Seeing Constable’s Clouds:

White Paper

Level 1 Digital Humanities Advancement Grant

Award No. HAA-271801-20

Presented to the National Endowment of the Humanities by

Elizabeth Mansfield (PI), James Z. Wang (co-PI), Jia Li,
John Russell, Catherine Adams, George S. Young,
and Zhuomin Zhang

31 December 2022
Table of Contents

Project Summary.........................................................................................................................3

Computer Vision and Art History: Charting a Way Forward.........................................................4

Humanists and Scientists: Learning to Collaborate.................................................................5

Seeing Constable’s Clouds: The Research Question.................................................................6

Applying Computer Vision to an Art Historical Research Question..........................................7

Initial Results Prompt New Hypotheses and Further Experiments.........................................8

Art History with Computer Vision............................................................................................9
Project Summary

Seeing Constable’s Clouds (HAA-271801-20) enlists computer vision as part of an interdisciplinary approach to research on the paintings of British landscape artist John Constable (1776 – 1837). Constable was noted for the naturalism of his scenes of the British landscape and for his striking depictions of atmospheric effects. Constable’s depictions of clouds were perceived by his contemporaries (as well as by present-day viewers) as more life-like than clouds painted by other artists associated with the Realist movement in Europe. The art historical question animating this project seeks to understand how Constable’s clouds elicit such confidence in viewers: do human viewers perceive Constable’s paintings of clouds to be realistic because they accurately document ephemeral meteorological phenomena or because they are aesthetically persuasive to viewers accustomed to the visual language of 19th-century European landscape painting? In other words, are Constable’s clouds accurate, or do they just seem accurate to viewers who are adept with European pictorial conventions? To answer this, we developed a novel feature fusion-based convolutional neural network (CNN) model for cloud type classification, and then trained the model using a National Oceanographic and Atmospheric Administration (NOAA) dataset containing over 2,500 natural cloud photographs with corresponding labels of cloud types (cumulus, nimbus, etc.). We then applied the CNN to our dataset of paintings of clouds, which had been expertly labeled according to cloud typology by a meteorologist. The experimental results showed John Constable’s clouds are highly classifiable by cloud type and approximate more consistently the formal features of clouds documented in meteorological photographs. Datasets of clouds painted by five other artists were then prepared for assessment by the computer. Our experiments rated clouds painted by Pierre-Henri de Valenciennes (1750-1819) and Constable as the most similar to real-world clouds as documented in photographs.

Our interdisciplinary research group is based at Penn State and includes art historian Elizabeth Mansfield and computer scientist James Wang, who served as co-PIs on this project. Additional computational expertise is provided by Jia Li, a Professor of Statistics and (by courtesy) Computer Science. Experiments for this project were carried out by Information Sciences and Technology PhD student Zhuomin Zhang. John Russell, Digital Humanities Librarian and Assistant Professor, contributed project conceptualization, data management expertise, and project management assistance. George Young, Professor of Meteorology and a specialist in observational and predictive meteorology as well as AI, undertook dataset labelling and participated in machine learning verification. Data curation and structuring was overseen by Catherine Adams of the Center for Virtual/Material Studies.

1 This white paper is a condensed and revised version of the article “Techniques of the Art Historical Observer,” Nineteenth-Century Art Worldwide (Spring, 2022).
Computer Vision and Art History: Charting a Way Forward

Computer vision has proven a useful tool for technical analysis of art, yielding valuable information about condition, artistic processes and materials as well as insights into provenance and authenticity. Yet, when it comes to art historical interpretation, computer vision has been applied more tentatively. Certainly, access to the resources necessary to pursue this research—especially opportunities for collaboration with computer scientists—is an obstacle for many. NEH Digital Advancement grants are among the few sources of funding for such endeavors. But few art historians seem to be lamenting these barriers, suggesting that computer vision research is not seen as especially promising in contrast to other emerging approaches grounded in technical analysis. This apparent apathy is likely due at least in part to a disciplinary ambivalence toward computer vision’s key affordance: pattern recognition. The ability to show convincingly that one thing looks like another thing is not incidental to the practice of art history: the discipline earned its academic credibility in the late 19th and early 20th centuries precisely by deploying methodologies that revealed and sought to explain formal relationships, whether in a given artist’s oeuvre, within national or ethnic or racial borders, or across cultures and through time. The capacity to recognize and compare formal features was essential for style-based approaches to art historical analysis as it was for Giovanni Morelli’s method of connoisseurship. It is easy to understand why computer vision might be viewed as a belated tool for art historical investigation. Connoisseurship is now rarely viewed as an art


