DHQ: Digital Humanities Quarterly
2023
Volume 17 Number 2

Distant Reading and Viewing: “Big Questions” in Digital Art History and Digital Literary Studies

Ruta Binkyte <ruta_dot_binkyte_at_gmail_dot_com>, Inria Saclay - Île-de-France Research Centre, Institut Polytechnique de Paris https://orcid.org/0000-0001-8781-3730

Introduction

While digital literary studies are well established, digital art history is only taking its first baby steps. As distant reading has contributed to revolutionizing traditional literary studies, distant viewing is lagging behind. Machines are taking their first looks at visual artworks in the research projects that, at this stage, are confusingly bordering domains of computer science and art history. This essay attempts to better demarcate these borders by looking at possible questions that are both art-historical in nature and impossible or hard to answer without computational techniques. For this task, we compare the distant reading with distant viewing and explore the experience of the “older sister” discipline of digital literary studies. 1

Related Work

Although articles dedicated to defining digital art history acknowledge digital literary studies as its predecessor in digital humanities, the systematic research comparing the two disciplines is minimal. [Bishop 2018] points out parallels between the disciplines by providing critical arguments against digital art history. [Arnold and Tilton 2019] analyze similarities and differences between natural language processing and computer vision in their essay on “distant viewing.” However, the authors focus on methodological aspects of machine vision rather than epistemology and research questions. 2

This essay primarily relies on the articles on digital art history ([Rodríguez-Ortega 2020], [Drucker et al. 2015] [Bender 2015]), critique of the quantitative approach in art history ([Pollock 2014]), and theoretical and methodological inquiries on digital literary studies ([Stefan et al. 2015] [Moretti 2000] [Coste 2018]). We review the literature to identify similarities in goals and challenges in digital art history and digital literary studies and how the experience of the latter can help to define research questions most suitable for using machine-aided vision in art history. 3

Digital Art History and Digital Literary studies

Digital art history and digital literary studies are the disciplines under the umbrella term of digital humanities – an area broadly described as an intersection of humanities disciplines and computing or digital technologies [Drucker 2021] [Burdick et al. 2016]. The inception of this interdisciplinary field is often linked to the work of Roberto Busa on computer-generated concordance of the works of Thomas Aquinas in 1940 and has long time been synonymous with the study of text [Schneider 2012] [Dalbello 2011]. Therefore, it comes as no surprise that digital literary studies have already earned their status through acknowledged contributions to the more general domain of digital humanities. In contrast, digital art history is still trying to prove itself [Joyeux-Prunel 2008]. Digital scholars in digital humanities like Stephen Ramsay, Franco Moretti, Matthew Jockers, Geoffrey Rockwell, and others have established the use of digital tools as part of the interpretative process and means of approaching literary works in novel ways through the methodology of distant reading [Earhart 2015]. Digital art history, however, relied on the processing of textual meta-data and, until recently, did not experience the deep transformation of the relationship between object and subject in its research. Although computational techniques of signal processing aiding art historians in authorship attribution and dating tasks have been around for more than four decades [Stork 2009], only recent advances in computer vision have brought a possibility of research parallel to distant reading in digital art history. The key milestones for computer vision were the introduction of Convolutional Neural Network algorithms and the creation of the ImageNet image database in 2012, which enabled the 4

training of deep learning algorithms for object recognition. They were later generalized for other tasks, including style recognition in artistic images [Bar et al. 2014]. Possibilities to challenge human sight when looking at art images have attracted computer scientists and art historians alike [Ginosar et al. 2015] [Lang and Ommer 2018] [Manovich 2012] [Schich et al. 2008] [Schich 2015].

Distant Reading and Distant Viewing

5 The idea of distant reading is linked to the emergence of open data initiatives and digitization of texts, for example, Google Books, Internet Archive PHI, and Perseus Digital Library [Stefan et al. 2015]. Distant reading in the field of literature was introduced by Franco Moretti [Moretti 2000] and Matthew Jockers's "macroanalysis" [Jockers 2013] as a possible answer to the challenges of massive text corpora. While Moretti mostly focuses on visualization techniques (like Manovich in digital art history), Jocker draws analogy from macroeconomics and emphasizes statistical analysis methods [Stefan et al. 2015]. The analog term in visual culture - *distant viewing* - already appeared in 2015, although it referred to statistics of artworks metadata, not the computer-vision analysis of images [Bender 2015]. A new formulation of distant viewing methodology already relies on machine-aided vision [Arnold and Tilton 2019]. The idea of analyzing artistic images at a scale that came before the articulation of the concepts is enabled by digitization and opening of visual image archives - Getty Research Institute, Google Art Project, and others [Drucker et al. 2015]. Some interesting studies on social media image content or classical paintings have entered the art history scene [Manovich 2012] [Schich et al. 2008] [Schich 2015].

