; Waldman & Devitt, 1988; Moriarty & Popovich, 1991; Barrett & Barrington, 2005a; Banning & Coleman, 2009).

Table 2.2 summarizes some common coding criteria of visual bias used in previous research (Moriarty & Garramone, 1986; Moriarty & Popovich, 1991; Waldman & Devitt, 1988; Verser & Wicks, 2006), illustrated with photographs of the two presidential candidates in the 2016 U.S. presidential election. These criteria can be grouped into three categories: (1) a politician’s own nonverbal behaviors, such as facial expression, hand gesture, and eye status; (2) contextual features, such as other subjects and other people in the picture, and (3) structural features, such as color and composition.
These studies have revealed that the ideological slant in news outlets can indeed manifest itself in visual forms, although its extent is often not substantial (Moriarty & Popovich, 1991; Banning & Coleman, 2009). For example, in an analysis of three newsmagazines’ (*U.S. News and World Report*, *Time*, *Newsweek*) photographic coverage of the 1984 presidential campaign, Moriarty and Garramone (1986) used a list of ten features to determine whether a photo portrayed a candidate positively or negatively. However, they did not detect substantial differences among the media representations of these candidates. In a later study that examined the same three newsmagazines’ coverage of the 1988 presidential campaign, Moriarty and Popovich (1991) extended that list to a total of 15 image attributes—to name a few, expression, activity, family, and camera angle—in addition to the sheer number of images. They found that the Republican candidate, George Bush, was treated only slightly favorably.

Other studies have employed similar attributes to determine the favorability of a photograph featuring a candidate (Banning & Coleman, 2009; Waldman & Devitt, 1988). For example, Waldman and Devitt (1988) used five features—expression, activity, interaction, background, and camera angle—to examine photographs from five newspapers featuring the two candidates in the 1996 presidential campaign. They found that Bill Clinton received slightly better pictorial treatment. In an analysis of photographs used by the 2000 presidential candidates on their own campaign websites, Verser and Wicks (2006) also proposed several criteria in addition to the
criteria used by Moriarty and Popovich (1991), for example, eye contact and camera focus. Their analysis revealed that the two candidates—Al Gore and George W. Bush—presented themselves differently in photographs. For example, Gore was often shown interacting with others while Bush was portrayed as less active and often shown alone.

In addition to still photographs, some studies have also looked at the potential bias in television coverage of political candidates (Banning & Coleman, 2009; Grabe & Bucy 2009). For instance, Banning and Coleman (2009) examined television shots of the 2000 presidential candidates and showed that the coverage was fairly balanced between the two candidates, with George W. Bush receiving slightly more favorable visuals. Grabe and Bucy (2009) analyzed television coverage of major presidential candidates between 1992 and 2004, incorporating four visual aspects of bias (visual editing techniques, camera distance, camera angle, and camera movement). Their study concluded that the Republican Party received more favorable television coverage (Grabe & Bucy, 2009).

This approach, however, still has several methodological weaknesses that need to be addressed. Previous studies using a list of visual attributes to code media content often treat these attributes as equally important, and sometimes, construct an overall index of visual bias with all the attributes equally weighted (Banning & Coleman, 2009; Moriarty & Popovich, 1991; Waldman & Devitt, 1988). As some scholars have already mentioned (Barrett & Barrington, 2005a; Groeling, 2013), this approach is
problematic, as these features do not contribute equally to the favorability of images:

for example, interviews showed that audience members predominantly used facial expressions and gestures to determine the favorability of news photographs of politicians (Lobinger & Brantner, 2015). Additionally, while the effects of some visual features on voters’ perceptions of candidates have been extensively documented (e.g., camera angle, see Mandell & Shaw, 1973; McCain, Chilberg, & Wakshlag, 1997; Mignault & Chaudhuri, 2003), the effects of other features are mostly based on scholars’ own interpretations and have rarely been empirically tested.

To summarize, the many features proposed by previous studies should not be regarded as equally sensitive proxies of media bias, nor should their effects be taken for granted.
<table>
<thead>
<tr>
<th>Activity</th>
<th>Favorable</th>
<th>Unfavorable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dynamic behavior</td>
<td>such as speaking, shaking hands, kissing babies</td>
<td>Lethargic or passive activity, such as listening, reading, dozing</td>
</tr>
<tr>
<td>Posture</td>
<td>Standing tall and upright</td>
<td>Bowed, slumped, or leaning on something</td>
</tr>
<tr>
<td>Arms</td>
<td>Arms head high or above</td>
<td>Arms at side, at rest, or folded</td>
</tr>
<tr>
<td>Hands</td>
<td>Gesturing or doing something</td>
<td>Hands at side, or at rest</td>
</tr>
<tr>
<td>Eyes</td>
<td>Eyes looking directly at camera or at someone</td>
<td>Eyes up, down, or closed</td>
</tr>
<tr>
<td>Expression</td>
<td>Cheerful or confident</td>
<td>Unhappy, worried, or tired</td>
</tr>
<tr>
<td>Crowd interaction</td>
<td>Cheering crowd or attentive colleagues</td>
<td>Candidate alone or with attentive crowd or colleagues</td>
</tr>
<tr>
<td>Props</td>
<td>Campaign symbols like flags, bunting and the presidential seal or knowledge symbols</td>
<td>No symbols</td>
</tr>
<tr>
<td>Dress</td>
<td>Dignified suit and tie</td>
<td>Sports clothes or shirtsleeves</td>
</tr>
<tr>
<td>Accompany</td>
<td>With family, spouse, or running mate.</td>
<td>Alone</td>
</tr>
<tr>
<td>Camera angle</td>
<td>Looking up at candidate</td>
<td>Looking down on candidate</td>
</tr>
<tr>
<td>Camera distance</td>
<td>Close-up or the head and shoulders</td>
<td>A long shot with the full figure</td>
</tr>
</tbody>
</table>
Chapter 3 Proposed approach

To summarize, current studies on measuring visual bias still have a few limitations. Theoretically, many studies miss a link to the fields of social cognition, visual persuasion, and political psychology. For example, some studies on visual bias often operationalize portrayals of politicians as a single dimension ranging from positive to negative—yet, as some scholars have noted, voters perceive politicians on different dimensions (Grabe & Bucy, 2009; Bligh et al., 2012). Methodologically, many studies on visual bias either rely on a limited number of coders’ subjective perceptions of favorability or ask coders to rate images on a list of objective features and arbitrarily treat all the features as equally important.

To address these issues, this dissertation plans to combine both subjective ratings of favorability and objective measures of visual features.

(1) First, I plan to review previous research in media bias, social cognition, visual persuasion, and computer vision to identify a pool of visual features that potentially relate to biased portrayals of politicians and could be coded by computer vision techniques—this feature selection process will be detailed in Chapter 4 and Chapter 5.

(2) Next, I will use crowdsourced workers to generate subjective ratings of images’ favorability of portraying politicians and apply computer vision techniques to code objective features of images. Thus, using subjective ratings as criteria, we can identify a subset of visual features that actually matter in shaping viewers’
perceptions and sizes of their impacts.

(3) I then examine how computationally coded visual features—particularly those bringing effects on viewers’ perceptions of favorability—are distributed in media outlets with different political affiliations and social media representations of women and men politicians.

(4) At last, by combing results from both content analysis and viewers’ reactions, this dissertation reveals not only systematic biased patterns in media content but also their potential effects on viewers. By examining how various features are distributed in visual representations of politicians, we can generate new insights about visual framing, media practices, and gender norms.

An overview of computer vision methods

In this chapter, I will first review a few areas in computer vision that could potentially be applied in social science research.

Supervised learning and convolutional neural network

In general, machine learning algorithms can be classified into supervised and unsupervised learning (Müller & Guido, 2016). Regarding supervised learning,

---

1 There are some emerging branches of machine learning that go beyond these two mainstream areas, such as self-supervised learning and reinforcement learning (Chollet, 2019), which will not be covered in this chapter.
researchers have prior knowledge of output values for certain inputs and develop an algorithm that produces the desired output when given an input (Müller & Guido, 2016). Typical tasks of supervised learning include optical character recognition, image classification, and speech recognition (Chollet, 2019). Regarding unsupervised learning, researchers have no prior knowledge of the input and output. Instead, researchers have to identify the structure or patterns inside the data. Dimensionality reduction and clustering are two common categories of unsupervised learning (Chollet, 2019).

Figure 3.1 An artificial neural network with an input layer, two hidden layers, and an output layer

In the last decade, artificial neural networks (ANN) have become the most popular method for supervised learning tasks in computer vision such as image classification (Chollet, 2019). An artificial neural network is based on a collection of
connected neurons or nodes, which are usually organized in layers (see Figure 3.1 for an example). A deep neural network (DNN) is a type of ANN with multiple layers between the input and output layers. Typically, DNNs are feedforward networks in which data flow from the input layer to the output layer: each neuron receives signals from neurons in the previous layer, performs some computations, and passes its signal to neurons connected to it in the following layer (Chollet, 2019; Rosebrock, 2017).

Figure 3.2 Illustration of the VGG16 model

![VGG-16 Diagram](image)

*Note.* Image from Neurohive (2018).

A convolutional neural network (CNN) is a class of DNN that is frequently applied to the analysis of visual data. A CNN contains several specific types of layers, such as convolutional layers and max-pooling layers, which allow it to handle multi-dimensional data of images (for more details, see Chollet, 2019; Joo & Steinert-Threlkeld, 2018; Rosebrock, 2017). Figure 3.2 presents one popular model of CNN, VGG16, proposed by Simonyan and Zisserman (2014), that performs well in image classification tasks. CNNs can learn local patterns from raw pixels of image data. These patterns are (1) translation invariant, meaning that after a CNN learns a certain pattern in one part of an image, it can identify it anywhere in the image, and (2)
spatially hierarchical, meaning that after one CNN layer learns some simple patterns such as edges and shapes, the following layer can pick up more complicated patterns that are made of these simple patterns (Chollet, 2019, p.123). Figure 3.3 visualizes how a CNN learns to recognize objects (Zeiler & Fergus 2014). First, the neural network learns to identify simple, basic patterns such as edges and color patches. As the network goes deeper, it learns to recognize more complicated, advanced features, which are then used to classify images.

Figure 3.3 Feature visualization of a CNN trained on ImageNet

![Feature visualization of a CNN trained on ImageNet](image)

*Note.* Image from LeCun (2015).

Researchers who wish to computationally analyze a diversity of visual content often need to train neural networks on thousands, if not millions, of labeled images to achieve a satisfactory level of accuracy. However, researchers often do not have access to a large dataset to begin with. Fortunately, there are several transfer learning
methods that can deal with relatively small datasets, for example, feature extraction
with a pretrained network and fine-tuning (Chollet, 2019; Jean et al., 2016). A
pretrained network is “a saved network that was previously trained on a large dataset,
typically on a large-scale image-classification task” (Chollet, 2019, p. 143). Today,
many pretrained models are publicly available, which are usually trained on the
ImageNet dataset, a dataset containing 1.4 million labeled images of 1,000 categories.
As noted earlier, CNNs are able to recognize advanced and complicated features.
These features are originally used to classify images in one context, for example,
animals or everyday objects, but they can then be repurposed for a different task in
image classification. By using advanced features from pretrained models and training
the model with a new classifier, instead of training a brand-new model from raw
pixels of images, researchers are able to achieve good accuracy with relatively small
datasets (Chollet, 2019).

A few scholars have started to harness the power of CNNs for image analysis in
social science research (for a review, see Joo et al., 2018). Won et al. (2017) collected
a dataset of 40,764 images from Twitter and other online resources and used CNNs to
successfully identify protest-related images as well as the levels of violence
associated with protest images. Gebru et al. (2017) applied CNNs to recognize
vehicle models in Google Streets View images and estimated income, race, education,
and voting patterns in neighborhoods based on the information about vehicles. Using
satellite imagery, Jean et al. (2016) combined CNNs and transfer learning techniques
to predict consumption expenditure and asset wealth in African countries. These studies emerging from a wide range of disciplines together demonstrate computer vision enables social scientists to extract useful insights from the large amount of digital visual data circulated in today’s media environment.

**Unsupervised learning and image clustering**

As noted earlier, unsupervised learning uncovers hidden patterns in data without pre-existing labels. For example, when dealing with textual data, researchers can apply topic modeling to automatically detect topics in a collection of documents (e.g., news articles, social media posts, scientific abstracts). Social scientists have applied topic modeling for a wide range of tasks, for example, classifying political candidates’ tweets into different topics (Guo et al., 2016) and uncovering topics from scientific abstracts (Griffiths & Steyvers, 2004).

How can we automatically detect visual topics from a large set of visual images? In this dissertation, I propose two approaches (Figure 3.4). The first approach is to get descriptive tags from pre-trained computer vision libraries or online services, which I will detail below. For example, Clarifai labels an image with 20 potential descriptive tags. After an image is transformed into a collection of textual tags, one can applied well-established natural language processing techniques such as topic modeling to uncover topics in a collection of images.
Another approach is to apply clustering algorithms to extracted features from pre-trained models. Feature extraction from pre-trained models has commonly used for supervised learning purposes, but computer scientists have recently noted that this method can also be applied in unsupervised learning contexts, for example, grouping objects into similar-looking categories (Guérin et al., 2017). When fed into a pre-trained neural network, each image is transformed into a vector of thousands of extracted features. Clustering algorithms such as k-means can then be used to detect clusters within a collection of images. A few scholars have started to adopt this approach in social science research. For example, applying clustering algorithms on extracted features from a Resnet50 neural network trained with ImageNet data, Zhang and Peng (2019) were able to come up with ten categories from protest images on Weibo. These categories included crowds, banners, vehicles, screenshots, etc. (Table 3.1).
Table 3.1 Exemplar images in categories from Zhang and Peng (2019)

<table>
<thead>
<tr>
<th>Banners</th>
<th>Police</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1" alt="Exemplar Image 1" /></td>
<td><img src="image2" alt="Exemplar Image 2" /></td>
</tr>
<tr>
<td><img src="image3" alt="Exemplar Image 3" /></td>
<td><img src="image4" alt="Exemplar Image 4" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Crowd</th>
<th>Injuries and violence</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image5" alt="Exemplar Image 5" /></td>
<td><img src="image6" alt="Exemplar Image 6" /></td>
</tr>
<tr>
<td><img src="image7" alt="Exemplar Image 7" /></td>
<td><img src="image8" alt="Exemplar Image 8" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Vehicles</th>
<th>Screenshots of texts</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image9" alt="Exemplar Image 9" /></td>
<td><img src="image10" alt="Exemplar Image 10" /></td>
</tr>
<tr>
<td><img src="image11" alt="Exemplar Image 11" /></td>
<td><img src="image12" alt="Exemplar Image 12" /></td>
</tr>
</tbody>
</table>

**Computer vision libraries and applications**

As the field of computer vision evolves, an increasing number of commercial computer vision services have become available to ordinary users as well as researchers. With these online services, one can upload an image to a computer vision service via an API (application programming interface) and get the analysis back.

**Object recognition.** For an image, object recognition services return a list of descriptive tags that best match its content. Table 3.2 provides a comparison of three object recognition services—Clarifai, Imagga, and Google Vision—and summarizes their results on the same two pictures. At a first glance, the tags returned by these services are not completely accurate. For instance, for the picture of Hillary Clinton surrounded by a group of supporters, the algorithms cannot accurately distinguish well between a political gathering and a sporting event. For the photo of Donald
Trump making a speech, the algorithms sometimes mistakenly identify the person as a businessman (Table 3.2). However, the results do reveal some valuable information from the pictures, for example, whether the politician is alone or surrounded by a crowd. In addition, after extracting descriptive tags for each photo, we can also apply tools in computational linguistic analysis, such as semantic network analysis or topic modeling, to make better sense of the data (e.g., Manikonda & De Choudhury, 2017). This might be a better approach than looking at each tag individually. In image classification tasks, CNNs are trained to return a list of tags with highest probabilities, which are often semantically connected and visually similar. For instance, tags such as business, politician, and leader often appear together, which might be due to that they tend to describe individuals wearing formal suits in professional settings. By grouping semantically similar tags, one can come up with an overall theme that summarizes a category of tags, even if one particular tag is not very accurate.
Table 3.2 A comparison of four object recognition services

<table>
<thead>
<tr>
<th>Service</th>
<th>Top Tags (Confidence Levels)</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clarifai</td>
<td>Flag (0.999), Chair (0.995), Administration (0.994), Democracy (0.989), Politician (0.986), League (0.976), Leader (0.973), Presidential (0.960), Meeting (0.960), Home (0.950), People (0.947), Election (0.940), Freedom (0.887), Battle (0.848), Pride (0.839), Candidate (0.837), Business (0.829), Diplomat (0.813), Military (0.760), Diplomacy (0.751)</td>
<td>Confidence levels provided by computer vision services are shown in brackets. Based on the documentation provided by Imagga, a higher confidence means that the algorithm is more likely to be sure that a tag is correct. How they are exactly calculated, however, is not disclosed to users.</td>
</tr>
<tr>
<td>Imagga</td>
<td>Executive (65.59%), Businessman (53.94%), Business (45.71%), Man (44.30%), Male (37.95%), Office (35.54%), Suit (34.03%), Corporate (33.14%), People (30.24%), Men (29.07%)</td>
<td>Crowd (87.52%), People (22.84%), Man (19.21%), Group (17.32%), Team (16.85%), Male (16.51%), Adult (14.60%), Silhouette (13.31%), Audience (13.26%), Event (12.51%)</td>
</tr>
<tr>
<td>Google Vision</td>
<td>Speech (92%), Profession (83%), Official (81%), Public Speaking (74%), Suit (74%), Spokesperson (72%), Party Leader (60%), Diplomat (58%), Orator (57%), Entrepreneur (55%), Politician (53%), Staff (51%)</td>
<td>Crowd (98%), People (97%), Audience (91%), Fan (83%), Event (71%), Interaction (69%), Product (66%), Public Event (64%), Competition Event (62%), Recreation (51%)</td>
</tr>
</tbody>
</table>

**Note.** Confidence levels provided by computer vision services are shown in brackets. Based on the documentation provided by Imagga, a higher confidence means that the algorithm is more likely to be sure that a tag is correct. How they are exactly calculated, however, is not disclosed to users.