3 This is true of art historians. Scientists, on the contrary, have been energetically applying computer vision to evaluate works of art. Examples are too numerous to cite. An indication of the scope of research is suggested by the more than 1500 results produced by a search of Penn State University Libraries’s holdings and journal subscriptions using the combined search terms “computer vision” and “artwork.” As the “state of the field” essays by Wang et al. and Foka make clear, projects involving art historians represent a tiny fraction of this research output.

4 The chief example here is technical art history, which has become a major trend in art historical research in the past decade. Technical art history poses similar barriers to scholars in that it often requires access to expensive instruments and benefits from collaborative research teams with diverse forms of expertise.

5 It is difficult to reflect on the potential for computer vision to aid in the study of 19th-century art without calling to mind Jonathan Crary’s account of the emergence of the modern viewing subject around 1800. In Techniques of the Observer (1990), Crary observed a shift away from a kind of stable, Cartesian conception of vision toward a subjective, somewhat disembodied viewing subject at the start of the 19th century. An observer that exemplified Foucault’s “empirico-transcendental doublet.” The art historian—as a distinct viewing subject—arguably came into being in the midst of this transition and has been negotiating the competing allures of objective, authoritative vision and subjective, unique insight ever since. Parallels between the moment Crary identifies around 1800 and our current negotiation of the advent of computer vision and machine learning bear consideration that is beyond the scope of the present report.
historical end in and of itself but rather as subsidiary to historical interpretation and analysis. Given this historiography, suspicion of interpretive methodologies that rely on schema and patterns is understandable. Art historians are wary of becoming re-enchanted by a reductive formalism. Even so, to replace intellectual curiosity with methodological vigilance risks normalizing a narrow understanding of both the practice of art history and the research potential of computer vision.6

When applied to image analysis, recent machine learning involves preparing a neural network by training it to recognize salient features shared across different images. This is often referred to as deep learning. Applied to art historical research, deep learning can readily be used for iconographic analysis or to identify recurrent motifs—potentially useful when applied to large image datasets. What would have taken a human viewer hours of work sorting and comparing images takes the computer seconds. Developing a neural network to conduct such rapid assessments, however, requires expertise that most art historians do not possess.

Humanists and Scientists: Learning to Collaborate

Art history research involving digital or computational methods is by necessity a collaborative endeavor. Diverse sorts of specialized knowledge are required.7 In our case, the research team included an art historian, a digital humanities librarian, a data and image management specialist, a meteorologist, and three computer scientists (including one whose areas of expertise include computational and statistical modeling of meteorological phenomena). At our first few meetings, along with discussing how computer vision could be used to analyze Constable’s clouds, the group spent a lot of time just looking at his paintings and oil sketches. These close-looking sessions began in a seminar room with projections of hi-res digital images and eventually took us to the Yale Center for British Art where we could see in person some of Constable’s cloud studies and examine the skies in several of his finished landscapes and consult with additional specialists.8

Flexibility and patience are essential for collaborations between humanists and scientists. Our research cultures as well as our methodologies are quite different. To make any progress, everyone needs to be willing to depart from accustomed paths for research in their disciplines,

