ARTICLE

Introduction to the Special Issue on Images as Data

Andreu Casas
Department of Communication Science, Vrije Universiteit Amsterdam

Nora Webb Williams
Department of Political Science, University of Illinois at Urbana-Champaign

Why a Special Issue on Images as Data?

Visual information (primarily still images and videos) is crucial for the study of many current communications, political, and social phenomena. Yet research leveraging large corpora of visuals to answer social science questions is still scarce, especially relative to the explosion of research using "big data" text-as-data methods. This special issue fosters innovative theoretical and methodological research in the area of images as data. The featured articles use computational methods to analyze, from a social science perspective, large quantities of images as well as videos. Three overarching motivations fueled the creation of this special issue on "Images as Data" at *Computational Communication Research.*

    First, visuals are becoming a more frequent form of communication, especially online. Social media is a good reflection of this new communication paradigm. Not only have social media platforms become more central to people's lives (more people, and more often, use social media to interact with each other and to learn about what is going on in the World), but the platforms have become more visual. In the last few years, we note an increase in visuals on messages posted on the more "traditional" social media platforms such as Twitter, Facebook and Weibo. And "younger" platforms such as Instagram, Tik Tok, and Douyin are even more specifically designed around visual communication. Other communication platforms have followed this visual trend. For example, online news outlets have also increased the amount of visuals they attach to news stories. Posting images or videos online comes at a tiny fraction of the cost of printing images in a newspaper; and in return, images and videos often help news media get people's attention,

attract readership, and frame the news story. In sum, images and videos are a central part of many phenomena of interest to social scientists today (e.g. news values, framing, agenda-setting, dis/mis-information, mobilization, etc.), and we should do more to include visuals in our analyses.

Second, a growing amount of visual data is available to social scientists. As communications become more visual, researchers interested in studying these phenomena have also had increased access to images and videos for analysis. Repositories of visual data from public communications are gradually flourishing. The TV News Archive[1] from the Internet Archive (mostly focused on news from English-speaking countries; see Dietrich and Ko (2022) in this issue) and the media collection of the Institute for Sound and Vision in the Netherlands are good examples, as is the Wesleyan Media Project, which collects political advertising in the US (see Neumann et al. (2022) in this issue). Visuals can also fill the research needs of scholars who are not necessarily interested in studying communications. As some examples, open-source satellite images are on the rise, which can provide measures of economic development in areas for which little data is available;[2] and many art institutions are digitizing artwork and creating open data sets of historical imagery that can, for example, facilitate the study gender roles across time and places.[3] In sum, the growing volume of visual data can help social scientists answer both new and long-standing questions in the literature.

Third, new computational methods are available for automated image analysis. Fueled by the deep learning revolution, the last few years have seen major advancements in "computer vision," the subfield of computer science (and related disciplines) concerned with automatically analyzing images. There have been key improvements in crucial research tasks, such as object recognition, image classification, face recognition, facial traits analysis, etc (see Steinert-Threlkeld and Joo (2022), in this issue, for an introduction to these tasks). These new methods allow social scientists to automatically analyze large quantities of images. With new tools we can more efficiently study how different groups or news organizations frame particular issues; how people process and learn particular information; the presence of representation biases in public communications; and the correlates of economic development, among many other topics. However, in order to take full advantage of these methods, social scientists first need to learn about how these methods work and for what research purposes, and under what conditions, they can be useful.

We believe that this special issue (a) highlights excellent work that addresses the three issues described above; (b) will inspire future work

with images as data; and (c) helps scholars overcome some of the current limitations emphasized above. The images-as-data field in the social sciences is still in its infancy. Currently there are some very strong pieces providing overviews of which computer vision methods can be of use to social scientists (e.g. Torres and Cantú, 2021; Webb Williams et al., 2020), and a few (although of high quality) research articles using computer vision techniques to answer relevant theoretical questions (Dietrich 2021; Boussalis et al., 2021; Peng, 2018, 2020; Won et al., 2017; Xi et al., 2020; Zhang and Pan, 2019). However, we think that **more research using computer vision methods in the social sciences is needed in order to encourage further work in this area**. This special issue serves as a home for innovative research papers that serve as examples of the many data and methods available, and the wide range of theoretical questions one can answer with images as data. We hope this inspires many others to take this research route in the years to come.