**Face detection.** Face detection is a specific example of object recognition that specializes in identifying human faces in digital imagery (Wong et al., 2001). Being
applied in a wide range of contexts such as surveillance, security, and social media, face detection might be one of the most mature computer vision techniques. Today, many commercial face detection services are also at researchers’ disposal, such as Face++, Microsoft Azure, and Google Vision. Besides identifying a face, face detection services also return the locations of detected faces in images, usually shown as face rectangles and face landmarks, which are the positions of important face components, such as the right corner of the left eye (see Figure 3.5).

Figure 3.5 Illustration of (a) face rectangle and (b) face landmarks

Facial recognition. Facial recognition refers to the process of recognizing identities of faces in digital imagery (Klare et al., 2012). Available web services provide two ways of identifying people in images. The first method is using “face verification” or “face search.” Face verification compares two faces and determines if they belong to the same person. Face search compares an input face to other faces in an existing face set and returns a face in the set that best matches the input face. Using this method, researchers need to prepare a face set beforehand that contains all
the identities they wish to target. For a new face, the service compares it to those in the existing face set and returns the identity that best matches the face. The second method is often named as “celebrities” mode. Some computer vision services have their own databases of celebrities and public figures. After the user uploads an image to a service, the service detects the face in it and compares the face to those in its own database.

**Facial analysis.** In addition, some face detection services also provide a list of facial attributes, such as age, gender, eye openness, hair color, and emotional expressions. Table 3.3 compares the facial attributes returned by five popular computer vision services. Some attributes can be calculated based on the spatial relationships among facial landmarks, such as eye openness. Other attributes, such as age, gender, and emotions, however, require the algorithms to be trained on a large pre-labeled dataset.

**Emotion detection.** Emotion detection is a specific example of facial analysis that identifies the facial expressions in images (Susskind et al., 2007). Most computer vision services use a emotion classification system proposed by Ekman (1992), categorizing facial expressions into one of the six basic emotions—happiness, sadness, disgust, anger, fear, and surprise—or neutral (i.e., absence of emotions). However, some services do not follow this convention. For example, Microsoft Azure has an additional emotion category called contempt and Google Vision does not provide measures of fear and disgust (Table 3.3).
<table>
<thead>
<tr>
<th></th>
<th>Face++</th>
<th>Microsoft</th>
<th>Google</th>
<th>Amazon</th>
<th>SightHound</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face rectangle</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Face landmarks</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Demographic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Age</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Face quality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Blur/sharpness</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Exposure/brightness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Occlusion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Noise</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td><strong>Eye status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eye openness</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Eye gaze</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Eye glasses</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Moth openness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Head pose</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Emotion</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smile</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin status</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Facial hair</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Hair</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Makeup</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Accessory</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

*Note.* Updated June 2019. For discrete emotions, A = angry, D = disgusted, F = fear, H = happy, Sa = sad, Su = surprised, N = neutral. Information was retrieved from:
https://console.faceplusplus.com/documents/6329465
https://westus.dev.cognitive.microsoft.com/docs/services/563879b61984550e40cbbe8d/operations/563879b61984550f30395236
https://cloud.google.com/vision/docs/detecting-faces
https://wwwww.sighthound.com/docs/cloud/detection/
**Accuracy and bias.** It is important to keep in mind the potential inaccuracies in different computer vision applications. Regarding face detection, it seems to be a relatively mature technique that has satisfactory accuracy. For example, Abdullah et al. (2015) claimed that the algorithms used in their study achieved an accuracy of 100% in discovering faces in their test dataset. In a study on Instagram photos (Bakhshi et al., 2014), the percentage of agreement between results from a face detection service (Face++) and crowdsourced workers’ judgment achieved 97% for whether an image has a face or not, 96% for having a female or a male face, and 93%–99% for categorizing a face into different age groups.

As for emotion detection, currently available computer vision tools can satisfactorily detect the presence of positive emotions like happiness but does not distinguish well among other types of facial expressions. For example, the computational algorithm in Abdullah et al., (2015) could detect smiles in images with a high accuracy rate (area under the ROC curve = 0.92). In an examination, a computer vision service, Microsoft Azure, performed well regarding detecting happiness (precision = 98%) and neutral expression (95%), but its performance dropped sharply for other emotions like sadness (66%), surprise (65%), anger (60%), disgust (23%), and fear (20%). In comparison, another computer vision service,  

---

2 Area under the ROC curve is a measure of accuracy. .90-1 = excellent, .80-.90 = good, .70-.80 = fair, .60-.70 = poor, .50-.60 = fail.
Sightbound, performed relatively better on labeling negative emotions (Dehghan, Ortiz, Shu, & Masood, 2017) (Figure 3.6). Given that emotional analysis has not been adequately tested, this dissertation uses crowdsourced workers to code the emotional expressions in images and compare human-coded results to results from different emotion detection services.

Figure 3.6 Confusion matrix tables between human-coded emotions (columns) and computer vision-calculated emotions (rows) for two computer vision APIs

![Confusion Matrix](image)

Note. Image from Dehghan et al. (2017).

There might also be potential biases in face detection and facial recognition algorithms regarding different demographic groups (Buolamwini & Gebru, 2018; Klare et al., 2012; Phillips et al., 2011; O’Toole et al., 2012). The accuracy of face

---

3 Rows represent emotion types predicted by computer vision algorithms and columns represent emotion types coded by humans. For example, 0.69 on the top left corner of the left figure means that among all the expressions identified as anger by humans, 69% were correctly classified as anger by the emotion detection service and the rest were misclassified as other types of emotions.
recognition algorithms largely depends on the training dataset. If the training dataset predominantly comprises faces from a certain demographic group, the algorithms trained on this specific dataset might not be very accurate for faces from other demographic groups (Klare et al., 2012). A study that compared face recognition algorithms from different countries showed that algorithms developed in East Asian countries were better at discerning Asian faces whereas algorithms in Western countries were better at recognizing Caucasian faces (Phillips et al., 2011). An analysis of six facial recognition algorithms has also shown that these algorithms had lower accuracy for females, Blacks, and individuals in the 18–30 age group (Klare et al., 2012).

**Computational aesthetics**

In addition to identifying content in images, computer vision tools can also extract aesthetical features from images, such as brightness, contrast, composition, color, texture, blur, and complexity (Bakhshi & Gilbert, 2015; Liu et al., 2016; Peng & Jemmott, 2018; Totti et al., 2014). These features are often irrelevant to the actual content in images, but can still meaningfully impact our perceptions of images (Dhar et al., 2011; Gelli et al., 2015; Guerini et al., 2013; Liu et al., 2016; Peng & Jemmott, 2018). For example, Peng and Jemmott (2018) demonstrated that computationally calculated features such as arousing–relaxing color, feature complexity, and color variety could successfully predict aesthetical appeal and likability of food images.
In addition, we can combine aesthetical features with results from face detection to formulate more sophisticated visual attributes. For example, research has shown that facial skin tone, such as lightness and redness, can shape viewers’ perceptions of individuals (Alter et al., 2016; Weaver, 2012; West et al., 2014). Therefore, we can locate the facial area in an image with face landmarks from facial detection and use aesthetical features specifically of that area to predict media bias in portraying politicians.

**Coding favorability**

This dissertation recruits crowd-sourced workers to rate how favorable a photo depicts a politician and uses their subjective ratings as criteria to determine which computer vision features actually make a difference in shaping viewers’ impressions of politicians. Yet, as noted earlier, simply asking a few human coders to rate how positively and negatively a photo portrays politicians might suffer from the issue of subjectivity (Groeling, 2013; Zeldes et al., 2008). This problem becomes more salient when a limited number of coders are used. For example, Barrett and Barrington’s (2005a) used three coders and Hehman et al. (2012) had six coders.

**Crowdsourcing**

With the advent of crowdsourcing platforms like Amazon Mechanical Turk (MTurk) and CrowdFlower, it has become possible for researchers to involve human coders
from a diversity of demographic and socioeconomic backgrounds to code media content on a large scale (Budak et al., 2016). For example, Budak et al. (2016) recruited 749 crowdsourced workers to code topic areas and ideological positions of 10,502 news articles.

Compared with traditional studies that usually involve research assistants to code media content, crowdsourcing provides several methodological advantages. First, this approach allows more coders to rate the same piece of content with relatively inexpensive costs (Berinsky et al., 2012; Buhrmester et al., 2012; Budak et al., 2016). Furthermore, compared with convenience samples like undergraduates in researchers’ own institutions, crowdsourced workers are often more diverse in their demographic and socioeconomic backgrounds (Berinsky et al., 2012; Budak et al., 2016).

Nevertheless, crowdsourced workers might differ from the general population regarding certain demographic characteristics and psychological traits. Compared with nationally representative samples, MTurkers tend to be younger, more liberal, and more likely to be Democrats (Berinsky et al., 2012). Given that coders’ perceptions of politicians are related to their political affiliations and gender attitudes, I plan to include an equal number of Democrats and Republicans, men and women to code the same content to balance out the impacts of coders’ prior attitudes. Crowdsourcing still relies on human coders’ subjective judgment of content. However, if there are an enough number of human raters from diverse demographic and ideological backgrounds to code each piece of content, averaging their ratings should
largely cancel out the impacts of raters’ pre-existing attitudes.

**Dimensions of favorability**

With a few exceptions (Hehman et al., 2012), most studies examining partisan media bias only treat bias as a single dimension that ranges from “unfavorable” to “favorable” (Moriarty & Popovich, 1991; Barrett & Barrington, 2005a; Budak et al., 2016). In contrast, most studies on gender bias have moved beyond focusing on the tonality of media coverage and started to examine how women and men politicians are covered qualitatively differently in the media (Devitt, 2002; Dunaway et al., 2013; Meeks, 2012).

As Grabe and Bucy (2009) argued, “visual analyses should move beyond the ‘positive versus negative’ index measures and investigate more specific and nuanced character frame–building dimensions” (p. 104). Simply coding the favorability of a piece of news content on one dimension might not capture the nuances in how different visual features shape viewers’ evaluations of politicians. For example, showing negative emotions like sadness in pictures might make a candidate look weak but also sympathetic in certain situations (Madera & Smith, 2009). A bottom-view camera angle might make a politician look powerful but simultaneously unapproachable and apathetic (Kepplinger, 1982). Therefore, we need to consider not only the favorability of a photograph but also potential dimensions of favorability.

Research on social cognition has converged on the claim that two fundamental dimensions are underlying our judgments of people (Abele et al., 2008; Cuddy et al.,
2008; Fiske et al., 2007; Rosenberg et al., 1968). Different schools of researchers
have assigned various names to these two dimensions, to name a few, communion
versus agency, warmth versus competence, socially versus intellectually good–bad,
social desirability versus social utility (for a review, see Abele et al., 2008). One
dimension (e.g., community, warmth) includes positive traits like being good-natured,
trustworthy, tolerant, friendly, and sincere, and negative traits like being unsociable,
cold, and dishonest; the other dimension (e.g., agency, competence) differentiates
positive traits like being capable, skillful, intelligent, and confident from negative
traits like being submissive, naïve, and inefficient (Abele et al., 2008; Cuddy et al.,
2008; Fiske et al., 2007; Rosenberg et al., 1968).

One popular theory, the stereotype content model (SCM), regards warmth and
competence as two central dimensions of social perceptions (Fiske et al., 2002, 2007;
Cuddy et al., 2007, 2008). These two dimensions do not always overlap. As illustrated
in Figure 3.7, while some social groups are consistently evaluated as negative (e.g.,
homeless) or positive (e.g., middle class) on both dimensions, other groups receive
incongruent stereotypes in a way that they are perceived as favorable on one
dimension but unfavorable on the other dimension (e.g., Asians, the elderly).

Making this distinction between warmth and competence is particularly
important when we take gender into account. Scholars have noticed that warmth and
competence largely overlap with traits that are stereotypically associated with
femininity and masculinity (Abele et al., 2008; Cuddy et al., 2008). Warmth is often
perceived as an advantage for women politicians, whereas competence is often associated with men politicians. Women leaders can be perceived as powerful and capable, but this often leads to a sacrifice on the warm dimension, resulting in a “competent but cold” impression (Fiske, Cuddy, & Glick, 2002).

Figure 3.7 Social groups mapped on the stereotype content model

Note. L = low, H = high, C = competence, W = warmth. Based on a survey of a representative sample of U.S. adults. Figure from Cuddy et al. (2007).

It is worth noting that some scholars do not completely agree with this two-dimensional model of social perception and have proposed some revisions to it. For example, some argue that both warmth and competence dimensions have two sub-dimensions. Regarding the warmth dimension, social warmth (e.g., sociable,
warm) should be separated from morality (e.g., trustworthy, sincere) (Goodwin et al., 2014; Goodwin, 2015). Regarding the competence dimension, competence (e.g., intelligent, competent) should be separated from dominance (e.g., powerful, dominant, assertive) (Chen et al., 2014). Furthermore, scholars have also proposed that attractiveness/charisma is a third dimension in perceiving politicians (Pancer et al., 1990).

In summary, this dissertation treats favorability as a multi-dimensional concept instead of a positive–negative dimension. Particularly, this dissertation asks crowd-workers to rate general favorability of politicians in images as well as some specific dimensions of perceived impressions, such as warmth and competence. The analysis also takes into account various trait dimensions in previous research and conducts an exploratory analysis on the dimensionality in perceptions of politicians.
Increasingly, media outlets are explicitly labeling their political affiliations to market themselves (Groeling, 2013). In the hope of exposing media consumers to more balanced and diverse viewpoints, scholars have made various attempts to automatically label political slants in media content (Garimella et al., 2017; Park et al., 2009). Given the difficulty in analyzing images on a large scale, prior research examining content bias across media outlets has been mostly limited to textual data (Groeling, 2013). Yet, visual content proliferates in the digital media environment and is widely used in political communication (Grabe & Bucy, 2009; Verser & Wicks, 2006). Empirical studies also demonstrate that visual attributes such as facial expressions and face size can effectively shape our impressions and voting preferences towards politicians (Mutz, 2007; Tiedens, 2001). This research gap regarding visual bias should require special attention today.

Applying computer vision techniques such as facial detection and emotional analysis, this study examines 13,026 images from 15 news websites covering the 2016 U.S. presidential candidates—Hillary Clinton and Donald Trump—regarding various visual features. Beyond this methodological contribution, this study also advances our understanding of visual bias in the following ways. First, while extensive research has been devoted to quantifying the direction and magnitude of partisan bias, relatively limited studies have examined the forms of bias in visual
content. Based on the ideological positions of media outlets established by prior research (Budak, Goel, & Rao, 2016; Flaxman, Goel, & Rao, 2016; Mitchell, Gottfried, Kiley, & Matsa, 2014), this study reveals the specific portrayals that media outlets use to convey their partisan views about the two candidates. In addition, this research also recruited crowdsourced workers to rate a subsample of images on their perceptions of media slant and impressions of the two candidates. By doing so, this study further illuminates how different visual representations adopted by partisan media potentially affect audiences of varying ideologies.