---

6 To cite just one obvious example, formal comparisons are still used to help identify pottery sherds. For a case study on how computer vision was used for this work, see Jun Zhou, Haozhou Yu, Karen Smith, Colin Wilder, Hongkai Yu, Song Wang, “Identifying designs from incomplete, fragmented cultural heritage objects by curve-pattern matching,” *Journal of Electronic Imaging* 26(1) 011022 (5 January 2017) https://doi.org/10.1117/1.JEI.26.1.011022.
7 The likelihood that a single researcher might possess all the requisite art historical, technical, and scientific knowledge and skills needed to carry out this kind of research is very small. Hence the expression, “DH unicorn.” Such researchers exist, but they are rare. To move research forward in this area, collaborations are essential.
8 One of the benefits of these kinds of collaborations is that, contrary to what one might assume, it actually promotes a kind of slow, close looking that can be revelatory in and of itself. Our March 2019 visit to the Yale Center for British Art also benefitted from conversations with Damon Crockett, Jessica David, Nicholas Robbins, and Scott Wilcox.
and the collaborative endeavor takes precedence over individual research aims. Further informing our collaborative endeavor was the fact that a key member of our team was a PhD student, so we had to be especially mindful of the impact our research and workflow could have on her degree progress and professional ambitions. Humanities scholars contemplating collaborations like ours will want first to consider whether this kind of work is right for them. Cross-disciplinary collaborations can be frustrating, especially for humanities scholars who are used to working alone and having complete control over their research agenda.9 One of the first things our project team had to do was to work together to refine the central research question. Fortunately, a research question suitable for computer vision readily emerged from the existing scholarship on Constable’s clouds.

Seeing Constable’s Clouds: The Research Question

An intriguing (and contested) explanation for the striking naturalism of Constable’s clouds was put forward by Kurt Badt in his 1950 book, John Constable’s Clouds. Badt argued that Constable’s ability to paint clouds improved dramatically around 1821 and that this change was due to the artist’s belated introduction to cloud taxonomy. The classification of clouds into different types—cumulus, cirrus, stratus, etc.—was proposed by British chemist Luke Howard at an 1802 lecture in London and published the following year.10 In Badt’s view, the difference in quality he discerned in clouds painted before 1821 and those made later could only be explained by a profound change in Constable’s understanding of clouds around that time—a new understanding based on a belated encounter with Howard’s cloud taxonomy. Badt had no direct evidence to support his contention. Constable doesn’t mention Luke Howard’s meteorological studies in his correspondence and there’s no evidence that he ever possessed Howard’s book or attended his lectures.11 Given the rapid and widespread popularization of Howard’s system in the years after its 1803 publication, it’s altogether possible that Constable was familiar with the taxonomy well before 1820. But it is just as possible that Constable learned about Howard’s taxonomy later. Over the years, Badt’s theory has been alternately

---

9 For an insightful account of interdisciplinary, collaborative work with scientists written by an anthropologist of research cultures, see Park Doing, Velvet Revolution at the Synchrotron: Biology, Physics, and Scientific Change (Cambridge: MIT Press, 2009).


11 As evidence that Constable was aware of Luke Howard’s taxonomy, Badt cites the artist’s 12 December 1830 letter to George Constable that apparently accompanied a “book you lent me long ago.” The letter then discusses “My observations of clouds and skies” and observes, perhaps in reference to the book he has just returned, “Forster’s is the best book; he is far from right, but still has the merit of breaking much ground.” Badt, 50. The first chapter of Thomas Forster’s Researches about Atmospheric Phenomena (1813) is an explication of “Mr. Howard’s Theory of the Origin and Modifications of Clouds.” But Badt himself raises doubts about Forster being the source of Constable’s knowledge of the taxonomy, instead speculating that Constable most likely read Luke Howard’s 1818 book, The Climate of London, Deduced from Meteorological Observations, around the time it was published—not long before the painter embarked on his most sustained period of skying in 1821–22. Badt cites no evidence for this other than the book’s appeal to “a wider public” than Howard’s Essay on the Modifications of Clouds, first published in 1803. Badt, 50–61.
dismissed and accepted by art historians. His argument relies mostly on Badt’s own visual assessment of the relative naturalism of Constable’s clouds before and after 1821, so it’s precisely the kind of art historical conversation to which computer vision might helpfully contribute.

**Applying Computer Vision to an Art Historical Research Question**

To test Badt’s theory, we first sought to determine whether Constable’s clouds could be readily classified by type. In other words, do Constable’s clouds exhibit the defining features that Luke Howard ascribed to cumulus, nimbus, stratus, and cirrus? The first step in this experiment was to train the CNN to distinguish clouds by type using a large dataset of cloud photographs created by NOAA that is labeled according to the taxonomy first proposed by Howard and still in use today. Once the CNN had been trained to accurately distinguish “cumulus” from “cirrus” from “stratus,” it was ready to assess Constable’s clouds by scoring each painted cloud as a probable match with one of the standard types. But, before putting the trained CNN to work on Constable’s clouds, we needed to create a ground truth against which to compare the CNN’s output. For this, we enlisted the expertise of a meteorologist to identify Constable’s clouds by type.