6 Similarly to distant reading, distant viewing includes statistical techniques and visualization and enables a macro view of big image data [Arnold and Tilton 2019]. There are, however, important differences. One of them is the semantic gap between raw pixels of an image and structured features extracted by a deep learning algorithm [Arnold and Tilton 2019]. Other – inherent materiality of most of the classical objects of art history and digitization error [Drucker et al. 2015]. Those two factors make the interpretation and role of the researcher even more implicit in the methodology of distant viewing. Both distant reading and distant viewing enable data-driven discovery; however, this essay attempts to clarify questions that this approach can anticipate.

Statistics versus Interpretation

7 As the very term “Distant Viewing” is inspired by its predecessor “Distant Reading”, the criticism towards this new approach is similar to that of new literary study methodologies. The reoccurring motif in the criticism is juxtaposing the quantitative and qualitative approach with binaries of counting or reading, statistics, or critical thinking. Still, as Mark Algee-Hewitt points out, the goal of computational humanities is to find a balance between them [Algee-Hewitt 2019]. According to Coste, the quantitative approach to literature has qualitative implications; the scope achievable with distant reading transforms the relationship with the unit observed, making it comparable with the big literary data corpora [Coste 2018]. On the one hand, we have a reality conditioned by our senses (close reading), and on the other – by the instruments (distant reading). However, it is the researcher that creates knowledge through the act of interpretation, not the machine. The same is true about the field of digital art history. The idea that statistical methods do not involve the interpretation or critical thinking is a misconception rooted in the popular imagination, with labels such as “dry statistics” instead of intuition and insight. Every statistical model is rooted in the qualitative assumptions of the world. For example, the mathematical model of image and style has an implicit concept of an image and modes of vision [Rodríguez-Ortega 2020]. The other extreme, pointed out by Claire Bishop in the critical article on digital art history, is overly enthusiastic reliance on empiricism or data as bearing objective truth and abandoning interpretation and search for causality [Bishop 2018]. Here we explore how art history can be enriched with approaches that consciously apply computational methods together with theory and interpretation.

Flat comparisons versus the big questions

8 The first successful attempt to use algorithms to identify artistic influences in paintings made the headlines promising computers will push art historians out of business [Sparkes 2014] [Emerging Technology 2015]. Computer scientists-led project explored a database of more than 80,000 paintings by more than 1,000 artists created over 15 centuries. The

authors have built models for attributing artistic styles and spotting influence links among the artists from algorithmically extracted visual features in the artworks. Apart from being able to match well-acknowledged links between styles and artists (for example, expressionism and fauvism), the algorithms also showed a never-before-seen visual resemblance between Frederic Bazille's Studio 9 Rue de la Condamine (1870) and Norman Rockwell's Shuffleton's Barber Shop (1950) [Saleh et al. 2016]. It did not take long for the community of art historians to react. The findings were dismissed as representing an outdated "connoisseurial art history" which ignores social, ideological, economic, and political components to form "larger narratives" [Pollock 2014]. As evident from the authors' response to criticism [Elgammal 2014], here we see the case where art history is used to test and develop algorithms rather than to answer big questions about art history. But can it work the other way around? Can distant viewing enable the eyes of an art historian to see differently and answer or raise new questions through the process of interpretation? As Matthew Lincoln points out, good research starts with art-historical rather than computer science questions [Drucker 2013]. But are all art-historical questions equally answerable with massive data and computer algorithms? In the following paragraphs, we will explore the cases from literary studies and digital art history to identify questions that can bring qualitative value to the discipline and transform the relationship with analyzed artworks.

Challenging the Canon

Pollock calls computational “influence tracing” a disguised way of protecting the canon [Pollock 2014]. The ongoing project “Images' Contagions” aims to prove quite the opposite. The Artl@s research group of art historians, engineers, and cognitive scientists approach the massive corpora of digitized images and related textual data with Deep Learning algorithms for pattern and object recognition. In addition, qualitative methods from history, visual studies, and cognitive sciences are applied [Joyeux-Prunel, B. 2012]. The project's objectives are to identify the most recurrent images and circulations of their copies and visual quotations. Iconographic and stylistic similarities and influences are analyzed by applying a model of epidemiology of images inspired by biology and based on mathematical network analysis. According to the project's authors, this approach, together with the available big image data, is the opportunity to challenge the “old geopolitical model of prescriptive centers and imitative peripheries” [Joyeux-Prunel, B. 2012]. The macro level of distant viewing can help to challenge the ideological defense of canon better than traditional art history which favors monographic studies [Joyeux-Prunel 2008] According to Franco Moretti, close reading limits one to the “canonical fraction” of Western literature, let alone world literature [Moretti 2013]. 9