Theoretical framework

Measuring partisan media bias in visual content

This study defines partisan bias as systematic patterns in media content that favor one political party, candidate, or ideology over another (Groeling, 2013; Waldman & Devitt, 1998). First, bias should be distinguished from slant in a specific piece of media content, as bias should be “systematic, rather than anecdotal, episodic, or fleeting” (Groeling, 2013) (p.133). In addition, as Entman (2007) argues, to establish media bias, researchers need to show “patterns of slant that regularly prime audiences…to support the interests of particular holders or seekers of political power” (p. 166). Our operationalization of bias thus retires the notion that it is a deviation from the truth but instead incorporates its potential impact on audiences. If a piece of content is determined as favoring certain political actors, it is reasonable to expect
that viewers exposed to it should form favorable impressions about the targets or at least perceive the intended slant. While prior research has proposed other types of partisan bias such as coverage bias (which focuses on the volume of coverage), this study focuses on presentation bias, which specifically deals with the favorability of media coverage toward one party or ideology over the other (D’Alessio & Allen, 2000; Groeling, 2013).

As I discussed in chapter 2, previous attempts at measuring visual slant can be broadly categorized into two approaches. Some have used more objective measures by coding visual content on a checklist of features that researchers predefine as (un)favorable treatment of politicians. The criteria used to determine visual slant usually include a politician’s nonverbal behaviors (e.g., facial expressions, hand gestures, activity), contextual features (e.g., photographic settings, other objects and people in the same picture), and structural features (e.g., camera angle, color) (Grabe & Bucy, 2009; Moriarty & Popovich, 1991; Waldman & Devitt, 1988; Verser & Wicks, 2006). This “checklist” approach offers nuanced understandings of how bias is embodied in specific visual portrayals, although, as scholars have noted, different visual features do not contribute equally to the favorability of images (Barrett & Barrington, 2005a). It also remains unknown if these attributes selected by scholars indeed influence audience interpretations (Lobinger & Brantner, 2015).

Others have used more subjective measures by instructing coders to rate the favorability of media content. For example, in Barrett and Barrington (2005a), three
coders rated photos of politicians on a “highly unfavorable”—“highly favorable” scale.

In Hehman, Graber, Hoffman, and Gaertner (2012), six coders evaluated candidates in photos on warmth and competence. With crowdsourcing platforms like Amazon Mechanical Turk, scholars can recruit more coders from diverse demographic and ideological backgrounds to assess bias on a large scale (Budak et al., 2016).

Nevertheless, this approach still relies on coders’ subjective interpretations and answers only *how much* but not *how* media content is biased. As Grabe and Bucy (2009) argued, “visual analyses should move beyond the ‘positive versus negative’ index measures and investigate more specific and nuanced character frame–building dimensions” (p. 104).

These two approaches can be regarded as not only different methods for quantifying bias but also two routes of conceptualizing bias that complement each other. Partisan bias should first be established as systematic patterns of differential treatment of political actors in media content. However, differences alone do not guarantee favorability; it requires additional efforts to demonstrate that these patterns indeed (dis)advantage certain actors among audiences. This study thus integrates a content analysis that investigates if the visual coverage of candidates does vary by media outlets with a survey that asks crowdsourced workers to rate the favorability of these pictures. By doing so, this research hopes to reveal visual cues that (a) reveal media outlets’ ideological positions and (b) influence audience perceptions of favorability.
Selection of visual features

The first step in the approach this research proposes is to select visual features that could be captured by computer vision applications on a large scale and that should theoretically reflect media bias. Based on previous scholarship on visual bias, social cognition, and political psychology, this study focuses on the following features:

Facial orientation. Current facial detection algorithms often show facial orientation in three angles: pitch refers to the extent of a head bowing down or raising up; roll reflects the extent of a head tilting to the left or to the right; yaw shows the extent of a face turning to the left or to the right (Figure 4.1). The pitch angle can be seen as a proxy of camera angle, a criterion frequently used in prior research to indicate visual favorability. A low camera angle (face pitching upward) is regarded as a better portray compared with a high angle (face pitching downward), as it could convey a sense of dominance and power (Grabe & Bucy, 2009; Waldman & Devitt, 1988). It is unclear, however, if the other two angles—roll and yaw—are related to visual bias.
Face size and location. As a cue for judging interpersonal distance, face size might pose mixed impacts on perceptions of other people. On one hand, compared with a long shot, a close-up portrait makes a person seem closer to viewers, thus appearing more intimate or dominant (in the sense of face-to-face confrontation) (Archer et al., 1983; Grabe & Bucy, 2009). On the other hand, an extreme close-up might be a negative portrayal, as it resembles an extremely close physical distance that violates the notion of personal space. It also brings a person’s face under detailed scrutiny, revealing skin flaws or awkward expressions (Grabe & Bucy, 2009). In addition, prior research also claims that featuring a candidate dominating the photo or as the center of attention positively portrays the candidate (Verser & Wicks, 2006). A photo that locates a politician’s face closer to the center should cast the person in a better light.
**Facial expressions.** Facial expressions of emotion, motions or positions of facial muscles that convey the emotional state (Ekman & Friesen, 2003), have also been used to evaluate visual bias. Looking happy or confident is usually coded as a positive representation of politicians while frowning, looking sad, worried, or tired is seen as negative (Moriarty & Popovich, 1991; Waldman & Devitt, 1988). Happy faces are also perceived as more trustworthy, attractive, and dominant (Oosterhof & Todorov, 2008; Knutson, 1996; Sutherland et al., 2013). Besides coding facial expressions on a positive–negative spectrum, scholars have also argued for distinctions among discrete emotions. Ekman and Friesen (2003) identified six basic facial expressions: happiness, sadness, fear, anger, surprise, and disgust. Grabe and Bucy (2009) also distinguished among anger/threat, fear/evasion, and happiness/reassurance. Different negative emotions may produce distinct impressions: for example, individuals posing anger or disgust are often perceived as more dominant and powerful than those showing sadness or fear (Knutson, 1996).

**Eye and mouth status.** Eye status is another criterion frequently used to evaluate the favorability of photographs. Looking directly at the camera or at someone in pictures is coded as a positive portrayal of the candidate, whereas closed eyes portray a politician negatively (Moriarty & Popovich, 1991; Verser & Wicks, 2006). Research has shown that eye openness could make people look more intelligent and attractive (Talamas, Mavor, Axelsson, Sundelin, & Perrett, 2016). Mouth openness should also be an indicator of favorability. Prior research has regarded exhibiting dynamic
behaviors such as speaking as a favorable depiction of politicians (Verser & Wicks, 2006). Mouth openness also reflects the intensity of smiling, although it can also indicate yelling and shouting that convey aggressiveness.

**Skin condition.** Skin condition has rarely been examined in visual bias literature but its effects are frequently documented in face perception research. First, darker facial skin often leads to more negative judgment of a person (Alter, Stern, Granot, & Balcetis, 2016). An analysis of news articles showed that how positively an article portrayed a person correlated with the person’s skin lightness in the article’s visuals (Alter et al., 2016). Facial skin coloration is another important cue in impression formation. Skin redness and yellowness can be used as cues for inferring health status (Stephen, Smith, Stirrat, & Perrett, 2009). Signaling increased blood flow, facial redness is also linked to perceptions of attractiveness, dominance, and aggressiveness (Fetterman, Robinson, Gordon, & Elliot, 2011; Stephen, Oldham, Perrett, & Barton, 2012). Last, faces with healthier skin are often rated as more attractive, whereas skin imperfections such as wrinkles and uneven pigmentation make a person look older, less healthy, and less attractive (Fink, Grammer, & Matts, 2006; Jones, Little, Burt, & Perrett, 2004).

**Other people.** In prior research, presenting a cheering crowd or attentive colleagues together with a politician is often coded as a positive representation, whereas featuring the politician alone, with inattentive crowds or colleagues is seen as negative (Moriarty & Popovich, 1991; Verser & Wicks, 2006). Being accompanied
by other people in pictures would also make a person appear more attractive than
being photographed alone (Walker & Vul, 2014). Therefore, the presence of other
people in pictures, as well as their facial expressions and eye openness, can also be
proxies of media bias.

**Visual bias across liberal and conservative media outlets**

Having proposed a list of computer vision features that should theoretically reflect
media bias, this study then asks how bias would be embodied in differential
portrayals of politicians. Past research on partisan bias often focuses on whether the
media deviates from the norm of balance and exhibits an overall liberal or
conservative bias, but no consistent patterns have emerged. For example, regarding
television news, D’Alessio and Allen’s (2000) meta-analysis revealed a detectable but
small pro-Democrat bias, whereas Grabe and Bucy (2009) found a persistent
pro-Republican one. Given the rise of partisan media outlets that explicitly favor one
side instead of sticking to the norms of objectivity and balance, scholarship has
gradually shifted to quantifying bias at the individual media outlet level (Budak et al.,
2016). Regarding visual bias, research has shown that media outlets like newspapers
(Barrett & Barrington, 2005a) and websites (Hehman et al., 2012) indeed publish
photos that portray candidates they endorse more favorably than the candidates’
opponents. Therefore, we should expect that liberal media would portray Clinton
better than Trump and conservative media would act reversely. However, it remains
unknown what visual cues are adopted by media outlets of various political affiliations in their (un)favorable treatment of politicians, which is one focus of this study.

The analysis of partisan bias is further complicated by that the past election had both a female and a male candidate. Although this research can uncover differences in visual representations of Trump and Clinton, it is difficult to attribute these differences solely to partisan bias or gender bias, as some attributes related to favorability are also linked to gender. For example, faces are often shown as more prominent in visual depictions of men than those of women, a phenomenon termed as “face-ism” in prior research (Archer et al., 1983). Gender is also stereotypically linked to different emotions: anger and disgust are seen as more masculine, signifying aggressiveness and dominance, whereas happiness, sadness, and fear are seen as more feminine, showing friendliness or weakness (Plant, Hyde, Keltner, & Devine, 2000). Therefore, instead of focusing on the overall contrast between Clinton and Trump, this study asks if the differential treatment between the two candidates would vary by media outlets’ ideological positions, thus uncovering visual cues that signal their political orientation.

RQ1: Among the following visual features, which were used by liberal and conservative media to differently portray Clinton and Trump: facial orientation (pitch, roll, yaw angles), face size, face location, facial expressions (e.g., happiness, anger, sadness), eye openness, eye gaze direction, mouth openness, skin condition (lightness,
redness, yellowness, health), presence of other people, other people’s facial
expressions and eye openness?

**Favorability and its dimensions**

As noted earlier, the notion of “favorability” in media bias implies that biased content
should indeed advantage certain political actors among media consumers. One
important function of visual media in political communication is to convey cues that
help us judge politicians’ traits and characters (e.g., warmth, competence), which in
turn could influence voting preferences (Caprara & Zimbardo, 2004). Nevertheless,
the majority of media bias studies often see bias as a single
unfavorable-versus-favorable spectrum, which might not capture the diverse effects
of visual portrayals on viewers’ impressions. For example, showing the negative
emotion of anger might make a person look unfriendly but simultaneously dominant
(Knutson, 1996; Tiedens, 2001). In addition, different traits also correspond to gender
stereotypes: women are expected to be more friendly and kind, whereas men are
perceived to be more assertive and aggressive (Prentice & Carranza, 2002). This
study thus asks if different visual representations of politicians actually affect viewers’
perceptions of media slant as well as evaluations of politicians on separate trait
dimensions.

Prior research is still divided on what specific dimensions govern our judgment
of people from visual portrayals. Research in person perception claims that two
fundamental dimensions underlie our judgments of people: one dimension (communion) captures traits related to perceived intent and could be further divided into two sub-dimensions, warmth (e.g., sociable, friendly) and morality (e.g., trustworthy, sincere); the other dimension (agency) captures traits about perceived ability and incorporates two sub-dimensions as well, competence (e.g., intelligent, competent) and dominance (e.g., dominant, assertive) (Abele et al., 2016). Evaluating politicians could be seen as a specific case of person perception. Caprara and Zimbardo (2004) also found a two-factor structure in judging personalities of politicians, energy and agreeableness, which largely overlap agency and communion.

Research in face perception also proposes that we use multiple dimensions to infer traits from human faces. The first dimension, labeled as valence, incorporates traits related to warmth, morality, and competence, indicating an overall favorability in impressions. The second dimension reflects only dominance (Oosterhof & Todorov, 2008). Attractiveness has also been proposed as a third factor in perceiving human faces (Sutherland et al., 2013). In this study, news photographs of politicians include not only their faces but also social information such as their interactions with other people. Therefore, this study first examines the structure underlying viewers’ perceptions of candidates in images and then investigates if visual features influence these dimensions differently. Here, this study proposes two research questions regarding the potential effects of various visual features.

**RQ2**: Among the visual features proposed in RQ1, what features could best predict
viewers’ judgment of media slant in images of the two candidates?

RQ3: Among warmth, morality, competence, dominance, and attractiveness, what dimensions underlie people’s perceptions of candidates in news photographs? (RQ3a) And what visual features proposed in RQ1 could best predict audience perceptions of these dimensions? (RQ3b)

**Method**

**Data preparation.** Prior research has already placed a list of popular news websites on the liberal–conservative spectrum. This research combined insights from several recent studies: one that averaged crowdsourced workers’ perceived slant of each media outlet’s news articles (Budak et al., 2016) and two using aggregated political orientation of each outlet’s audience as a proxy of its ideological position (Mitchell et al., 2014; Flaxman et al., 2016). The sample included eight liberal sites (Daily Kos, Slate, The New York Times, The Huffington Post, The Washington Post, MSNBC, BBC, CNN), four relatively neutral sites (USA Today, Reuters, NBC News, The Wall Street Journal), and three conservative sites (Fox News, Breitbart, and TheBlaze). With Google search, the study searched for images of the two candidates limited to a specific site (e.g., “Hillary Clinton site:cnn.com”). A total of 20,702 still images were retrieved during November 25–29, 2016.⁴ All images were transformed

⁴ Although BBC is a British outlet, it attracts a sizable U.S. readership and has been classified as left-leaning by prior research (Budak et al., 2016; Mitchell et al., 2014),
to JPEG format and large images were resized so both width and height did not exceed 1000 pixels.

With a computer vision service, Face++, this study then identified images with visible faces of the two candidates. The analysis first prepared a face set that included the two candidates. Next, for each image, the analysis detected if this image contained faces (facial detection). For each detected face, Face++ compared this face to faces in the face set and returned the most similar-looking face with a confidence score of the two faces belonging to the same person (facial recognition). Pictures without faces of Clinton or Trump were excluded. Based on the URLs of images returned by Google, this study then determined the date of each photo. The majority of URLs embedded the date when an article or a photo was published. For remaining images, a combination of web scraping and manual checking was used. Images from 2015–

thus being included in our sample. Two strategies were used to address the concern about the potential bias in search engine results: (1) this study predefined a list of media sources and limited each search to one source, so whether Google prioritized certain sources should have limited impact on the results; (2) all the images returned by Google were retrieved, so whether Google ranked certain types of entries ahead of others should not generate substantial biases in the data.

The face set included the two candidates and their family members (e.g., Ivanka Trump), running mates (e.g., Tim Kaine), competitors (e.g., Bernie Sanders), third-party candidates (e.g, Gary Johnson), and other figures who frequently appeared in the campaign (e.g., Barack Obama). For each person, the face set included about 10 images that covered a diversity of facial expressions and luminance conditions. To validate results from Face++, the study manually coded 400 images selected from all the retrieved images on whether the image had visible faces of Clinton and Trump. The analysis tried a series of thresholds (.60, .65, .70, ..., .95) as cut-off points to accept if the face should be determined as Clinton and Trump. The overall accuracy rate was maximized when threshold .75 was used both for identifying Clinton (accuracy = 97.75%, false negative = 2%) and Trump (accuracy = 97.25%, false negative = 1.5%).
2016 ($N = 13026$; $6543$ for Clinton) were kept in further analysis (Figure 4.2).

**Figure 4.2 Data preparation procedures**

- **20791 search entries**
  - Animated GIFs and broken URLs excluded
- **20702 still images**
  - Facial detection
- **19581 images containing faces**
  - Facial recognition
- **14326 images containing faces of Clinton or Trump**
  - Date determined based on URLs
- **13026 images from 2015–2016**
  - Face set

**Participant ratings.** From Amazon Mechanical Turk, 596 U.S. crowdsourced workers who had completed at least 100 tasks with an approval rate above 98% were recruited to rate a subset of 1200 images randomly selected from the sample (40 for each candidate and each outlet). Each participant rated a random set of 20 images for each candidate. On average, each image received $19.9$ ratings ($SD = 1.66$, range =
Ratings of each image were averaged across participants and used in the later analysis (Budak et al., 2016; Oosterhof & Todorov, 2008).