The CNN then took its turn assessing the paintings. Constable’s clouds performed well in terms of conforming to typological characteristics, and the computer’s assignment of cloud types largely matched the meteorologist’s. So the training of the CNN appeared to be successful. The results suggested that Constable recognized the distinctive features of cumulus, stratus, and cirrus clouds and noted those features in his paintings. While these results might weigh in favor of Badt’s argument, the CNN (like the meteorologist) did not register a change in Constable’s performance in this area around 1821.

To add texture to the experiment, we expanded our dataset to include another two hundred images of clouds painted by Constable’s known emulators and other contemporaries noted for

---


13 Zhang’s account of the results as well as precise (including mathematical) descriptions of the programming of the CNN, her analysis of the results, and the steps she took to confirm the results can be found in Zhuomin Zhang, Elizabeth C. Mansfield, Jia Li, John Russell, George S. Young, Catherine Adams, and James Z. Wang, “A Machine Learning Paradigm for Studying Pictorial Realism: Are Constable’s Clouds More Real than His Contemporaries?,” arXiv preprint arXiv:2202.09348 (2022), https://doi.org/10.48550/arXiv.2202.09348.

14 Meteorologist George Young’s assessment of Constable’s depiction of specific types of clouds was the first measure of the artist’s accuracy in this vein, the CNN confirmed Young’s analysis.

15 Our dataset of Constable clouds includes 86 individual paintings comprising both cloud studies and studio paintings where at least one-third of the composition is given to a cloudy sky. Of these, 21 were created before 1821. We plan to analyze our results further both in relation to chronology and also in relation to studio versus plein-air execution.
their illusionistic skies and their _plein air_ paintings of clouds. The results produced by the CNN again aligned with the meteorologist’s ground truth: Constable’s clouds were “better” than his contemporaries in that they could be classed pretty consistently by their features into Howard’s typology. Interestingly, Henri de Valenciennes’ clouds (mostly painted in the 1780s and 1790s) scored nearly as well as Constable’s in terms of typological consistency. This result, while not dispositive of Badt’s argument, does suggest that empirical study and long practice of _plein air_ painting of clouds could equip artists striving for a certain kind of naturalism with the ability to depict clouds accurately by type—even if they are not familiar with Howard’s taxonomy.

**Initial Results Prompt New Hypotheses and Further Experiments**

Now confident in the CNN’s ability to compare individual clouds—both photographic and painted—across our dataset, we decided to run an experiment in which the painted clouds were assessed according to their similarity or difference from photographic clouds from NOAA, proceeding from the hypothesis that photographs of clouds are the most realistic images of clouds. Here, it’s worth pausing a moment to acknowledge what is presumed by this hypothesis. Photographs—even unartful, climate research photographs—are no less forms of representation than are paintings. NOAA’s cloud photographs may not involve a human photographer (many are taken automatically), but they are nonetheless mediated by the positions of the cameras, by the decisions that led to the photographic project in the first place, by whatever editing may have been done, and by the technology of digital photography itself. This list of ways in which photographs of clouds are different from real clouds could go on. That said, our experiment was founded on the reasonable belief that photographic clouds are the most empirically faithful visual representations of clouds we have. Additionally, the representational biases of photography are not irrelevant to the study of naturalism of cloud paintings by Constable and his contemporaries. The invention of photography in the 1820s and 1830s was a response to already emergent European notions about what representation could or should be as much as it was an outcome of specific technological innovations and aspirations. Naturalism emerged from this same milieu. That artists like Constable might have been working toward a mode of representation in some ways akin to that produced by photography is probable. So to posit photographs as a kind of standard for comparison raised