Moreover, as with any research (and computational) method, there are starting costs to learning how they work before one can implement them in a research project. This is particularly the case with computer vision methods, given that the jargon is often very computer-science and machine-learning specific, and state-of-the-art libraries and packages are available mostly in Python while social scientists are often more used to working in the R programming language. Thus **the special issue aims to lower start-up costs for scholars interested in using images-as-data methods in their research**. We include workshop articles in which the authors clearly walk readers through novel computer vision methods that will be of use to many social scientists. In addition, we asked the authors of both workshop and research papers to provide open source code to not only replicate the findings reported in the papers, but also to demonstrate how to implement the computer vision tool/s used in the papers. We strongly believe that the combination of illustrative and inspiring research articles, workshop articles detailing particular methods, and open source code that shows, step by step, how to implement a wide range of computer vision methods, will advance research on images-as-data in the social sciences in the future.

In the Special Issue

As mentioned above, we solicited two types of articles for the special issue. First are traditional "research" articles, where authors are advancing knowledge about specific phenomena. Second are "workshop" articles, where authors are introducing new tools or methods for image analysis.

Our hope was that the workshop format would give a home to work that can be difficult to place in other publication outlets. It reflects our belief that describing and explaining new methods represents a significant scholarly contribution. Although all contributors submitted replication material and provided open-source data and code along with their papers, sharing raw data for some of the contributions turned out to be rather complicated for privacy or copyright issues. Nevertheless, even in such cases, the authors have produced detailed instructions and sample analyses to demonstrate their pipelines. These replication materials will be a fruitful resource for future researchers.

We received a total of 31 extended abstracts in response to our call for submissions. Of those, 19 were in the research track and the remaining 12 were for workshop articles. We requested full papers from 18 of the abstracts, all of which were then subject to blind peer-review. In the end we accepted a total of 10 articles. Of the accepted articles, 7 are designated as research and 3 are workshop. Decisions about which abstracts to pursue, which full papers to send for review, and which papers to accept were very difficult. The quantity of important and interesting work in this field is growing at a rapid pace; it is an exciting time in the images-as-data field and we sincerely hope that every paper we rejected is eventually published.

In the remainder of this section we discuss common threads across the accepted papers. We begin with thematic threads before turning to methodological commonalities. Finally, we address common challenges faced by the authors.

Thematic similarities
There are three general topic areas that emerged in the special issue papers, despite not specifying topical themes in advance: representation, persuasion, and explaining new methods. Many of the papers touch on one or more of these themes.

The first theme, representation, encompasses self-representation as well as the representation of individuals by others. To a lesser extent, and not always explicitly, these pieces are engaged with questions of framing. Neumann et al. (2022) analyze political ads in the United States, for example, to examine how politicians of different genders use different gestures – they find that men tend to use more assertive body language. Jürgens et al. (2022) also address differences in gender representation, analyzing 16 million faces from 6 years of German television and finding significant differences in age and gender representation between men and women. Staying in the sphere of popular media but examining racial dynamics instead of gender, Malik et al.

(2022) analyze popular movies to see how often different races/ethnicities are represented. Dietrich and Ko (2022) study how often Dr. Anthony Fauci appears in media coverage of the COVID-19 pandemic, addressing how different media organizations make different choices about whose voices to include in news coverage.

Some of the papers mentioned above also touch on the theme of persuasion, as the images are intended to have some effect on the audience (e.g. the ads in Neumann et al. (2022) are supposed to persuade voters). Yet a couple papers are more specifically focused on this theme. Lu and Pan (2022), for example, analyze trends on the Douyin platform in China (similar to TikTok) to see what tactics Chinese regime accounts use to boost popularity. Chen et al. (2022) study conspiracy videos posted on YouTube to address commonalities between these videos, as well as differences between clips sponsoring versus debunking such conspiracies (this piece naturally also touches on the representation/framing theme).

The final theme encompasses articles that have a primary focus on explaining new methods and tools. As mentioned in the section below, all of the special issue articles are using new and innovative methods. This subset of articles, however, are more explicitly intended to describe new tools. This topic includes the three workshop articles in the special issue, as well as one research article. The articles all include substantive applications, some of which relate to the themes of representation and persuasion. Peng (2022), for example, provides an introduction to a Python package that extracts common aesthetic features from images. The application in the article is to Instagram representations of US politicians Hilary Clinton and Donald Trump during the 2016 election. Steinert-Threlkeld and Joo (2022) provide a wide-ranging introduction to the basics of automated content analysis (if you are new to these methods, we recommend this as a starter piece). Baldwin and Schmälzle (2022) introduce a new tool for tracking various "characters" across images while Wu and Mebane (2022) describes an innovative method for combining text and image data for supervised classification.

A wide range of new methods

Given the nature of the special issue call, it is no surprise that all of the articles are using computer vision tools to analyze large quantities of image data (and we do mean *large* quantities of data – some of the observation counts and file sizes are astounding). As you read the articles, you will encounter a wide range of helpful packages (mainly in Python), as well as a sense of the huge lifts the authors have performed to collect, clean, and analyze their data.