The broader horizons provide a different relationship with the artwork but are unreachable to the human eye. It is even more true if one adds user-generated content on social media as a contemporary art form [Manovich 2012]. Therefore machine-aided vision comes as a tool that opens important possibilities, even if it does not allow the intimacy of close reading or close viewing. As Bender points out, studying the spread of visual quotations requires a quantitative approach [Bender 2015]. 10

Counting and Counter-intuitive insights

Reading or viewing at a large scale is an unachievable task for humans not equipped with algorithmic tools. Another one is precise quantification. The application of algorithmic graph analysis and network visualization technique to classic plays by Shakespeare and Sophocles has revealed that main characters do not necessarily have the central role in the character network of the play, for example, Cesar in Shakespeare's *Julius Caesar* [Coste 2018] [Grandjean 2015]. One may argue that simple counting does not change the role of the main character. However, it can encourage a more detailed look at characters' space and how literary means are used to create centrality. Arnold and Tilton applied a similar approach to analyzing visual culture. The authors use a deep learning face recognition algorithm to study the appearance of main characters in situational comedies *Bewitched* (1964–1972) and *I Dream of Jeannie* (1965–1970). The authors start by locating the faces in each frame, then identify the main characters and cluster them based on camera angles and distances [Arnold, Tilton, and Burke 2019]. The study aimed to inquire about the gender roles promoted by television. It has challenged the intuitive perception of both Samantha in *Bewitched* and Jeannie in *I Dream of Jeannie* as leading characters by showing that Jeannie appears on the screen much less than Samantha. 11

Furthermore, Jeannie is often visually framed as a supporting rather than the main character [Arnold, Tilton, and Burke 12

2019]. This revelation challenges dominating views that two TV series represent the same female roles [Arnold, Tilton, and Burke 2019]. The study is an example of how digital methods allow massive data analysis that would be tedious work, if at all possible, without computational aid. However, the central insight of the study would not be possible without an art historian's intuition and knowledge of visual tools for building character in TV drama.

Conclusions

Traditional art historians often critique distant viewing, and distant reading is controversial in literary studies. However, both methods provide possibilities unachievable with conventional methods. Recent examples of computer algorithms in art history suggest that this methodology allows the transformation of the relationship with the art object rather than mere counting. The analog research cases from literary studies strengthen this claim. While traditional methods have their place in the discipline, the digital approach is superior in working with the massive scope of data in global or longitudinal studies. It has the advantage of questioning traditional beliefs and well-established canons with quantitative analysis. Therefore, computational tools can help to reposition standard objects in their historical or ideological contexts. Computational tools become even more relevant when new objects – internet and social media content – begin to capture the attention of researchers in art history. However, the future challenges include establishing rigorous interdisciplinary frameworks and evaluation criteria together with openly available high-quality datasets. In addition, developments in institutional support, availability of multidisciplinary training, and technical infrastructure are important in developing distant viewing in digital art history research.

Acknowledgements

The author is thankful to Dr James Stewart who provided useful discussion and feedback when writing this article.

Works Cited

Algee-Hewitt 2019 Algee-Hewitt, M. “Criticism, Augmented”. *Computational Literary Studies: A Critical Inquiry Online Forum*, 2019. https://critinq.wordpress.com/2019/04/01/computational-literary-studies-participant-forum-responses/

Arnold and Tilton 2019 Arnold, T., and Tilton, L. “Distant viewing: analyzing large visual corpora”. *Digital Scholarship in the Humanities*, Volume 34, Issue Supplement_1, 2019, pp. i3–i16, https://doi.org/10.1093/llc/fqz013

Arnold, Tilton, and Burke 2019 Arnold, T., Tilton, L., and Berke, A. “Visual style in two network era sitcoms”. *Journal of Cultural Analytics*, Volume 4, Issue 2, 2019, 11045. https://doi.org/10.22148/16.043

Bar et al. 2014 Bar, Y., Levy, N., and Wolf, L. “Classification of artistic styles using binarized features derived from a deep neural network”. *European conference on computer vision*, Springer, 2014, pp. 71-84.

Bender 2015 Bender, K. “Distant viewing in art history. a case study of artistic productivity”. *International Journal for Digital Art History*, Issue 1, 2015. doi:10.11588/dah.2015.1.21639

Bishop 2018 Bishop, C. “Against digital art history”. *International Journal for Digital Art History*, Issue 3, 2018.

Burdick et al. 2016 Burdick, A., Drucker, J., Lunenfeld, P., Presner, T., and Schnapp, J. (2016). *Digital_Humanities*, MIT Press, 2016.

Coste 2018 Coste, F. “La littérature en numérique”. *La vie des idées*, 2017. https://hal.science/halshs-01590655/

Dalbello 2011 Dalbello, M. “A genealogy of digital humanities”. *Journal of Documentation*, Volume 6, 2011, pp. 480-506.