Each image was rated on the following questions. (1) Perceived slant: on five-point scales, participants rated how an image negatively or positively (1 = extremely negative, 5 = extremely positive) and how unfavorably or favorably (1 = extremely unfavorable, 5 = extremely favorable) portrayed a candidate ($\alpha = .99$; $M_{\text{Clinton}} = 3.32, SD = 0.47; M_{\text{Trump}} = 2.84, SD = 0.58$). (2) Traits perceptions: on five-point scales (1 = strongly disagree; 5 = strongly agree), participants rated in each image whether the candidate looked friendly, warm, honest, trustworthy, dominant, assertive, competent, intelligent, attractive, youthful, and healthy. These adjectives correspond to various constructs proposed in prior research (Table 4.3; Abele et al., 2016; Oosterhof & Todorov, 2008; Sutherland et al., 2016). (3) Perceived facial expression: Given the subjectivity in human perceptions of facial expressions, this study also asked 396 participants to rate a subsample of 800 images to determine the extent to which the candidates’ face displayed anger, disgust, fear, happy, sad, surprise, and neutral emotion (1 = not at all, 5 = to the full extent). These ratings

---

Participants reported their gender (female = 50.0%), age ($M = 39.3, SD = 12.6$), race/ethnicity (multiple answers allowed: Hispanic = 7.6%; African American = 8.6%; Asian = 8.1%; White = 83.2%), education (high school or less = 7.9%; some college = 21.0%; Associate degree = 13.7%; Bachelor’s degree = 37.4%; Master’s degree or higher = 19.8%), political ideology ($M = 3.73, SD = 1.78$) (1 = extremely liberal, 7 = extremely conservative), and party affiliation ($M = 3.86, SD = 2.13$) (1 = strong Democrat, 7 = strong Republican). The last two ($r = .82$) were combined into one political orientation scale. A screening survey sampled a roughly equal number of women and men, Democrats (34.3%) and Republicans (33.8%).

71
were used to validate emotional analysis results from computer vision services.

**Computer vision analysis.** For each face, Face++ provided the yaw, roll, and pitch angles that represented *facial orientation*. Face++ also provided each face’s location in the picture as a facial rectangle that bounded the face region (Figure 4.2). *Face size* was calculated as the ratio between the size of the face rectangle and the size of the image. *Face location* was calculated as $1 - 2 \times \frac{d_1}{d_2}$, in which $d_1$ was the distance between the face rectangle’s center and the picture’s center, and $d_2$ was the length of the picture’s diagonal. A higher value indicated that the face was closer to the center. Face++ also provided the extent of *eye openness* and *mouth openness* as well as *eye gaze direction* as a three-dimensional vector. The analysis calculated the angle between the gaze direction and the image, so a higher value approaching 90 degrees would indicate the eye was looking towards the camera rather than looking elsewhere. Values for eye openness and eye gaze direction were averaged across both eyes.

Regarding *facial expressions*, prior research has shown that computer vision tools can accurately detect happiness but might not well identify other facial expressions (Dehghan, Ortiz, Shu, & Masood, 2017). The analysis thus compared emotions detected by four popular emotion analysis services—Microsoft, Face++, Sighthound, and Google Vision—with participants’ perceived emotions (Table 4.1). All services detected happiness well with Microsoft performing best ($r = .85$). Microsoft also predicted anger ($r = .54$) and neutral emotion ($r = .63$) relatively accurately. Yet, the
detection of other emotions such as fear and sadness was not satisfactory. Therefore, scores of anger, happiness, and neutral emotion from Microsoft were used and other emotions were excluded from the analysis. In case that Microsoft failed to detect a candidate’s face in a picture, Face++’s results were used.

Table 4.1 Correlations between human-perceived and computer vision services-detected facial expressions

<table>
<thead>
<tr>
<th></th>
<th>Face++</th>
<th>Microsoft</th>
<th>Sighthound</th>
<th>Google Vision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anger (.88)</td>
<td>.29***</td>
<td>.54***</td>
<td>.31***</td>
<td>.23***</td>
</tr>
<tr>
<td>Disgust (.81)</td>
<td>.25***</td>
<td>.36***</td>
<td>.09*</td>
<td>NA</td>
</tr>
<tr>
<td>Fear (.50)</td>
<td>.08*</td>
<td>.19***</td>
<td>.11**</td>
<td>NA</td>
</tr>
<tr>
<td>Happiness (.96)</td>
<td>.79***</td>
<td>.85***</td>
<td>.71***</td>
<td>.80***</td>
</tr>
<tr>
<td>Sadness (.64)</td>
<td>.04</td>
<td>.28***</td>
<td>.18***</td>
<td>.17***</td>
</tr>
<tr>
<td>Surprise (.77)</td>
<td>.24***</td>
<td>.39***</td>
<td>.23***</td>
<td>.36***</td>
</tr>
<tr>
<td>Neutral (.74)</td>
<td>.46***</td>
<td>.67***</td>
<td>.36***</td>
<td>NA</td>
</tr>
</tbody>
</table>

Note: Analysis was performed on images containing only one face of the candidate (N = 791). The sample size varied as some faces detected by Face++ were not detected by other services. NA = not available. Interrater reliability for each item is provided in brackets. * p < .05. ** p < .01. *** p < .001.

Face++ provided facial landmarks of each face, which were locations of important face components, such as face contour, eyes, and mouth (Figure 4.2). Based on facial landmarks, the analysis then identified the facial skin region as the facial region excluding eyebrows, eyes, nose, and mouth. The image was transformed into the CIELab color space. Lightness, redness, and yellowness of facial skin were calculated as the average L, a, b values of pixels inside the facial skin region (Stephen et al., 2012). In addition, Face++ also returned the likelihoods about a face’s skin
condition. *Skin health* was calculated as the difference between the likelihood that the skin was healthy and average likelihood that the skin was having different types of ill conditions such as dark circles and stains.\(^7\)

The number of other people’s faces detected by Face++ was used to indicate if a candidate was presented with other people or not. If the candidate was accompanied by multiple individuals in an image, means of other faces’ happiness and eye openness were used as predictors. For a small portion of images with multiple faces of the same candidate (e.g., Trump standing in front of a screen showing his face) (0.9%), face-related visual attributes were averaged across the candidate’s faces. The number of the candidate’s face was included as a control variable. A few computationally calculated aesthetical features were also included as control variables, including *brightness* and *contrast*, measured as the mean and the standard deviation of all pixels’ perceived luminance values; *colorfulness*, based on the combination of R, G, B pixels in the RGB color space (Peng & Jemmott III, 2018); *image size* and *aspect ratio*, measured as the product and the quotient of an image’s width and height.

---

\(^7\) To assess the accuracy in other computer vision features, on three-point scales (2 = in the middle/can’t decide), this study also manually coded a subset of 150 images on pitch (1 = head bowing, 3 = raising; \(r = .60\)), roll (1 = head tilting right, 3 = left; \(r = .72\)), and yaw (1 = facing right, 3 = left; \(r = .78\)) angles of the face; eye (\(r = .71\)) and mouth openness (1 = closed, 3 = open; \(r = .67\)); eye gaze direction (1 = looking elsewhere; 3 = nearly vertical to the image; \(r = .50\)); and skin health (1 = visible skin flaws, 3 = smooth and healthy; \(r = .48\)), which well correlated with computationally calculated features.
Table 4.2 Distribution of visual features across candidates and media outlets

<table>
<thead>
<tr>
<th></th>
<th>Candidate</th>
<th>Media outlet</th>
<th>Candidate × media outlet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Face orientation: Pitch</td>
<td>.14***</td>
<td>-.01</td>
<td>.02</td>
</tr>
<tr>
<td>Roll</td>
<td>-.11***</td>
<td>.02</td>
<td>-.02</td>
</tr>
<tr>
<td>Yaw</td>
<td>-.02</td>
<td>.01</td>
<td>-.01</td>
</tr>
<tr>
<td>Face size</td>
<td>.14***</td>
<td>.06***</td>
<td>-.12***</td>
</tr>
<tr>
<td>Face location</td>
<td>-.09***</td>
<td>.18***</td>
<td>-.02</td>
</tr>
<tr>
<td>Facial expressions: Happiness</td>
<td>-.46***</td>
<td>-.06***</td>
<td>.08***</td>
</tr>
<tr>
<td>Anger</td>
<td>.42***</td>
<td>.02†</td>
<td>-.06**</td>
</tr>
<tr>
<td>Neutral</td>
<td>.09***</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>Eye openness</td>
<td>-.30***</td>
<td>-.01</td>
<td>.04†</td>
</tr>
<tr>
<td>Eye gaze</td>
<td>-.22***</td>
<td>.03*</td>
<td>-.01</td>
</tr>
<tr>
<td>Mouth openness</td>
<td>-.11***</td>
<td>.00</td>
<td>-.03</td>
</tr>
<tr>
<td>Skin condition: Skin lightness</td>
<td>-.19***</td>
<td>.03**</td>
<td>-.04†</td>
</tr>
<tr>
<td>Skin redness</td>
<td>.40***</td>
<td>.01</td>
<td>.02</td>
</tr>
<tr>
<td>Skin yellowness</td>
<td>.08***</td>
<td>.04**</td>
<td>.02</td>
</tr>
<tr>
<td>Skin health</td>
<td>-.09***</td>
<td>-.05***</td>
<td>.07**</td>
</tr>
<tr>
<td>Number of other people’s faces</td>
<td>-.12***</td>
<td>-.03**</td>
<td>.07**</td>
</tr>
<tr>
<td>Other faces’ happiness</td>
<td>-.04</td>
<td>-.01</td>
<td>.02</td>
</tr>
<tr>
<td>Other faces’ eye openness</td>
<td>.13***</td>
<td>.02</td>
<td>.00</td>
</tr>
</tbody>
</table>

Note: $N = 13026$ ($N = 4939$ for other faces’ happiness and eye openness). Candidate (0 = Clinton, 1 = Trump). Media outlet (1 = liberal, 2 = neutral, 3 = conservative). Each line represents one regression model. Standardized coefficients are shown. † $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

Results

The study first examined if the differential treatment between Clinton and Trump would differ between liberal and conservative media regarding various visual features (RQ1). The analysis constructed a candidate variable (0 = Clinton, 1 = Trump) and a media outlet variable (1 = liberal media, 2 = relatively neutral, 3 = conservative media) based on ideological positions of media outlets quantified by prior research (see method section). A series of moderated multiple regressions were conducted
(Jaccard & Turrisi, 2003), with each regression using the candidate variable, the media outlet variable, and their interaction to predict one visual feature (Table 4.2). For each visual feature, a statistically significant coefficient of the candidate variable would imply that media overall portrayed the two candidates differently regarding that feature. As noted earlier, this difference could be a mix of partisan and gender bias. A significant interaction would imply that the size of this difference was moderated by media outlets’ political orientation, suggesting the presence of partisan bias.

Figure 4.3 Distribution of visual features across media outlets and candidates

Note: DK = Daily Kos. HP = The Huffington Post. NYT = The New York Times. WP = The Washington Post. UT = USA Today. WSJ = The Wall Street Journal. BB = Breitbart. Scales: happiness/anger [0–100], skin health [−100–100]; face size [0–100%].
In overall media coverage, compared with Clinton, Trump had larger faces (β = .14), showed less happiness (β = -.46) but more anger (β = .42), and was portrayed with less healthy facial skin (β = -.09) and with fewer other people’s faces (β = -.12, all ps < .001). As indicated by significant interactions, these gaps regarding face size (β = -.12, p < .001), happiness (β = .08, p < .001), anger (β = -.06, p = .002), skin health (β = .07, p = .002), and number of other faces (β = .07, p = .001), narrowed or reversed as media’s political orientation moved from liberal to conservative (Figure 4.3), implying that these attributes were adopted by outlets to differentially portray the two candidates. For example, regarding happiness, the gap between Clinton and Trump was 32 overall (on a 1–100 scale), being wider in liberal media (34) than in conservative media (27). The gap was most pronounced in Daily Kos (46) and least in Breitbart (23), two sites situated at two extreme ends of the ideological spectrum.

Face size served as another example. While liberal and relatively neutral outlets almost universally portrayed Trump with larger faces than Clinton, the gap diminished or reversed in conservative sites such as Fox News, TheBlaze, and Breitbart. A similar pattern occurred for skin health. With two outliers (The New York Times and The Huffington Post), liberal media, especially Daily Kos and Slate, portrayed Clinton with healthier facial skin, but this gap narrowed in conservative media.

Having shown what visual features were adopted by media outlets of varying positions as signals of their political leanings, the analysis then investigated if these
features indeed impacted viewers’ perceptions (RQ2). The interrater reliability (IRR) of participants’ ratings was calculated based on intraclass correlation coefficients (see Kim, 2014, p.167). The IRR for perceived slant was quite high (.87), suggesting a high degree of agreement among participants. An ordinary least squares (OLS) regression used computer vision features to predict averaged perceived slant, controlled for which candidate an image featured (Figure 4.4). Given that raters were randomly assigned to a large pool of images, it was unlikely that the raters’ characteristics would be exactly the same across all the images. The model thus controlled for the number of raters assigned to each image and aggregated characteristics of each image’s raters, including percentages of raters who were women and white, and means of raters’ age, education level, and political orientation. Detected neutral emotion highly correlated with happiness ($r = -0.71, p < .001$) and was therefore removed from the model. All variance inflation factors were below 3.

Among all the attributes, expression of happiness had the largest effect size ($\beta = .48$). Large face size ($\beta = -0.19$) and expression of anger ($\beta = -0.15$) were rated as negative portrays of candidates whereas skin health ($\beta = .09$; all $ps < .001$) positively contributed to favorability. These results showed a large overlap between visual features that differentiated liberal and conservative media and features that influenced audience perceptions of favorability. Based on unstandardized coefficients, a completely happy face, a completely angry face, and a face with perfectly healthy skin would impact a picture’s favorability by $+0.8$, $-0.7$, $+0.4$ on a 5-point scale. In
addition, mouth openness (β = .10, p < .001) and other faces’ happiness (β = .07, p = .002) also slightly enhanced perceived favorability.

Table 4.3 Principal component analysis of perceived traits

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Items</th>
<th>All (N = 1200)</th>
<th>Clinton (600)</th>
<th>Trump (600)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Communion (warmth)</td>
<td>Friendly</td>
<td>.87</td>
<td>-.38</td>
<td>.85</td>
</tr>
<tr>
<td></td>
<td>Warm</td>
<td>.89</td>
<td>-.35</td>
<td>.87</td>
</tr>
<tr>
<td>Communion (morality)</td>
<td>Honesty</td>
<td>.95</td>
<td>-.01</td>
<td>.94</td>
</tr>
<tr>
<td></td>
<td>Trustworthy</td>
<td>.96</td>
<td>-.01</td>
<td>.95</td>
</tr>
<tr>
<td>Agency (dominance)</td>
<td>Dominant</td>
<td>.29</td>
<td>.92</td>
<td>.29</td>
</tr>
<tr>
<td></td>
<td>Assertive</td>
<td>.35</td>
<td>.89</td>
<td>.35</td>
</tr>
<tr>
<td>Agency (competence)</td>
<td>Competent</td>
<td>.92</td>
<td>.24</td>
<td>.88</td>
</tr>
<tr>
<td></td>
<td>Intelligent</td>
<td>.90</td>
<td>.22</td>
<td>.85</td>
</tr>
<tr>
<td>Attractiveness</td>
<td>Attractive</td>
<td>.94</td>
<td>-.12</td>
<td>.93</td>
</tr>
<tr>
<td></td>
<td>Youthful</td>
<td>.85</td>
<td>-.24</td>
<td>.85</td>
</tr>
<tr>
<td></td>
<td>Healthy</td>
<td>.93</td>
<td>-.02</td>
<td>.92</td>
</tr>
</tbody>
</table>

Variance explained (%) | 66.5 | 22.4 | 62.2% | 26.6 | 61.1% | 25.1 |

Note: Interrater reliability for each item is provided in brackets. Factor loadings larger than .5 are shown in bold.

Next, this study looked at the dimensionality in the evaluation of traits from photographs (RQ3a). The IRR in perceiving different traits ranged from .64 to .89, indicating reasonably high agreement among participants, particularly for traits related to warmth, competence, and attractiveness, but less for dominance and morality (Table 4.3). A two-factor structure emerged from exploratory factor analyses. Two terms related to dominance (dominant and assertive) loaded on one factor (IRR = .68; α = .95; MClinton = 3.25, SD = 0.45; MTrump = 3.25, SD = 0.46). Warmth, morality, competence, and attractiveness did not form distinct concepts as some prior research
had implied; instead, these traits converged on another factor and was referred as 
valence (IRR = .77; α = .98; M_{Clinton} = 3.09, SD = 0.51; M_{Trump} = 2.57, SD = 0.41) 
(Table 4.3).