What may be surprising to readers, however, is the amount of video content analyzed in the special issue. It is clear from the articles that scholars are hungry to move beyond static images to better understand video content; so much of modern communication occurs though moving pictures. That said, most of the analyses of video in the special issue use a strategy of sampling frames from videos for analysis. This strategy is powerful and allows the authors to dig into important and interesting questions. However, as we discuss in more detail below, this also is an area where the subfield can expand.

Readers should also, we think, be impressed by the merging of image and text content in these articles. Many of the authors wrestle with (or leverage) the combined task of understanding images and text. We were very pleased to see these analyses, as so often texts and images are analyzed in isolation. Using both comes with many challenges but also a very large upside, as the authors demonstrate.

Common methodological challenges

In addition to all using new tools, the articles all also face methodological challenges that are common for computational social science research. Here we highlight three: validation, replication, and rapidly changing technologies.

Validation is the bane of big data work. How can we be sure that the concepts we see in images are what others would see? How much do we trust our image annotators? How much do we trust algorithms that may have been trained on biased data? All of the authors in the special issue have had to address questions along these lines. The need to provide evidence of careful validation is clear for work in this area.

Replication in some ways goes hand in hand with validation. If an image classification schema is valid, another scholar should be able to replicate that classification. One major challenge to replication faced by the article authors is the ability (or lack thereof) to share raw data for privacy or copyright reasons. This can be frustrating to see after noting the heroic efforts of the authors to collect and organize their data – it is a challenge to not be able to easily share the raw images, videos, and texts for others to use for replication and extension.

A final challenge arises due to the fast-paced nature of changes in the field. Not only are new computer vision algorithms popping up like mushrooms, but access to data and tools are also changing frequently. For example, a new way to access Twitter data means that code that previously ran beautifully now stalls before collecting any data. Or a package update might change a tool's functionality. A solution to this is to rigorously document which version

of tools scholars are using. Another possible suggestion is an online forum for rapidly publishing information on new tools that can easily be updated. Keeping up with the pace of change in the field may require new thinking about publishing models (even beyond the workshop article option). As editors we appreciated the CCR model of allowing authors to post working papers as early as possible in the review/acceptance stages – this is a great step to making sure that peer-reviewed publishing accelerates to meet the rate of technological changes.

What Comes Next for the Subfield?

As the papers in this issue demonstrate, the sky is the limit for fascinating research using images as data. We can already track and describe patterns of behavior on social media; better explain political campaigning techniques; analyze trends in popular media; and understand propaganda efforts. We can only begin to imagine all of the new research that will emerge in the next few years with the advancement of automated techniques for image analysis. There are many areas that are ripe for future research with images as data – we highlight a select few here.

First is remote-sensing research, or using picture taken of earth at a distance, usually by satellites. Images in this sphere can help us track land use patterns, climate change, and even electrification (by looking at nighttime lights images). We have the computational tools to process these images; it is up to social scientists to creatively use remote-sensing data to answer pressing research questions.

Second, this special issue makes clear that the future of images-as-data research is multi-modal. Many of the authors are already explicitly working in a multi-modal space, combining images with text, or video with text. These are important advances as there are huge gains to be won by not treating images in isolation, but instead pairing them with text, audio, and motion. In particular, many of the authors working with video data acknowledge that the strategy of extracting and analyzing single still frames is useful, but that there is much more than can be done. Methods that leverage motion and change between frames (as Neumann et al. (2022) do, or see Dietrich 2021) are on the cutting edge of helping us to better understand images as they appear in our daily lives. Video analysis is here and the tools are ready for our explorations.

Third, we are particularly pleased that the research in this SI covers multiple countries, including the US, China, and Germany. Yet we note that this is an area where there is still much work to be done, as much of the

current work focuses on the US and highly developed countries. The challenges and opportunities of working with images from a variety of cultural and political background cannot be overstated. Some aspects of visual communication may be universal, but others are not. Our understanding of social phenomena from a range of cultures will be enhanced by analyses of images from many different places and times.

Finally, there is still work to do to make the innovations in methods accessible to a wide range of scholars. Our colleagues in computer science, data science, statistics and informatics continue to develop new and exciting tools and techniques for unsupervised, supervised, and semi-supervised pipelines. Social scientist methodologists also contribute to these developments. A major task for the images-as-data subfield is to teach others how to use these new techniques accurately, efficiently, and ethically. We can accomplish this by publishing papers describing new techniques but also by actively training new scholars with specialized workshops and seminars.

Conclusion and Thanks

The wide range of articles in this special issue should serve as inspiration for new and exciting research in the field of images as data. The articles in the issue are excellent examples of what we can learn by using images as data – and there is so much more yet to be discovered.