Drucker 2013 Drucker, J. “Is there a “digital” art history?”. *Visual Resources*, Volume 29, Issue 1-2, 2013, pp 5-13.

Drucker 2021 Drucker, J. *The Digital Humanities Coursebook: An Introduction to Digital Methods for Research and Scholarship*, Routledge, 2021.

Drucker et al. 2015 Drucker, J., Helmreich, A., Lincoln, M., and Rose, F. “Digital art history: the American scene”. *Perspective [En ligne]*, Volume 2, 2015. https://doi.org/10.4000/perspective.6021

Earhart 2015 Earhart, A. *Traces of the old, uses of the new: The emergence of digital literary studies*, University of Michigan Press, 2015.

Elgammal 2014 Elgammal, A. “Computer science can only help — not hurt — art historians”, *The Conversation*, 4 December 2014.

Emerging Technology 2015 Emerging Technology “The Machine Vision Algorithm Beating Art Historians at Their Own Game”. *MIT Technology Review*, 2015. https://www.technologyreview.com/s/537366/the-machine-vision-algorithm-beating-arthistorians-at-their-own-game/

Ginosar et al. 2015 Ginosar, S., Haas, D., Brown, T., and Malik, J. “Detecting people in cubist art”. *AI Matters*, Volume 1, Issue 3, 2015, pp. 16-18.

Grandjean 2015 Grandjean, M. “Network visualization: mapping Shakespeare's tragedies”. *The Routledge Companion to Digital Humanities and Art History*, Routledge, 2015, pp. 338-357.

Jockers 2013 Jockers, M. L. *Macroanalysis: Digital methods and literary history*, University of Illinois Press, 2013.

Joyeux-Prunel 2008 Joyeux-Prunel, B. “L’histoire de l’art et le quantitatif. Une querelle dépassée”. *Histoire and mesure*, Volume 23 Issue XXIII-2, 2008, pp. 3-34. =

Joyeux-Prunel, B. 2012 Joyeux-Prunel, B. “ARTL@ S: A spatial and trans-national art history Origins and positions of a research program”. *Artl@ s Bulletin*, Volume 1, Issue 1, 2012

Lang and Ommer 2018 Lang, S., and Ommer, B. “Attesting similarity: Supporting the organization and study of art image collections with computer vision”. *Digital Scholarship in the Humanities*, Volume 33, Issue 4, 2018, pp. 845-856.

Manovich 2012 Manovich, L. “How to compare one million images?”. *Understanding digital humanities*, Palgrave Macmillan, London, 2012, pp. 249-278

Moretti 2000 Moretti, F. “Conjectures on world literature”. *New left review*, Volume 1, Issue 54, 2000.

Moretti 2013 Moretti, F. *Distant reading*. Verso Books, 2013.

Pollock 2014 Pollock, G. “Computers can find similarities between paintings — but art history is about so much more”. *The Conversation*, Volume 22, 2014.

Rodríguez-Ortega 2020 Rodríguez-Ortega, N. “Image Processing and Computer Vision in the Field of Art History”. *The Routledge Companion to Digital Humanities and Art History*, Routledge, 2020.

Saleh et al. 2016 Saleh, B., Abe, K., Arora, R. S., and Elgammal, A. “Toward automated discovery of artistic influence”. *Multimedia Tools and Applications*, Volume 75, Issue 7, 2016, pp. 3565-3591.

Schich 2015 Schich, M. “Figuring out art history”, 2015. https://doi.org/10.48550/arXiv.1512.03301

Schich et al. 2008 Schich, M., Lehmann, S., and Park, J. *Dissecting the canon: visual subject copopularity networks in art research*, arthistoricum.net, 2008.

Schneider 2012 Schneider, M. “Steven Ramsay. Reading Machines: Toward an Algorithmic Criticism. Champaign IL: University of Illinois Press”. *Digital Studies/le Champ Numérique*, Volume 3, Issue 1, 2012. http://doi.org/10.16995/dscn.245

Sparkes 2014 Sparkes, M. “Could computers put art historians out of work?”. *Telegraph*, 2014. https://www.telegraph.co.uk/technology/news/11041814/Could-computers-put-arthistorians-out-of-work.html

Stefan et al. 2015 Stefan, J., Franzini, G., Cheema, M. F., and Gerik, S. “On close and distant reading in digital humanities: A survey and future challenges”. *Eurographics Conference on Visualization (EuroVis)*, 2015

Stork 2009 Stork, D. G. “Computer vision and computer graphics analysis of paintings and drawings: An introduction to the literature”. *International Conference on Computer Analysis of Images and Patterns*. Springer, 2009.