The analysis then looked at what visual features impacted audience perceptions 
on these two factors (RQ3b). Interestingly, the valence dimension highly correlated 
with perceived slant ($r = .93, p < .001$). Most features predicting perceived slant also 
influenced judgment on the valence dimension, including face size ($β = −.16$), 
happiness ($β = .43$), anger ($β = −.14$), mouth openness ($β = .10$), skin health ($β = .08$), 
and other faces’ happiness ($β = .07$, all $ps < .001$). The number of other faces also 
slightly increased valence ($β = .04, p = .03$) (Figure 4.4).

In contrast, the effects of visual attributes on dominance, which correlated with 
perceived slant only to some degree ($r = .23, p < .001$), showed a different pattern. A 
few attributes influenced dominance in the same direction of predicting valence, such 
as face size ($β = −.14, p < .001$) and mouth openness ($β = .09, p = .01$). Yet, anger ($β 
= .25$) increased dominance while happiness ($β = −.13$, both $ps < .001$) acted 
negatively, which were in the opposite directions of predicting valence. A completely 
happy and angry face would impact dominance by $−0.1$ and $+0.8$ on a 5-point scale. 
Skin redness ($β = −.09, p = .01$) and eye gaze ($β = .08, p = .02$) also had small effects 
on dominance (Figure 4.4).
Figure 4.4 Effects of visual features on perceived slant, valence, and dominance

Note: Regarding other faces’ happiness and eye openness, mean substitution was applied to photos with only the candidate’s faces. Candidate (1 = Trump, 0 = Clinton).

Discussion

In summary, this research advances our understanding of visual bias in the following ways. First, by integrating the objective checklist and subjective rating approaches developed in prior research, this study extends conventional analysis, which often focuses on the magnitude and direction of bias, to understanding how partisan media bias is constructed in visual portrayals of politicians and how these portrayals then
influence audience interpretations. Regarding audience perceptions, this research also extends visual favorability from a single positive–negative spectrum to a two-factor space that includes valence and dominance. Different visual cues can exert both similar and reversed impacts on perceptions of these two dimensions.

**Partisan bias in visual content.** With a few exceptions, visual features that differentiate liberal from conservative media largely overlap features that impact viewers’ perceptions of slant, including facial expressions, face size, and skin condition. Facial expressions, and particularly happiness, play essential roles in shaping participants’ perceptions of media slant and impressions of politicians. Indeed, in face perception literature, perceived happiness in faces and valence-related traits such as trustworthiness, intelligence, and attractiveness often load on the same dimension in factor analysis (Oosterhof & Todorov, 2008; Sutherland et al., 2013). In interviews, Lobinger and Brantner (2015) also found that viewers heavily relied on politicians’ facial expressions to judge the slant in news images. Therefore, computationally detected happiness in politicians’ faces could be a simplified but efficient proxy of visual slant. This measure might be particularly useful for information platforms that primarily circulate visual data, such as Instagram and YouTube.

One genre of negative portrayals in media coverage might require further attention, which often features candidates with a large face occupying almost the
entire image, highlighting their skin flaws and negative emotional expressions. Prior research in face-ism, which often compares full-body with half-body images, has shown a positive influence of facial prominence on impression formation (Archer et al., 1983). The negative impact of face size found in this study somewhat contradicts face-ism but echoes what Mutz (2007) has termed as “in-your-face politics” in television discourse and Grabe and Bucy’s (2009) interpretation of extreme close-ups. Close-up shots of politicians make them seem to be in the face of viewers and create a sense of discomfort and uneasiness among viewers, which could intensify viewers’ preexisting negative feelings towards them. Furthermore, in fashion and advertising, media professionals can manipulate skin condition (e.g., removing wrinkles) to make models look more attractive in photos. In political coverage, media outlets might also intentionally make a politician look unfavorable by highlighting facial skin flaws. We should also note that both candidates examined here are white. The effects of skin condition might be more complicated if politicians of other races are considered, which could be an arena for future research.

**Audience perceptions of favorability.** Echoing prior research, this study reveals

---

8 Face size negative correlated with skin health ($r = -.21$) and happiness ($r = -.09$), implying that media combined these features to portray a candidate positively/negatively. Using regressions, tests of interactions suggested that the impact of happiness (but not skin health) on slant ($\beta = -.04, p = .048$) and valence ($\beta = -.05, p = .012$) was larger for smaller faces. Given the potential mixed effect of face size, the quadratic term of face size was also tested and was significant in predicting slant ($\beta = .09, p = .007$) and valence ($\beta = .09, p = .009$), implying that extremely large faces did not further reduce favorability compared with fairly large ones.
the multiple dimensions people use to evaluate politicians and highlights the need to study the effects of visual portrayals on different trait perceptions (Grabe & Bucy, 2009). Confirming previous research in face perception, a two-factor structure emerged from the data (Oosterhof & Todorov, 2008). While perceived slant converges with the valence dimension, dominance forms a separate dimension that should require further attention. These two factors do not completely converge. For example, some pictures intended to make candidates look bad by emphasizing their angry expression and aggressive behaviors like yelling and shouting, might simultaneously render them as more powerful and dominant.

Prior research has frequently shown the effects of trait perceptions on voting preferences (Caprara & Zimbardo, 2004), but where do people get their impressions of politicians? This study suggests that visual portrayals of politicians might be one source of trait perceptions. In the results, across different media outlets, Trump has expressed more anger and less happiness than Clinton, which should make him look less favorable and friendly but more dominant and aggressive. Indeed, a study conducted before the election showed that people considered Clinton as more caring and competent and Trump as more dominant (Kakkar & Sivanathan, 2017). Furthermore, trait perceptions on different dimensions might have distinct impacts on candidate preferences. Voters might prefer more dominant, aggressive leaders to more competent or caring ones when feeling threatened by outsiders or in uncertain situations (Kakkar & Sivanathan, 2017). Future work could experimentally examine
how dominance in visual portrayals influences viewers’ judgment of candidates and subsequent voting behaviors.

The diversity in audience interpretations of visual content might also require our attention. As Figure 4.4 shows, photos of candidates that matched (mismatched) viewers’ political orientation were rated as more positive (negative) independently from visual attributes. This pattern was especially salient for Trump, a controversial figure who invited polarized responses from participants. As one crowdsourced worker commented at the end of the rating survey: “It was very difficult to rank Trump anything but the lowest in honesty no matter what the picture was.” We also see that compared with slant and valence, viewers’ agreement on dominance was relatively low. This shows that crowdsourced workers can reach an agreement regarding certain evaluations, but there might be some inherent variability in people’s interpretation of visual dominance, which requires further study. Future research could also study how other individual characteristics such as political knowledge and visual literacy might affect how viewers attend to and process visual portrayals of politicians.
Chapter 5 Gender bias in politicians’ self-presentations

While traditional political communication is often mediated by news coverage, social media platforms allow politicians to circumvent media outlets and directly communicate with citizens (Dennis et al., 2016; Enli & Skogerbø, 2013). Despite that female leadership is increasingly visible in the political world (Powell & Graves, 2003), women politicians are still receiving biased media coverage that could potentially disadvantage them in campaigns and reinforce gender stereotypes (Aday & Devitt, 2001; Devitt, 2002; Kahn, 1994). Yet, women politicians might utilize the affordances of social media sites to create self-presentations that best benefit themselves and combat stereotypical coverage they receive from news outlets. This study addresses the issue of gender bias by empirically examining U.S. politicians’ self-presentations on a popular visually oriented platform: Instagram. With 37% of U.S. adults reporting using this site, Instagram has become the third most used social media platform in the United States, following YouTube and Facebook (Pew Research Center, 2019).

In this study, two aspects of gender bias are examined. First, this research applies

9 It might be worth clarifying that this study treats politicians’ gender as gender expressions, instead of politicians’ biological sex or gender identity. In other words, this study examines how politicians’ gender is represented to the public and how the public perceives it. In previous research, both “female”/“male” (Aaldering et al., 2018; Bystrom et al., 2001) and “women”/“men” (Kahn, 1993; Meeks, 2012) were frequently used to describe politicians’ gender. This study uses “women/men politicians” more frequently to highlight the social aspects of gender.
computer-assisted image analysis to image posts \((N = 59,020)\) published by 172 Instagram accounts of U.S. politicians and compares women and men politicians’ self-portrayals on this platform. Second, social media users can collectively signal their preferences towards politicians’ self-presentation strategies by liking, commenting, and sharing social media posts. The analysis also investigates if people’s reactions to politicians’ different self-presentations are contingent on politicians’ genders. Last, this study also makes a methodological contribution by developing a technique that combines computer vision techniques and semi-supervised topic modeling (Jagarlamudi, Daumé III, & Udupa, 2012) to efficiently detect “visual topics” from large-scale visual data.

**Theoretical framework**

**Gender stereotypes and traits**

Warmth and competence (also named as agency and communion in previous literature) are two fundamental dimensions that are related to gender stereotypes (Cuddy et al., 2007; Fiske et al., 2007). Women are stereotypically expected to display warmth-related traits, such as being sociable, friendly, compassionate, and family-oriented, placing more importance on caring, cooperation, and interpersonal relationships. In contrast, men are traditionally associated with competence or agency-related traits, such as being ambitious, aggressive, intelligent, and rational, striving for leadership and achievement (Deaux & Lewis, 1984).
In the political communication domain, one possibility is that media coverage features politicians in congruence with their gender stereotypes, disproportionately emphasizing traditionally “masculine” traits (e.g., strong leader, intelligent, tough) for men politicians and highlighting “feminine” traits (e.g., compassionate, caring) for women politicians. Yet, empirical examination of media content often deviates from this pattern. An examination of Dutch newspapers revealed that while men politicians received more coverage on leadership-related traits (i.e., political craftsmanship, vigorousness) than their women counterparts, which were in accord with gender stereotypes, they also got more media mentions of a traditionally feminine trait, communicative skills (Aaldering & Van Der Pas, 2018). Furthermore, some studies have shown that media coverage of politicians might be drawn to their gender-incongruent traits (Gidengil & Everitt, 2003). For instance, when covering the 1993 and 1997 Canadian election campaigns, reporters used more affectively negative and aggressive verbs (e.g., attack, blast, fire) to describe women candidate’s speech (Gidengil & Everitt, 2003). Last, a few studies also revealed virtually no gender bias in trait coverage. A study of local newspapers’ coverage of the 2010 U.S. House campaigns showed no differences between women and men candidates regarding media mentions of four traits, i.e., competence, leadership, integrity, and empathy (Hayes & Lawless, 2015).

A growing body of literature also demonstrates that politicians’ self-presentations do not necessarily follow traditional gender stereotypes (Fridkin & Kenney, 2014;
Lee, 2013). The strategic stereotype theory (Fridkin & Kenney, 2014) posits that politicians strategically adopt stereotypes that benefit themselves. Women politicians strive to exploit stereotypes that link women to caring and nurturing traits while downplaying stereotypes that associate men with traits related to competence and leadership. For example, women candidates in the 2012 U.S. House elections were more likely to post “attack-style” tweets than their men competitors on Twitter (Evans & Clark, 2016). Lee (2013) revealed that U.S. congresswomen tended to emphasize their masculine traits—for example, portraying themselves as tough fighters—than their men counterparts in official biographies.

Conveying traits is a key function of visual content in political communication (Grabe & Bucy, 2009). This research then asks if the gender stereotypes would also emerge in politicians’ self-presentations on Instagram. Several visual portrayals should contribute to the perception of communal traits, particularly emotional expressions and social interactions. First, happy facial expressions should make a person look more friendly and caring. In previous research, viewers’ perceptions of happiness or smiling are often highly correlated with warmth-related traits, such as being friendly, sociable, caring, and trustworthy (Oosterhof & Todorov, 2008; Sutherland et al., 2013; Peng, 2018). In addition, presenting the politician interacting with other people should also make the politician look more sociable than featuring the politician alone. Therefore, the number of other faces in a picture and other faces’ happy facial expressions can also signal communal traits.
Personalization and politicians’ use of social media

In addition, one common strategy of delivering a relatable and likable image among voters is to use personalization in political campaigns. Personalization is an increasingly prevalent phenomenon in contemporary politics. Scholars have argued that the personalization of politics has two distinct components: (1) *individualization*, meaning that candidates or politicians, instead of institutions, parties, or issues, become central in the political process, and (2) *privatization*, meaning that politicians are shown as private individuals instead of occupiers of public roles and their personal, non-political characteristics that become more relevant in the political process (Adam & Maier, 2010; Holtz-Bacha et al., 2014; Van Aelst et al., 2012). Since this study focuses on politicians’ own social media accounts—which are already highly individualized representations of politics—it is the privatization aspect of personalization that is more related to the purpose of this study.

Prior studies have demonstrated that women politicians often receive more treatment of personalization than men politicians in media coverage. Women politicians often receive more coverage on their personal information—for instance, personality, attire, appearance, and family—whereas men politicians receive more coverage on their issue positions (Aday & Devitt, 2001; Kahn, 1994). Analyses of media coverage of Elizabeth Dole and her male opponents during her run for the Republican presidential nomination in 1999 showed that compared with other male
contenders, Dole received more coverage on her dress/appearance, personality, background, qualification but less on her issue stances (Aday & Devitt, 2001; Heldman et al., 2005). Similar patterns have also been found for candidates running for U.S. governors and senators (Bystrom et al., 2001; Dunaway et al., 2013; Devitt, 2002; Kahn, 1994).

Politicians might also engage in self-personalization on social media platforms by strategically disclosing personal information to build an intimate connection with voters (McGregor et al., 2017; Meeks, 2016). Nevertheless, empirical research has demonstrated that politicians still use social media in a traditional, “politics-as-usual” manner, for example, broadcasting policy and issue positions, showing professional activities, or mobilizing voters, whereas personalization often appears in a small percentage of social media posts (Evans et al., 2014; O’Connell, 2018). What’s more, women and men politicians generally do not differ much regarding their incorporation of personalized content. McGregor et al. (2017) examined Facebook and Twitter posts from gubernatorial candidates in 2014, showing that men candidates (10.6%) having slightly more personalized posts than women candidates (5.1%). Evans et al. (2014) examined House candidates’ use of Twitter in the 2012 campaign and revealed that 27% of tweets were categorized as personal, with no significant difference detected between women and men candidates (Evans et al., 2014). Bene (2017) analyzed candidates’ Facebook posts during the Hungarian general election campaign in 2014 and found that only 3.8% of posts could be categorized as personal. Meeks (2016)
analyzed U.S. Senate candidates’ Twitter feeds in the 2012 election and showed that women (11.2%) and men (12.6%) were similar in their use of personalization.

O’Connell (2018) revealed that among Instagram posts made by U.S. congresspeople, only 8.16% were categorized as personal, whereas a majority of posts (69.37%) were classified as professional. In addition, there was no significant gender difference (O’Connell, 2018). In summary, if we do detect a difference between women and men politicians’ use of personalization on Instagram, it is likely that this gap is not substantial.

Furthermore, while some research on media bias implies that personalized media coverage might bring detrimental effects to politicians (Aday & Devitt, 2001; Kahn, 1994), viewers often react positively to personalizing politicians. The act of self-disclosure, revealing personal information about oneself, often makes a person more liked by others (Collins & Miller, 1987). Scholars claim that personalized content tends to be more emotionally engaging and interesting and audience demand might help explain the increasing personalization in media coverage of politicians (Kaid & Strömbäck., 2008). A series of empirical studies have also demonstrated that personalization can help political candidates (Kruikemeier et al., 2013; Lee & Oh, 2012; McGregor, 2018; Meeks, 2017). In experiments, participants exposed to personalized tweets from a candidate were more likely to feel a social presence of the candidate and to report a sense of parasocial interaction, compared with those seeing issue-related tweets (McGregor, 2018; Lee & Oh, 2012). Experimental participants
also perceived a personalizing politician on Twitter as more likeable and more capable of handling various political issues than a depersonalizing candidate (Meeks, 2017). Analyses of social media data showed that candidates’ personal posts were more likely to gather audience engagement such as likes and comments, indicating that users respond positively to more intimate and emotional content on social media (Bene, 2017; Larsson, 2017; Metz et al., 2019).