We are grateful to the article authors for submitting their work to this special issue. We are also very thankful for our pool of anonymous reviewers, who provided crucial feedback and helped strengthen the submissions. Finally, we thank the editorial staff of the journal for their constant support of the special issue. The COVID-19 pandemic and other whims of the universe led to many missed deadlines and a significantly-revised publication schedule for this issue – we are grateful for the patience and grace that all participants extended throughout this process.

Notes

1. https://archive.org/details/tvarchive
2. See this repository for some examples: https://github.com/chrieke/awesome-satellite-imagery-datasets
3. See here for some examples: https://pro.europeana.eu/pages/datasets/data/itemtype/paintings

References

Baldwin, J., & Schmälzle, R. (2022). A Character Recognition Tool for Automatic Detection of Social Characters in Visual Media Content. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.010.BAL

Boussalis, C., Coan, T. G., Holman, M. R., & Muller, S. (2021). Gender, Candidate Emotional Expression, and Voter Reactions During Televised Debates. *American Political Science Review, First View*, 1–16. https://doi.org/10.1017/S0003055421000666

Chen, K., Kim, S. J., Raschka, S., & Gao, Q. (2022). Visual Framing of Science Conspiracy Videos. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.003.CHEN

Dietrich, B. J. (2021) Using Motion Detection to Measure Social Polarization in the U.S. House of Representatives. *Political Analysis*, 29(2), 250–259. https://doi.org/10.1017/pan.2020.25

Dietrich, B. J., & Ko, H. (2022). Finding Fauci: How Visual and Textual Information Varied on Cable News Networks During the Covid-19 Pandemic. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.004.DIET

Jürgens, P., Meltzer, C., & Scharkow, M. (2022). Age and Gender Representation on German TV: A Longitudinal Computational Analysis. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.005.JURG

Lu, Y., & Pan, J. (2022). The Pervasive Presence of Chinese Government Content on Douyin Trending Videos. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.2.002.LU

Malik, M. I., Hopp, F. R., & Weber, R. (2022). Representations of Racial Minorities in Popular Movies: A Content-Analytic Synergy of Computer Vision and Network Science. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.006.MALI

Neumann, M., Franklin Fowler, E., & Ridout, T. N. (2022). Body Language and Gender Stereotypes in Campaign Video. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.007.NEUM

Peng, Y. (2018). Same Candidates, Different Faces: Uncovering Media Bias in Visual Portrayals of Presidential Candidates with Computer Vision. *Journal of Communication*, 68(5), 920–941. https://doi.org/10.1093/joc/jqy041

Peng, Y. (2020). What Makes Politicians' Instagram Posts Popular? Analyzing Social Media Strategies of Candidates and Office Holders with Computer Vision. *The International Journal of Press/Politics*, 26(1), 143–166. https://doi.org/10.1177/1940161220964769

Peng, Y. (2022). Athec: A Python Library for Computational Aesthetic Analysis of Visual Media in Social Science Research. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.009.PENG

Steinert-Threlkeld, Z. C., & Joo, J. (2022). Image as Data: Automated Content Analysis for Visual Presentations of Political Actors and Events. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.001.JOO

Torres, M., & Cantú, F. (2021). Learning to See: Convolutional Neural Networks for the Analysis of Social Science Data. *Political Analysis, First View*, 1–19. https://doi.org/10.1017/pan.2021.9

Webb Williams, N., Casas, A., & Wilkerson, J. D. (2020). *Images as Data for Social Science Research: An Introduction to Convolutional Neural Nets for Image Classification*. Cambridge University Press. https://doi.org/10.1017/9781108860741

Won, D., Steinert-Threlkeld, Z. C., & Joo, J. (2017). Protest Activity Detection and Perceived Violence Estimation from Social Media Images. *Proceedings of the 25th ACM International Conference on Multimedia*. https://doi.org/10.48550/arXiv.1709.06204

Wu, P. Y., & Mebane, W. R. J. (2022). MARMOT: A Deep Learning Framework for Constructing Multimodal Representations for Vision-and-Language Tasks. *Computational Communication Research*. https://doi.org/10.5117/CCR2022.1.008.WU

Xi, N., Ma, D., Liou, M., Steinert-Threlkeld, Z. C., Anastasopoulos, J., & Joo, J. (2020). Understanding the Political Ideology of Legislators from Social Media Images. *Proceedings of the International AAAI Conference on Web and Social Media*. https://doi.org/10.48550/arXiv.1907.09594

Zhang, H., & Pan, J. (2019). CASM: A Deep Learning Approach for Identifying Collective Action Events with Text and Image Data from Social Media. *Sociological Methodology*, 49(1), 1–57. https://doi.org/10.1177/0081175019860244