---

16 We limited our dataset to representations of clouds executed in oil paint for sake of consistency in medium. The artists included were Lionel Constable (1828–1887), Frederick R. Lee (1798–1879), and Frederick W. Watts (1800–1870); Pierre-Henri de Valenciennes (1750–1819), David Cox (1783–1859), and Eugène Boudin (1824–1898). The team’s meteorologist labeled these images according to cloud type—when discernable—and apparent conditions along with his subjective “realism” score. Those interested in learning more about the image dataset are invited to contact Elizabeth Mansfield, [https://arts.psu.edu/faculty/elizabeth-mansfield/](https://arts.psu.edu/faculty/elizabeth-mansfield/). Metadata for the paintings used in our research is available at GitHub, [https://github.com/ECMARTH/ConstablesClouds.git](https://github.com/ECMARTH/ConstablesClouds.git).


---
interesting possibilities for computer vision as one way to analyze immediately pre-
photographic and post-photographic pictorial strategies.

The results of this experiment also proved interesting. During the image translation process
from paintings to photos, the CNN-based encoder can disentangle the “style” features from
each painting. Here, “style” is used in reference to a set of formal features that can be
described mathematically and thus compared. For any given set of images (e.g. NOAA
photographs of clouds or painted clouds by Constable or painted clouds by all artists in the
dataset), distinctive formal features are identified by the CNN and these features can be further
used to measure formal similarity between paintings and photos.\(^{18}\) Henri de Valenciennes’
clouds were scored by the CNN as the most similar to photographic clouds. John Constable
ranked quite closely behind Valenciennes, with Constable’s emulators Lionel Constable,
Frederick W. Watts and Frederick R. Lee trailing them slightly.\(^{19}\)

**Art History with Computer Vision**

What might these experiments with computer vision contribute to the practice of art history?
For one thing, working with computer vision encourages greater attentiveness to one’s own
habits of looking, both those habits formed by disciplinary protocols as well as those arising
from personal biases. The myriad ways in which bias can influence or undermine AI-based
analyses of works of art became a particular concern for us, resulting in an article that proposes
best practices for identifying and mitigating the effects of bias.\(^{20}\) Second, experimentation with
computer vision might expand our understanding of 19th-century Realist painting and its
relationship to modernism. Helpful in this respect have been experimental results like the
comparable scores given to Henri de Valenciennes and John Constable in terms of their
approximation to photographs of clouds. The degree to which this finding provides accurate
information about Valenciennes’ and Constable’s relative “realism” matters less than the
invitation it extends to look more closely, not just at their art but also at their respective
historiographies. Art historical narratives tend to align Constable with the origins of naturalism
and avant-garde modernism in French landscape painting while characterizing Valenciennes as

\(^{18}\) Y. Bar, N. Levy, and L. Wolf, “Classification of Artistic Styles Using Binarized Features Derived from a Deep Neural
Classification Based on Learnt Deep Correlation Features,” IEEE Transactions on Multimedia 20, no. 9 (2018):
2491–2502.

\(^{19}\) Zhuomin Zhang is the lead author on an article that details her experiments. This forthcoming publication will
provide a more extensive account of the results as well as precise (including mathematical) descriptions of the
programming of the CNN, her analysis of the results, and the steps she took to confirm the results. In the
meantime, those interested in learning more about the computer science involved in the project are encouraged
to contact James Wang, https://ist.psu.edu/directory/jzw11. In addition to this forthcoming article, the preprint
cited in note no. 13, and the article published in Nineteenth-Century Art Worldwide cited in note no. 1, we also
presented our research at the Technological Revolutions and Art History Symposium, hosted by the Frick Art
Reference Library on 11 March 2021.

\(^{20}\) Z. Zhang et al., “Reducing Bias in AI-Based Analysis of Visual Artworks,” in IEEE BITS the Information Theory
a stalwart classicist and traditionalist. Yet, as the experiments with computer vision help to show, the two painters have a good deal in common when it comes to formal interests and representational strategies. If computer vision can point to avenues for further inquiry—in the case of our study, by confirming the relevance ongoing transnational research on European art around 1800 and by inviting a more complex understanding of Constable’s place in art history as both a precursor to modernism and a culminating painter of the 18th century—its use as a complement to existing techniques of art historical analysis has arrived.