**Face and face-ism**

Last, this study also pays attention to the presence of faces in images. Showing faces might be a simple strategy of cultivating close relationships with followers on social media, which might be particularly effective in a platform like Instagram that is characterized by a prevalence of selfies (Deeb-Swihart et al., 2017). Previous research has shown that on Instagram, pictures with faces are more likely to get likes and comments than pictures without faces, although it is unclear whether this tendency was due to people engaging with faces in general or faces they know of (Bakhshi et al., 2014). Studies in computer-mediated communication have also demonstrated that disclosing profile images can leave more favorable impressions among viewers and initiate communication of better quality (Feng et al., 2016; Tanis & Postmes, 2007). Yet, although there are some accounts of gender differences in selfie-posting behaviors (Dhir et al., 2016; Sorokowska et al., 2016), it is not clear whether women and men politicians would decide to show their own faces differently.
On a related note, faces are often shown as more prominent in visual depictions of men than those of women, a phenomenon termed as “face-ism” by researchers (Archer et al., 1983; Zuckerman, 1986; Copeland, 1989). Some scholars have argued that face-ism can be interpreted as a manifestation of “body-ism” that indicates a preference for viewing female bodies (Archer et al., 1983; Konrath & Schwarz, 2007). As our culture is still dominated by the “male gaze,” women are sexually objectified for men’s pleasure, which might lead to an overemphasis on women’s bodies instead of their faces (Zuckerman, 1986; Nigro, Hill, Gelbein, & Clark, 1988; Copeland, 1989; Fredrickson & Roberts, 1997). Using face-ism index as a measure of facial prominence (Figure 5.1), scholars have frequently found that women had lower levels of facial prominence than men across different contexts, including magazines, artworks, and entertainment programs (Archer et al., 1983; Copeland, 1989).

A few studies have also looked at whether face-ism prevails among pictorial representations of politicians. Sparks and Fehlner (1986) examined photos of the four 1984 presidential candidates appearing in Time and Newsweek but found that the then woman vice presidential candidate, Geraldine Ferraro, and other men candidates did not differ in terms of their facial prominence in images, contradicting previous findings. In a more recent study, Konrath, Au, and Ramsey (2012) examined politicians’ official online photographs from 25 cultures and discovered the pattern of face-ism in 18 of them. Szillis and Stahlberg (2007) also found the pattern of face-ism in online pictures of politicians in the German parliament.
As noted in the previous chapter, facial prominence might pose mixed impacts on our impressions towards other people. On one hand, compared with a long shot, a close-up portrait makes a person seem closer to viewers, thus appearing more intimate or dominant (Archer et al., 1983; Grabe & Bucy, 2009). On the other hand, an extreme close-up might be a negative portrayal, as it resembles an extremely close physical distance that violates the notion of personal space (Grabe & Bucy, 2009). As study 1 shows, media outlets used large face size to render negative portrayals of politicians and large face size indeed elicited negative evaluations from viewers.

In summary, this study proposes that gender bias in politicians’ self-representations might operate on two levels: whether there are systematic gender differences in politicians’ social media content and whether audience responses to different self-presentations are contingent on politicians’ genders. Therefore, I
propose two research questions.

**RQ1:** Do women and men politicians differ in their self-presentations on Instagram regarding the following aspects: the display of communal traits (facial expression of happiness, presence of other people and other people’s facial expressions), personalization, presence of their own faces and face size?

**RQ2:** Do audience members’ reactions (e.g., likes, comments) to different visual portrayal proposed in RQ1 differ by politicians’ genders?

**Data and method**

**Sample.** I generated a list of U.S. politicians in August 2018. The sampling procedure started with four groups of politicians: (1) presidential candidates in the 2016 U.S. election ($N = 29$), including 17 Republican, 6 Democratic, 4 third-party nominees (Gary Johnson, Jill Stein, Evan McMullin, and Darrell Castle), and two vice presidential candidates (Mike Pence, Tim Kaine), (2) Governors ($N = 50$), (3) Senators ($N = 104$), including four who ended their terms earlier, and (4) Cabinet members of the Trump administration ($N = 30$). Some politicians fell into multiple categories (e.g., Bernie Sanders). This study removed overlapping names and checked if each a politician had an Instagram account. Private accounts and accounts with fewer than 20 posts were excluded. One politician might have multiple accounts. A total of 176 accounts, representing 159 politicians (women = 19.5%, Democrat = 39.6%, Republican = 56.6%, mean age = 62.4), were kept in the sample.
Using Selenium, a web-scraping Python package, I retrieved each politician’s entire Instagram feed published before August 31, 2018, along with each post’s caption, publication date, and numbers of likes and comments it received. Videos were excluded from the dataset. Regarding “carousel” posts, a type of Instagram post that contained multiple photos for viewers to swipe through, only the first cover image was kept. The final sample included a total of 59,020 images. Each account had 337.1 posts on average ($SD = 416.3$, $Mdn = 220$).

**Facial analysis.** I performed facial analysis using Face++, a computer vision service specializing in facial recognition and analysis. The analysis contained two steps. I first prepared a target face set that included faces of all the politicians sampled in this study. For each image in our dataset, the facial recognition algorithm first detected if this image had a face or not. Next, for each detected face, the algorithm compared this face to all the faces in the target face set and returned the most similar-looking face with a confidence level. Based on facial analysis, I measured the number of faces belonging to the politician who owns the account, the number of other faces, the average happiness in the politician’s faces and other people’s faces, and how much area the politician’s face occupied the whole image, measured as the ratio between the size of the facial rectangle and the whole image.

**Visual topics.** This study proposes a semi-automated way to classify visual themes from a large number of visual images. First, I resorted to Clarifai, an image tagging service to transform each image into a list of twenty descriptive tags (Table
5.1) and then applied natural language processing techniques to further analyze the visual tags. Topic modeling is a common approach to uncover hidden topics from a large corpus of textual documents (Guo et al., 2016). Latent Dirichlet Allocation (LDA), for example, is a widely used technique to generate topic model solutions. Yet, due to its unsupervised nature, scholars often encounter situations where topics generated by algorithms are not meaningful or cohesive (Chang et al., 2009).

Table 5.1 Examples of Clarifai tags

| Leader administration people politician group election offense chair home adult parliament man business portrait mayor several religion candidate presidential many |
| text desktop illustration education abstract business facts graphic paper medicine number research science school conceptual alphabet type knowledge symbol internet |
| snow winter cold recreation skier ice people adventure woman fun mountain man goggles adult group_together child climber powder wear climb |

Recent approaches allow researchers to incorporate prior knowledge into the discovery of hidden topics (Jagarlamudi, Daumé III, & Udupa, 2012). Using this semi-supervised approach, researchers redefine *seed words* for each topic and nudge
the algorithm to generate solutions that align with researchers’ pre-conception of topics. This approach is particularly useful for this study. Based on previous content analyses about media coverage of politicians and politicians’ social media campaigns, we have already known that some “topics” should deserve our attention. For example, O’Connell (2018) conducted a manual content analysis of Instagram posts made by members of the United States and revealed three frequently used image types: 

personal posts (8.16%) that features a member’s personal life, with examples including “pictures of a child’s graduation, pictures of meals, or pictures of the member participating in a recreational activity like outdoor sports”; professional posts that captured the professional life of a member, including “meetings with constituents, shots taken at Congressional hearings, photos of a member interacting with the press, and so on”; and text posts, which were “infographics, memes, screen captures of tweets, news articles or press statements, or quotes superimposed upon a stock background image” (p.4). Therefore, adopting this semi-supervised topic modeling approach could ensure that certain topics would be included in the final topic-modeling solution.

When applying topic modeling, researchers usually have to remove stop words, which are commonly used words that are not relevant to the subject of inquiry. These words, if remaining in the dataset, often appear in multiple topics and do not help distinguish different topics. Many software programs provide a standard set of English stop words (e.g., a, for, the), which are not applicable to the context of this
study (Lucas et al., 2015). To generate a list of stop words to be removed from the data, I first calculated the frequencies of tags in the corpus. Visual tags that had appeared in more than 15% of images (e.g., *people, man*) were treated as stop words and then excluded. In addition, to generate seed words for each category, I ran a series of *unsupervised* topic modeling to identify which tags were used by Clarifai to describe different visual themes.

After settling down stop words and seed words, I then applied *semi-supervised* topic modeling to generate the topic solutions (Jagarlamudi et al., 2012). To decide the optimal number of topics, I repeatedly ran the analysis with the number of topics varying from 5 to 20. Although some metrics, such as perplexity and held-out likelihood, have been proposed to evaluate the congruence in topic models and to help determine the optimal number of topics, research has shown that topic solutions performing better on these measures are often less semantically interpretable to human raters (Chang et al., 2009). Therefore, I evaluated the consistency and representativeness of topics in each solution based on both (1) to what extent the

---

10 A total of 27 tags were removed with the threshold of 15%. These tags included generic descriptors of people or image content (e.g., *people, adult, man, group, woman, many, no person, one, travel, indoors, outdoors, furniture*) and tags that were highly relevant to politics (e.g., *leader, administration, election, politician*). There was no rule-of-thumb of determining the threshold in this situation, so the analysis also tried other thresholds such as 20% (removing 18 tags) and 10% (removing 44 tags). Choosing a higher threshold did not substantially alter the results except that a few common tags would show up in multiple categories. However, given that each image was described by 20 tags, choosing a lower threshold could remove all the tags from certain images and fail to assign topics to these images.
descriptive tags assigned to each topic were semantically coherent and comprehensible, and (2) to what extent the images classified into each topic indeed formulated a coherent and meaningful theme. Based on comparisons across different solutions, I found that the solution with 12 topics gave us most interpretable results, although choosing adjacent numbers of topics did not substantially alter the results.

Table 5.2 provides a summary of topics, descriptive tags, and representative images. These 12 topics could be further grouped into two large categories. The institutional politics category featured politicians in governmental buildings, legislative chambers, and press conferences, whereas the office/meeting category featured politicians in offices and conference rooms. Both categories reflected politicians’ professional activities. Relatedly, the rally/crowd, military, and commerce/industry also spoke to politically related activities, although events portrayed in these images, such as speaking to a crowd and visiting factories, took place in a variety of settings. The text/illustration category was a unique type of posts that politicians published to get their messages out, such as delivering policy positions and mobilizing people to vote. All these six categories (61.1%) were somewhat more related to the professional side of politicians’ lives.

In comparison, the other six categories were more related to the personal or private aspect of politicians’ lives. The sports category featured politicians in stadiums and sports fields or showed their interest in sporting teams. The family/friend/pet category showed politicians in private settings and intimately
interacting with their family members, friends, or pets. The *nature, architecture,* and *food* categories showcased the moments politicians found worth sharing. The *performance/stage* category was somewhat mixed. It usually featured people performing on a stage. Although many images were about musicians and other celebrities, some of them also portrayed politicians in conference settings, so this category was a mix of both professional and non-political images.
<table>
<thead>
<tr>
<th>Topic</th>
<th>Prevalence</th>
<th>Descriptive tags</th>
<th>Representative images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutional politics</td>
<td>16.8%</td>
<td>[parliament, candidate, presidential, press conference] democracy, banking, offense, ceremony, military, school, congress</td>
<td><img src="https://example.com/image1.jpg" alt="Image 1" /> <img src="https://example.com/image2.jpg" alt="Image 2" /> <img src="https://example.com/image3.jpg" alt="Image 3" /> <img src="https://example.com/image4.jpg" alt="Image 4" /></td>
</tr>
<tr>
<td>Office/meeting</td>
<td>12.9%</td>
<td>[office, conference room, teamwork] furniture, cooperation, sit, school, banking, technology, desk, table</td>
<td><img src="https://example.com/image5.jpg" alt="Image 5" /> <img src="https://example.com/image6.jpg" alt="Image 6" /> <img src="https://example.com/image7.jpg" alt="Image 7" /> <img src="https://example.com/image8.jpg" alt="Image 8" /></td>
</tr>
<tr>
<td>Rally/crowd</td>
<td>7.7%</td>
<td>[rally, crowd] offense, drag race, flag, battle, police, school, city, religion</td>
<td><img src="https://example.com/image9.jpg" alt="Image 9" /> <img src="https://example.com/image10.jpg" alt="Image 10" /> <img src="https://example.com/image11.jpg" alt="Image 11" /> <img src="https://example.com/image12.jpg" alt="Image 12" /></td>
</tr>
<tr>
<td>Category</td>
<td>Percentage</td>
<td>Keywords</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Military</td>
<td>5.9%</td>
<td>[military, army, police, veteran] war, offense, uniform, police, vehicle, ceremony, battle, soldier</td>
<td></td>
</tr>
<tr>
<td>Commerce/industry</td>
<td>6.1%</td>
<td>[industry, commerce, shopping, market] vehicle, transportation system, stock, exhibition, city, car, street</td>
<td></td>
</tr>
<tr>
<td>Text/illustration</td>
<td>10.3%</td>
<td>[illustration, text, symbol, design, horizontal, vertical, desktop, communication, vector] sign, internet, graphic, conceptual, design, facts</td>
<td></td>
</tr>
<tr>
<td>Family/friend/pet</td>
<td>10.4%</td>
<td>[family, child, friendship, dog, puppy, baby, couple, recreation, leisure, wedding] togetherness, happiness, fun, enjoyment, three, four</td>
<td></td>
</tr>
<tr>
<td>Category</td>
<td>Percentage</td>
<td>Relevant Keywords</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Sports</td>
<td>5.3%</td>
<td>[athlete, soccer, football] recreation, sports fan, crowd, race (competition), competition, championship, stadium, game</td>
<td></td>
</tr>
<tr>
<td>Nature</td>
<td>6.1%</td>
<td>[landscape, nature] travel, sky, water, summer, tree, wood, sunset, sea</td>
<td></td>
</tr>
<tr>
<td>Architecture</td>
<td>4.6%</td>
<td>[architecture, building] travel, city, sky, street, religion, urban, house, light</td>
<td></td>
</tr>
<tr>
<td>Food</td>
<td>3.6%</td>
<td>[food, drink] restaurant, table, celebration, furniture, one, desktop, Christmas, decoration</td>
<td></td>
</tr>
<tr>
<td>Performance/stage 10.1%</td>
<td>[music, performance, singer, concert, theater] recreation, ceremony, actress, stage</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* Tags in brackets are seed words for guided topic modeling.
Results

**Gender bias in visual self-presentations.** Our first research question (RQ1) investigates if gender differences existed in politicians’ self-presentations on Instagram. I aggregated photos’ characteristics by each politician ($N = 159$) and conducted a series of regression analyses to identify the effects of gender (coded as a dummy variable: women = 0, men = 1). Each politician’s age, party affiliation (coded as two dummy variables with Democrat as the reference group), and group (coded as three dummy variables with presidential candidates as the reference group) were entered as covariates.

We first looked at the results from facial analysis (Figure 5.2). On average, images from women politicians contained 0.46 more faces ($b = -0.46, p = .016$), 0.07 more politicians’ own faces ($b = -0.07, p = .008$), and 0.38 more faces of other people ($b = -0.38, p = .026$), compared with those from men politicians. Relatedly, women politicians were 6% more likely to post images containing their own faces than men politicians ($b = -0.06, p = .010$). Among images containing at least one face of the politician, women and men politicians did not differ regarding face size. In addition, women politicians displayed more happiness in photos than men politicians ($b = -0.19, p < .001$). The means of happy expressions (on a 0–1 scale) were quite high for politicians of both genders ($M_{\text{women}} = .77, M_{\text{men}} = .60$), which implied that politicians tended to deliver positive self-presentation on social media. Other faces from women
politicians’ accounts also displayed more happiness ($b = -0.07$, $p < .001$) than those from men politicians’ accounts. Overall, these patterns suggested that women politicians showed more sociable and communal traits than men politicians in our sample, being consistent with the gender stereotype literature.

Figure 5.2 Effects of politicians’ characteristics on image features

The analysis then looked at the topic distributions across women and men politicians. Figure 5.3 summarizes the results. Men politicians were 6.8% more likely to post images related to institutional politics ($b = 0.068$, $p < .001$), whereas women
politicians were 6.1% more likely to post images related to *family/friends* \( (b = -0.061, p < .001) \). Politicians of different genders also differed in terms of *rally/crowd* \( (b = -0.015, p = .037) \), *commerce/industry* \( (b = 0.017, p = .023) \) and *performancestage* \( (b = -0.016, p = .034) \), although these differences were generally small (less than 2%).

Women and men politicians did not differ significantly in terms of other post categories. Overall, the results indicated that women politicians were more likely to post personalized content on Instagram instead of traditional political content, although this difference was not large.
Audience responses to visual representations. I then ran multilevel regression analyses \(N = 59,020\) to predict what images features could lead to audience engagement for each image post. In addition to image characteristics such as the number of faces, the following variables were also included as fixed effects: (1) each post’s history (the number of days between the day when the account’s first post in the dataset was published and the day when this particular post was published) and duration (the number of days between this post’s published date and August 31, 2018); (2) characteristics of each post’s caption, including its length and the numbers of
mentions and hashtags; and (3) account characteristics, including numbers of followers, following, and posts. Accounts were entered as random effects. Two dependent variables, the numbers of likes and comments a photo received, were log-transformed.

The analysis first looked at the effects of whether the post emphasized politicians’ own faces. First, having faces had a positive effect on both likes ($b = 0.120, p < .001$) and comments ($b = 0.098, p < .001$). Given that the dependent variables were log-transformed, this implied that images with faces attracted 12.8% more likes and 10.3% more comments than images without faces. What’s more, featuring one politician’s face would lead to an increase in likes by 6.2% ($b = 0.060, p < .001$) and comments by 6.0% ($b = 0.058, p < .001$), respectively. However, one additional face not identified as the politician would decrease likes and comments by 1.1% ($b = -0.011, p < .001$) and 1.4% ($b = -0.014, p < .001$).

Next, the analysis looked at how audience members responded to different visual topics. The category institutional politics was treated as the reference group in the model. Compared with posts in the institutional politics category, images in the rally/crowd and military categories would increase viewers’ likes by 23.4% and 17.1% (both $ps < .001$). Images in the office/meeting, commerce/industry, and text/illustration categories also differed, though not substantially, from the reference category. Viewers preferred personalized topics. Compared with the institutional
politics category, images in the family/friend/pet, sports, nature, architecture, food, and performance/stage categories attracted more likes by 30.1%, 21.1%, 23.3%, 31.9%, 26.5% and 9.2%, respectively (all $ps < .001$) (Figure 5.4).

The pattern was slightly different for comments. First, categories related to professional categories did not differ substantially from the institutional politics category. One exception was the text/illustration category. When an image included a textual message, the number of comments increased by 25.7%. In addition, personalized categories such as family/friend/pet and food increased comments by 14.6% and 25.3%, respectively. Non-political categories such as architecture and nature, although being aesthetically appealing and likable, did not further lead to more comments (Figure 5.4).

Last, the analysis investigated the effects of politicians’ face size and facial expressions on audience engagement. I ran multilevel analyses only including images featuring politician’s faces ($N = 24,190$). The facial expression of happiness positively influenced likes ($b = 0.04, p < .001$) but did not affect comments, meaning that a completely happy would increase likes by 4%. Face size positively contributed to both likes ($b = 0.03, p < .001$) and comments ($b = 0.04, p < 0.001$): for one percent increase in face size, a post’s likes and comments would increase by 3% and 4%.
Figure 5.4 Effects of post features on audience engagement in multilevel analyses

Note. $N = 59,020$. Accounts ($N = 176$) were entered as random effects. Politician (gender, party affiliation, type, and age) and account characteristics (number of followers and followings) were included in the models as fixed effects but their effects were not illustrated, as some of these effects were too large to be shown in this figure.

**Are audience reactions gendered?** Next, I looked at if the effects of visual portrayals on audience engagement were moderated by politicians’ gender (RQ2).

The analysis first ran the multilevel models separately using images from women and men politicians. As I discovered before, women politicians posted more images in the *family/friend/pet* category, but fewer posts related to *institutional politics*. As revealed
in the results (Figure 5.5), audience members responded similarly to politicians of different genders when they posted images in the family/friend/pet category and the majority of other categories. We did see that men politicians seemed to benefit more from posting images related to commerce/industry, nature, and food. Therefore, the audience, in general, did not discriminate much regarding the effects of different visual characteristics on engagement except a few categories.

Figure 5.5 Effects of post features on audience engagement by gender

To test whether the regression coefficients were significantly different across two models, the analysis further conducted a multilevel regression using all observations and all variables used in Figure 5.4 as well as the interaction terms between gender (a dummy variable) and all the topic scores. These interaction terms indeed reached statistically significant.
Discussion

In summary, this research advances our understanding of gender bias in politicians’ self-portrayals on Instagram by investigating bias at two levels. This research both (1) compared women and men politicians’ self-presentations on Instagram and (2) examined if audience reactions to different portrayals were influenced by politicians’ genders. In addition, this research also demonstrates the potential of combining computer vision and natural language processing techniques to automatically uncover visual themes from a large amount of visual data.

Gender bias in politicians’ self-representations on Instagram. Our first observation is that the gender differences among politicians’ self-presentations are not as salient as previous research in media bias has suggested (Aday & Devitt, 2001; Devitt, 2002). Yet, this finding was in large consistent with prior studies on politicians’ use of social media platforms (Evans et al., 2014; McGregor et al., 2017; O’Connell, 2018). The most striking difference is that women politicians expressed more happiness than men politicians in their Instagram images with a gap of 19%. Posts from women politicians also contained more faces in general and their own faces and were likely to be about personalized category (family/friend). These findings indicate that women politicians indeed displayed more sociable and friendly traits in the sample than men politicians, although these differences were generally not large. Last, the study also failed to find a difference regarding facial prominence between
politicians of different genders, contradicting previous studies on face-ism (Archer et al., 1983; Copeland, 1989).

In addition, gender bias can also be manifested in audience responses to similar visual portrayals of politicians. The likes and comments can be seen as endorsements of social norms about what representations of genders are appropriate. Nevertheless, we also did not observe a substantial difference in terms of how people engaged with social media posts made by politicians of different genders. The effects of different visual portrayals on likes and comments were generally similar between women and men politicians. The perception bias was not salient in this study.

**Audience reactions to visual self-representations.** This research also examined how audience members responded to different visual representations of politicians. First, previous research has shown that on social media, pictures with faces are more likely to get likes and comments than pictures without faces (Bakhshi et al., 2014). It is unknown, however, if this tendency is due to that people engaging with generic faces or that people are attracted to account holders they are following. This research showed that although the presence of faces did increase audience engagement, it was politicians’ own faces that enhanced audience engagement, indicating that social media users react positively to images with faces primarily because they can recognize the faces in images and build a connection with these images.

In addition, I also looked at the effects of different content categories on audience
reactions, with an emphasis on the role of personalization. A significant portion of politician’s Instagram posts still fell into the traditional, “politics-as-usual” type of political communication, showcasing politicians’ professional activities and political statements, but this type of content generally was less successful in attracting followers’ engagement. Instead, we observe that personalized, non-political content attracted more likes and comments from audiences.

Previous research analyzing gender bias in media coverage of politicians often assume that personalization can disadvantage politicians and has revealed that news stories of women politicians tend to use a personal frame, emphasizing their non-political aspects, such as personality and family (Aday & Devitt, 2001; Devitt, 2002). Yet, recent experimental studies have shown that personalization might have positive effects on evaluations of political candidates, showing that potential voters might develop more emotional connections with candidates adopting a personalized campaign strategy (Lee & Oh, 2012; McGregor, 2018). Along with other studies using social media data (Bene, 2017; Metz et al., 2019), this study also shows that audience members react more positively to personalized content than political content. Certainly, it is unclear if audience likes and comments can transform into more substantial political outcomes such as voting, but these findings in combination should lead us to rethink the assumption that personalization only negatively affects political candidates.
This study also reveals that some content categories had divergent effects on the numbers of likes and comments, indicating that social media users might have different psychological motivations when liking and commenting on social media posts (Alhabash & McAlister, 2015; Peng & Jemmott, 2018). Alhabash and McAlister (2015) argued that liking is less cognitively demanding than commenting on a message. Additionally, liking signals a positive evaluation of a media message while comments can be of mixed sentiment. In the results, images conveying textual messages, while not getting many likes, led to more comments. This revealed that textual statements added to visual images could provoke more thoughts or cognitive efforts among viewers. In comparison, images of nature/landscape and architecture, while attracting a large number of likes, failed to get many comments. This type of images might be aesthetically likeable and serve as “eye candy” in users’ newsfeed, but does not provoke conversations among followers.

While the majority of studies on politicians’ use of social media have focused on Facebook and Twitter, this study provides valuable insights into Instagram, a popular platform that has been only infrequently examined in this line of scholarship (but see Larsson, 2017; O’Connell, 2018). Certainly, we should note how this particular platform might shape the observations of this study. As previous scholars have pointed out, each social media platform is characterized by a unique set of affordances and used by people with certain demographic characteristics and
psychological motivations (Alhabash & Ma, 2017; Bossetta, 2017; Oz et al., 2018).

Thus, certain self-presentation strategies might see appropriate and favorable on one platform might not work well in other venues. The majority of Instagram users turn to this site not for political content. Only 32% percent of Instagram users reported getting news on this platform, compared with 71% on Twitter and 67% on Facebook (Matsa & Shearer, 2018). This might contribute to the pattern that nonpolitical content performs better in getting people’s attention on this platform. With regards to interaction behaviors, Instagram prioritizes liking over other behaviors such as retweeting or reposting. Therefore, it might encourage more positive self-representations strategies. In comparison, negative and attacking-style messages should resonate more with other platforms such as Twitter. Indeed, Lee and Xu (2017) showed half of Clinton and Trump’s tweets during the campaign period were attacks and attacking posts also got more retweets and favorites. Future research might look beyond one specific social media platform and compare audience responses to politicians’ self-presentations across different sites.
Chapter 6 Concluding Remarks and Future Directions

In the introduction, I summarized the main goals of this dissertation as the follows: (1) to identify visual features that best discriminate positive and negative depictions of politicians, (2) to investigate two types of bias: partisan media bias, or how liberal and conservative media outlets portray the same politicians, and gender bias, or women and men politicians are presented differently in the pictorial domain, and (3) to examine the methodological potential of computer vision techniques in communication research. I then offered an overview of current advances in the field of computer vision that could be applied in social science research. The first study investigated partisan media bias in visual representations of two major presidential candidates in the 2016 US election, revealing visual features adopted by media outlets across the political spectrum to signal their ideological positions and visual features that effectively shaped viewers’ perceptions of favorability. The second study looked beyond media contexts and probed into politicians’ self-presentations on one visually oriented platform, Instagram, demonstrating how women and men politicians converged and diverged on their choices of self-presentation strategies and how social media users reacted to various types of self-portrayals. In this chapter, I open a discussion about this dissertation’s theoretical and methodical contributions as well as a few promising directions for future research.
Theoretical contributions

First, this dissertation offers us a chance to reflect upon the concept of media bias.

Previously, content analyses were frequently adopted in the media bias scholarship in the hope of exposing systematic patterns in media coverage of politicians of different political affiliations, genders, or ethnicities (Banning & Coleman, 2009; Moriarty & Popovich, 1991). Yet, scholars often make assumptions about what makes a favorable/unfavorable portrayal of politicians without actually testing the effects of certain media representations (but see Barrett & Barrington, 2005b; Boomgaarden, Boukes, & Iorgoveanu, 2016). As Entman (2010) argues, establishing content biases requires scholars to show “patterns of slant that regularly (and perhaps without an audience’s conscious awareness) promote support for some interests or actors who seek power and disapproval of their opponents” (p. 393). If some media content is regarded as biased, it should “assists such entities as political parties or interest groups in consistently persuading people to accept interpretations helpful to the favored actor for some significant period” (Entman, 2010, p.393). Therefore, in addition to establishing systematic, large-scale patterns in media treatment of politicians, we should further think about media bias in terms of biased content patterns’ potential influence on audience members. Consequently, in study 1, I look at not only how visual portrayals of presidential candidates differ across media outlets
on the liberal—conservative spectrum but also how these visual portrayals contribute
to crowdsourced workers’ perceived favorability and impressions. In study 2, instead
of looking at crowdsourced workers’ ratings, I use audience reactions on social media
(e.g., likes and comments) as proxies of audience favorability or interest. Examining
content patterns and audience reactions in combination offer us a chance to reflect
upon our assumptions about whether certain media representations should be
regarded as favorable or unfavorable.

One type of visual portrayal, personalization, serves as an example. Previous
research in the gender bias literature often assumes that personalized media coverage
could disfavor political candidates. However, empirical research also notes that
personalization can elicit more positive impressions among viewers and make
viewers feel more connected with a politician (Lee & Oh, 2012; McGregor, 2018).
Results from Study 2 also demonstrate that politicians can gain more audience
engagement when they post personalized, non-political images. Nevertheless, it is
possible that by personalizing themselves on social media, politicians might present
themselves as relatable and approachable, but this strategy might also render
themselves as less serious contenders and distract viewers from more substantial
issues such as politicians’ policy positions. Experimental studies have revealed an
*innuendo effect* in person perception: when media coverage of a politician only
focuses on one of the two trait dimensions, warmth and competence (i.e., only as
friendly or only as competent), viewers would rate the politician more negative on the other dimension (Koch & Obermaier, 2016). In addition, the positive effect of personalization was found to be contingent on viewers’ characteristics (Lee & Oh, 2012). While personalization might produce positive effects among less politically involved viewers, politically engaged voters could be less drawn to personalized coverage and further develop negative feelings towards politicians (Lee & Oh, 2012). Nevertheless, we should note that politicians on social media platforms rarely use personalized content alone. Echoing previous research (Evans et al., 2014; O’Connell, 2018), this dissertation reveals that this type of content is often outnumbered by more traditional, “politics-as-usual” posts. It is unlikely that personalized posts on Instagram would overshadow politicians’ delivery of policy stances or display of more politically relevant qualities. In summary, the assumption that personalization is a negative (or positive) portrayal should not be taken for granted. Furthermore, while previous research often compares personalization with issue-oriented presentation (e.g., McGregor, 2018), future studies should examine how different types of self-presentation strategies in combination—which represents a more common situation on social media—affect our evaluation of politicians.

Conclusions of this dissertation also support that we need to extend our conceptualization of bias from a positive-versus-negative spectrum to incorporate different components of favorability. Study 1 affirms that viewers use two dimensions
—valence and dominance—to evaluate politicians in images, in accord with previous research in face perception (Oosterhof & Todorov, 2008), Yet, a third dimension—attractiveness – youthfulness—that has been found in prior research (Sutherland et al., 2013; Vernon et al., 2014), has not emerged from the data. This might be due to that the study only sampled two politicians—Donald Trump and Hillary Clinton—and did not include a wide range of faces varying in age or attractiveness. Results also highlight that visual attributes often have complicated, and sometimes, divergent effects on different dimensions of trait perceptions (Grabe & Bucy, 2009). Some pictures making candidates look unfriendly by emphasizing their angry expression and aggressive behaviors like yelling and shouting, might also make them look more powerful and dominant. This observation should provide some practical implications. For example, designers of image-related services who wish to inform users of potential biases in visual content should be aware that visual bias can extend beyond one positive-versus-negative spectrum and illustrate how images can affect favorability in multiple dimensions.

This dissertation also highlights the importance of further bridging partisan bias and gender bias research. Gender bias might operate on multiple levels. First, media content might associate different norms with women and men politicians. In the first study, Trump was portrayed with larger faces, less happiness, more anger, and fewer people around than Clinton. This could result from an unfavorable treatment of
Trump across media outlets as well as gender bias in face-ism and gender stereotypes that expect women to be more friendly, sociable, and less aggressive (Archer et al., 1983; Prentice & Carranza, 2002). Indeed, in study 2, we also observed a similar pattern that women politicians exhibited more sociable traits than men politicians (displaying more happiness, being shown with other people). Moreover, the effects of visual cues on audience perceptions might differ between women and men politicians. Additional OLS regressions using photos of Clinton and Trump separately partially affirmed this possibility (Figure 4.4). For example, the positive impact of happiness on perceived slant was more salient regarding Clinton’s images than Trump’s, potentially reflecting the gender norm in facial expressions that it was more proper or rewarding for women than men to smile (Plant et al., 2000). The negative effect of face size on dominance was also more pronounced for Clinton than Trump. Nevertheless, in study 2, when the sample included more politicians of different genders, I failed to find substantial differences in people’s reactions towards women and men politicians’ self-portrayals.

**Methodological reflections**

Beyond contributing to a theoretical dialogue with the media bias literature, this dissertation also illustrates the potential of computer vision techniques in communication research. Overall, computer vision methods provides several
methodological advantages: (1) The use of computer vision tools greatly expands the scope of visual analysis and allow researchers to tackle the challenge of large-scale visual analysis; (2) it also provides standardized ways of coding certain visual attributes, such as skin tone, which could be compared across different studies; (3) it provides a combination of supervised learning and unsupervised learning techniques with which researchers can both code certain visual representations to test specific hypotheses and explore the hidden pattern in a large collection of visual data.

Together, computer vision tools better equip communication scholars to study visual content, which is becoming increasingly ubiquitous in our media environment.

Nevertheless, there are also several limitations researchers should be aware of. First, the majority of facial detection and analysis services only work well with frontal or near-frontal faces and might be unable to detect faces that are non-frontal (e.g., a profile face, a rotated face), partially blocked (e.g., wearing a hat), or of low image quality (e.g., blurred, too dark, too small). If algorithms fail to detect faces in a photo, this might already indicate an unfavorable portrayal of politicians (e.g., blocking a politician’s face). Figure 6.1 presents a few examples of photos that might be difficult for facial detection/recognition algorithms to analyze. Therefore, the findings of this dissertation should only apply to images featuring identifiable faces of politicians.
We should also be aware of the potential biases in computer vision algorithms. As noted in chapter 3, research has demonstrated that facial analysis algorithms might not work well for certain minority groups (Klare et al., 2012; Phillips et al., 2011; O’Toole et al., 2012). For example, Buolamwini and Gebru (2018)’s examination of three gender classification algorithms showed that while these algorithms performed relatively well (with error rates ranging from 0.0% to 12.0%) regarding light-skinned males, light-skinned females, and dark-skinned males, their accuracies dropped sharply for dark-skinned females (error rates = 20.8–34.7%). This low performance for certain minority groups was often due to that many current algorithms are not trained on demographically diverse datasets and can be improved by using more inclusive datasets (Buolamwini & Gebru, 2018; Klare et al., 2012). The same issue
might also happen regarding object recognition algorithms. For example, in study 2, we observed a 6.8% gap between women and men politicians’ images in the institutional politics category. If the object recognition algorithm was trained on a biased dataset that exaggerated the stereotypical association between men and politics, the actual gap should be even smaller.

To check if there were systemic biases in computational analysis of faces from various demographic groups, I selected 20 politicians of different genders and a diversity of racial/ethnical backgrounds and randomly sampled 20 images for each politician. 12 I manually coded (1) whether the politician’s face was present in the picture (any part of the politician’s face is shown) and (2) whether the politician’s face was completely shown (with two eyes, nose, and mouth visible to viewers), and then compared them with facial recognition results. First, there was no misidentification of identities in the facial recognition results (precision = 100%), meaning that if the algorithm determined that a picture featured a politician, it was

12 Certainly, race and ethnicity are fuzzy categories. I based the selection on Wikipedia pages that curated minority politicians:
https://en.wikipedia.org/wiki/List_of_minority_governors_and_lieutenant_governors_in_the_United_States
https://en.wikipedia.org/wiki/Asian_Americans_in_government_and_politics
https://en.wikipedia.org/wiki/List_of_Asian_Americans_and_Pacific_Islands_Americans_in_the_United_States_Congress
https://en.wikipedia.org/wiki/List_of_Hispanic_and_Latino_Americans_in_the_United_States_Congress
https://en.wikipedia.org/wiki/Category:Hispanic_and_Latino_American_state_governors_of_the_United_States
highly likely to be correct. However, the algorithm was not able to pick up all the pictures that featured the politician, which was largely due to that the politician’s faces were often incomplete and non-frontal (see Figure 6.2). In these pictures, politicians were usually interacting with other people and only showing profile faces. For complete faces, the facial recognition algorithm achieved a high recall rate ($M = 89.5\%$). although there was no clear pattern of discrimination against women politicians or politicians of racial minorities. A close inspection revealed photos that were misclassified mostly featured faces that were too small (Figure 6.3, a, b) and partially blocked (c, d). Overall, the results indicated that the facial recognition algorithm in study 2 showed no clear bias against minority politicians when politicians’ faces were complete and of good quality, but it would be inaccurate for politicians whose pictures frequently featured non-frontal, incomplete, and small faces.
Figure 6.2 Examples of pictures that have the politician’s face but the face is not complete

A  b  c

Figure 6.3 Examples of pictures featuring the politician’s complete face but the face was not recognized by the algorithm

A  b  c

E  f  g
Table 6.1 A comparison between manually coded results and facial recognition results

<table>
<thead>
<tr>
<th>Politician</th>
<th>Gender</th>
<th>Race/ethnicity</th>
<th>Face present (FP)</th>
<th>Face complete (FC)</th>
<th>Face recognized (FR)</th>
<th>FC/FP (%)</th>
<th>FR/FC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mazie K. Hirono</td>
<td>Woman</td>
<td>Asian</td>
<td>11</td>
<td>11</td>
<td>9</td>
<td>100</td>
<td>81.8</td>
</tr>
<tr>
<td>Tammy Duckworth</td>
<td>Woman</td>
<td>Asian</td>
<td>11</td>
<td>9</td>
<td>8</td>
<td>81.8</td>
<td>88.9</td>
</tr>
<tr>
<td>Kamala D. Harris</td>
<td>Woman</td>
<td>Black/Asian</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>100</td>
<td>80</td>
</tr>
<tr>
<td>Catherine Cortez Masto</td>
<td>Woman</td>
<td>Hispanic</td>
<td>18</td>
<td>14</td>
<td>12</td>
<td>77.8</td>
<td>85.7</td>
</tr>
<tr>
<td>Nikki Haley</td>
<td>Woman</td>
<td>Asian</td>
<td>15</td>
<td>12</td>
<td>12</td>
<td>80</td>
<td>100</td>
</tr>
<tr>
<td>David Ige</td>
<td>Man</td>
<td>Asian</td>
<td>12</td>
<td>9</td>
<td>9</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>Tim Scott</td>
<td>Man</td>
<td>Black</td>
<td>11</td>
<td>9</td>
<td>8</td>
<td>81.8</td>
<td>88.9</td>
</tr>
<tr>
<td>Cory A. Booker</td>
<td>Man</td>
<td>Black</td>
<td>10</td>
<td>10</td>
<td>9</td>
<td>100</td>
<td>90</td>
</tr>
<tr>
<td>Ben Carson</td>
<td>Man</td>
<td>Black</td>
<td>19</td>
<td>13</td>
<td>11</td>
<td>68.4</td>
<td>84.6</td>
</tr>
<tr>
<td>Marco Rubio</td>
<td>Man</td>
<td>Hispanic</td>
<td>11</td>
<td>6</td>
<td>6</td>
<td>54.5</td>
<td>100</td>
</tr>
<tr>
<td>Deb Fischer</td>
<td>Woman</td>
<td>White</td>
<td>11</td>
<td>10</td>
<td>10</td>
<td>90.9</td>
<td>100</td>
</tr>
<tr>
<td>Elizabeth Warren</td>
<td>Woman</td>
<td>White</td>
<td>14</td>
<td>10</td>
<td>9</td>
<td>71.4</td>
<td>90</td>
</tr>
<tr>
<td>Patty Murray</td>
<td>Woman</td>
<td>White</td>
<td>15</td>
<td>11</td>
<td>8</td>
<td>73.3</td>
<td>72.7</td>
</tr>
<tr>
<td>Jill Stein</td>
<td>Woman</td>
<td>White</td>
<td>15</td>
<td>14</td>
<td>13</td>
<td>93.3</td>
<td>92.9</td>
</tr>
<tr>
<td>Cindy Hyde-Smith</td>
<td>Woman</td>
<td>White</td>
<td>13</td>
<td>13</td>
<td>12</td>
<td>100</td>
<td>92.3</td>
</tr>
<tr>
<td>Ryan Zinke</td>
<td>Man</td>
<td>White</td>
<td>12</td>
<td>9</td>
<td>8</td>
<td>75</td>
<td>88.9</td>
</tr>
<tr>
<td>Sheldon Whitehouse</td>
<td>Man</td>
<td>White</td>
<td>14</td>
<td>13</td>
<td>13</td>
<td>92.9</td>
<td>100</td>
</tr>
<tr>
<td>Benjamin L. Cardin</td>
<td>Man</td>
<td>White</td>
<td>11</td>
<td>9</td>
<td>9</td>
<td>81.8</td>
<td>100</td>
</tr>
<tr>
<td>Richard J. Durbin</td>
<td>Man</td>
<td>White</td>
<td>11</td>
<td>7</td>
<td>6</td>
<td>63.6</td>
<td>85.7</td>
</tr>
<tr>
<td>Al Franken</td>
<td>Man</td>
<td>White</td>
<td>14</td>
<td>12</td>
<td>8</td>
<td>85.7</td>
<td>66.7</td>
</tr>
</tbody>
</table>

* There is no misidentification across all the photos examined (precision = 100%).
In addition, the computer vision services examined could not accurately identify facial expressions such as fear and sadness. As study 1 demonstrates, only happiness, anger, and neutral emotions achieved reasonable levels of accuracies. This is unfortunate, as different negative emotions often produce differing effects regarding the evaluation of valence and dominance (Knutson, 1996). Nevertheless, this low accuracy could partially be due to that some expressions were not prevalent in the dataset and some were highly ambiguous—raters themselves couldn’t agree upon which face should count as fearful (IRR = .50). This issue is less salient in the second study, as politicians on social media platforms predominantly use emotional expressions of happiness to render positive portrayals of themselves. One potential way to improve emotional detection is to harness the power of both existing computer vision libraries and supervised machine learning. For example, OpenFace is a library that provides the analysis of facial action units (FAUs), such as raising inner eyebrows, wrinkling nose, and jaw dropping (Baltrusaitis et al., 2018; Figure 6.4). FAUs are building blocks of facial expressions and different combinations of FAUs further signal emotional displays such as anger and disgust (Ekman & Rosenberg, 1997). To achieve better predictive performance on facial expressions of emotions, researchers might consider collecting a dataset of politicians’ facial expressions and then applying supervised machine learning on FAUs provided in OpenFace.
Visual features such as politicians’ gestures and activities might also reflect media bias (Moriarty & Garramone, 1986; Grabe & Bucy, 2009; Verser & Wicks, 2006) but could not be captured by the currently available computer vision services. In study 1, based on a crude examination of pictures that are perceived as most and least dominant, the use of hand gestures might be particularly important in shaping people’s perceptions of dominance. For example, images featuring politicians waving hands and pointing fingers are more likely to make them look more dominant than
images showing politicians in a static manner (Figure 6.5). This observation is consistent with some previous work: Joo et al. (2014) showed that participants’ perceptions of dominance were positively correlated with politicians’ hand gestures such as waving hands, shaking hands, and pointing fingers. Unfortunately, it was still difficult to use computer vision to accurately detect hand gestures in images: Joo et al. (2014) reported an average precision (AP) score of .351 for recognizing gestures.

Future directions

One promising direction is to look beyond still images of politicians and apply computer vision techniques to analyze moving images such as online videos. By January 2018, YouTube has become the most popular social media platform in the United States, with 73% of adults reporting using this site (Smith & Anderson, 2018). A significant portion (21%) of U.S. adults also consume news on YouTube, following Facebook (43%) and being followed by Twitter (12%) (Matsa & Shearer, 2018).
Previous scholarship has demonstrated that a few video-related visual characteristics, such as “lip flap” (a person mouthing words with no sound) and zooming in/out, can also reflect unfavorable or favorable media treatment (Grabe & Bucy, 2009). What’s more, audio analysis can also provide insights into the representations of politicians. Research has demonstrated that voters respond positively to political candidates with deeper voices (Klofstad, 2016). The analysis of multimedia in political communication should be a fruitful direction in this line of research.

Future research can combine textual analysis and visual analysis to advance the interrogation of media bias. In study 1, news stories associated with images should further influence how photos of politicians are picked up by media practitioners and interpreted by audience members. In study 2, the caption accompanying each post might further describe or explain what the image is about. Analyzing social media users’ comments made to these posts should also give us more insights into people’s interpretations of these images. We also see that a large number of Instagram images contain textual messages, which can be extracted with optical character recognition technique and further analyzed. Bringing tools in computational textual analysis, the next step in this line of research should be multi-modal analyses that investigate the interplay between textual and visual messages in media outlets or social media platforms.

Another direction is to apply visualization techniques in convolutional neural
networks (CNNs) to automatically detect visual regions that are linked to certain audience interpretations. Previously, scholars often adopt an approach that first proposes a list of visual features and then examines how these features are correlated with certain trait perceptions (e.g., Vernon et al., 2014). In study 1 and 2, I also pre-selected a list of visual features and investigated their associations with audience ratings or engagement. For future research, it is possible that to feed the data into a CNN and to use visualization techniques to recognize regions that activate people’s judgments (McCurrie et al., 2018; Zhou et al., 2016). For example, McCurrie et al. (2018) collected a dataset of over 6000 images of faces and applied CNNs to predict human judgments of four traits (i.e., trustworthiness, dominance, age, and IQ). The heat maps in Figure 6.6 visualize salient regions neural networks rely on to arrive at certain predictions. Joo and Steinert-Threlkeld (2018) also demonstrated that visualization techniques could be applied to diagnose crucial regions in protest-related images that lead CNNs to make certain classifications. Regarding politicians’ visual representations, future research can embrace this data-driven approach to test if participants attend to different visual cues, such as facial expressions, gestures, or background, to form impressions of politicians on different trait dimensions. In addition, we can also compare these findings from CNNs with results from eye tracking, a method that uses eye positions and movements to track how viewers allocate attention to different regions of visual stimuli (King et al., 2019;
Kruikemeier, Lecheler, & Boyer, 2018). In combination, these methods should help us better understand how visual attributes together shape viewers’ impressions of politicians.

Figure 6.6 Heat maps in a CNN that visualize important regions for predictions

![Heat maps](image)

*Note.* Image from McCurrie et al. (2018).

Aesthetical features might also play a role in influencing our perceptions of politicians and attracting audience engagement on social media. For instance, Enli (2017) analyzed Clinton and Trump’s campaign strategies on Twitter in the 2016 US presidential election and noticed a sharp distinction between a professional, traditional style in Clinton’s tweets and an amateurish but authentic style in Trump’s tweets. Trump’s posts were more likely to use an unconventional style of political campaigning, being characterized by frequent use of impolite and uncivil language, capital letters, and exclamation marks, thus rendering him as an “authentic outsider” that challenges established political norms. Will a similar pattern occur in the pictorial domain? Instead of publishing well-staged, professional-looking photographs with good composition and sharp focus, politicians might deliberately post more
amateurish-looking images marked by visual imperfections, such as blur, bad lighting, and poor composition, to cultivate an image of authenticity. The pursuit of amateurism and authenticity has also been observed in other communication contexts. For instance, Jenkins (2003) noticed that in amateur filmmaking, visual flaws such as “the abruptness in editing, the roughness of camera movement, the grittiness of film stock, and the unevenness of lighting” are regarded as indicators of authenticity (p.464). One future direction is to develop computer vision algorithms that capture amateurish-looking aesthetics in politicians’ photographs and investigate how they are used and how they can influence audience reactions. This direction can also help us better understand the theoretical links between personalization and authenticity.

We should also acknowledge that this dissertation is based on observational studies and correlational analyses. In study 1, I looked at whether the gap between two presidential candidates varied by media outlets of different ideological positions to isolate the presence of partisan bias. However, it is also possible that gender bias is a component of partisan bias, as liberal and conservative media outlets might adopt different representations of women politicians. Therefore, the upcoming 2020 presidential election provides a valuable opportunity for future research to examine how partisan and gender biases interplay in media coverage of politicians and extend the conclusions of this study, as more women are participating in the Democratic primaries. What’s more, certain visual representations might be correlated with each
other. Future scholars can use experiments to study how certain visual attributes influence viewers’ impressions and attitudes regarding politicians.

Last, not only does computer vision provide social scientists with advanced tools to analyze massive visual data, but it itself is also an important subject of social scientific inquiry. Computer vision has already been applied in a variety of settings that directly relate to the public interest, such as surveillance and automatic driving. The emergence of “deepfake,” a technique that generates images and videos based on existing materials using a machine learning technique (generative adversarial network), could also be used in political communication to spread misinformation and mislead public opinions (Gilmer, 2019; Klaas, 2019). The influence of computer vision technologies on our society should also be one promising direction for future research.
References


Boston, MA: IEEE.


Fridkin, K., & Kenney, P. (2014). *The changing face of representation: The gender of*


Klaas, B. (May 14, 2019). Deepfakes are coming. We’re not ready. Retrieved from https://www.washingtonpost.com/opinions/2019/05/14/deepfakes-are-coming-were-not-ready


Konrath, S., Au, J., & Ramsey, L. R. (2012). Cultural differences in face-ism: Male politicians have bigger heads in more gender-equal cultures. *Psychology of
Women Quarterly, 36(4), 476–487.


Larsson, A. O. (2017). Skiing all the way to the polls: Exploring the popularity of personalized posts on political Instagram accounts. Convergence, Online advanced publication.


Meeks, L. (2012). Is she “man enough”? Women candidates, executive political


Microsoft (u. d.). API reference. Retrieved from https://westus.dev.cognitive.microsoft.com/docs/services/563879b61984550e40bbbe8d/operations/563879b61984550f30395236

Microsoft (u. d.). Emotion API. Retrieved from https://dev.projectoxford.ai/docs/services/5639d931ca73072154c1ce89/operatio


Mutz, D. C. (2007). Effects of “in-your-face” television discourse on perceptions of a


clearer picture: The contribution of visuals and text to framing effects. *Journal of Communication, 65*(6), 997–1017.

Prentice, D. A., & Carranza, E. (2002). What women and men should be, shouldn't be, are allowed to be, and don't have to be: The contents of prescriptive gender stereotypes. *Psychology of Women Quarterly, 26*(4), 269–281.


cues to identity in CMC. *Computers in Human Behavior, 23*(2), 955–970.


Vernon, R. J., Sutherland, C. A., Young, A. W., & Hartley, T. (2014). Modeling first


