AI and the Digitized Photoarchive:

Promoting Access and Discoverability

Ellen Prokop, *National Gallery of Art USA*
X.Y. Han, *Cornell University USA*
Vardan Papyan, *University of Toronto Canada*
David L. Donoho, *Stanford University USA*
C. Richard Johnson Jr., *Cornell University USA*

Abstract—The Frick Art Reference Library in New York launched a pilot project with Stanford University, Cornell University, and the University of Toronto to develop an algorithm that applies a local classification system based on visual elements to the library's digitized Photoarchive. As a test case, the Cornell/Toronto/Stanford team focused on a dataset of digital reproductions of North American paintings and drawings and employed recent advances in artificial intelligence and machine learning to produce automatic image classifiers. The results of this preliminary experiment suggest that automatic image classifiers have the potential to become powerful tools in metadata creation and image retrieval.

INTRODUCTION

In January 2018, the Frick Art Reference Library (FARL) launched a pilot project in collaboration with Stanford University, Cornell University, and the University of Toronto to leverage recent advances in artificial intelligence (AI) and machine learning (ML) to develop an algorithm that applies a local classification system based on visual elements

Ellen Prokop, X.Y. Han, and Vardan Papyan served as co-first authors and contributed equally to this article. Ellen Prokop was recently digital art history lead, The Frick Collection and Frick Art Reference Library, New York, New York, and is currently museum specialist for pre-1900 French art, National Gallery of Art, Washington, D.C.; E-Prokop@nga.gov. X.Y. Han is a PhD student in operations research and information engineering, Cornell University, Ithaca, New York; xh332@cornell.edu. Vardan Papyan is assistant professor, Department of Mathematics, cross-appointed to the Department of Computer Sciences at the University of Toronto, Ontario; papyan@math.toronto.edu. David L. Donoho is professor of statistics and the Anne T. and Robert M. Bass Professor of Humanities and Sciences at Stanford University, Stanford, California; donoho@stanford.edu. C. Richard Johnson, Jr. is professor of electrical and computer engineering and the Geoffrey S.M. Hedrick Sr. Professor of Engineering at Cornell University, Ithaca, New York, and the senior research advisor to the Frick Art Reference Library; crj2@cornell.edu.

Art Documentation: Journal of the Art Libraries Society of North America, vol. 40 (spring 2021)
0730-7187/2021/4001-0001$10.00.

to FARL's digitized photograph study collection, or Photoarchive. Such a tool promises to prove valuable not only to FARL staff struggling to process a backlog of tens of thousands of reproductions from the Photoarchive that are digitized but not-yet-cataloged but also to those members of the international archive community seeking new ways to manage their digitized image collections. Significantly, AI's potential to streamline cataloging and expand access to and discoverability of digitized image collections has not yet been fully realized. A 2018 Ex Libris survey revealed that nearly 80 percent of research librarians were exploring the use of AI to enhance productivity, but only about 5 percent of these institutions were currently utilizing this technology.[1] Several concerns deter the widespread use of AI among such institutions. Aside from the expense of researching, developing, and implementing new technologies, there is the fear that the application of AI in the archive will reduce library staff, undermine traditional methods of scholarship and learning, and propagate misinformation and bias.[2] These concerns are legitimate; however, the collaboration between the Frick Art Reference Library and Cornell/Toronto/Stanford demonstrates that AI can fulfill a positive role in the formation of the digital archive, potentially resolving multiple challenges facing librarians, archivists, and digital heritage specialists, including budgetary concerns and the pressing issue of scale. As the results of this pilot project suggest, AI has the potential to improve the operational efficiency of librarians and archivists, reduce the possibility of introducing mistakes into the catalog, and help researchers refine their results within the online archive. In addition, AI can save staff time, which may then be devoted to strategic priorities such as improving user experience, assisting students and scholars, and attracting more diverse audiences.[3] For the Frick Art Reference Library specifically, embracing AI has enhanced the discoverability of an important but underutilized research collection, its Photoarchive. By increasing the Photoarchive's accessibility, this technology will, in short, add to the value of this resource. The present article introduces FARL's Photoarchive and reports on the first stage of the pilot project, which was completed in August 2020.

THE FRICK ART REFERENCE LIBRARY AND ITS PHOTOARCHIVE

The Photoarchive is a study collection of 1.2 million reproductions of fine and decorative arts of the Western tradition produced from the fourth through the twentieth centuries. This research collection, the founding collection of the library, was established

1. "How AI Can Enhance the Value of Research Libraries," *Library Journal*, April 15, 2019, https://www.libraryjournal.com/?detailStory=how-ai-can-enhance-the-value-of-research-libraries. Also see the white paper prepared by Ex Libris: "Artificial Intelligence in the Library: Advantages, Challenges and Tradition," https://cdn2.hubspot.net/hubfs/2909474/Ex%20Libris%20Artificial%20Intelligence%20White%20Paper.pdf. One institution that has explored the potential of computer vision to assist in the processing of digital photograph collections is Carnegie Mellon University (CMU) Libraries. In the summer of 2020, CMU Libraries staff developed a prototype that successfully isolates duplicates and tags photographs depicting similar content in CMU Archives' General Photograph Collection. For additional information on this pioneering project, see Julia Corrin, Emily Davis, Matthew Lincoln, and Scott B. Weingart, "CAMPI: Computer-Aided Metadata Generation for Photo Archives Initiative," Carnegie Mellon University, October 8, 2020, https://doi.org/10.1184/R1/12791807.v2.

2. Jackie Snow, "Bias Already Exists in Search Engine Results and It's Only Going to Get Worse," *MIT Technology Review*, February 26, 2018, https://www.technologyreview.com/2018/02/26/3299/meet-the-woman-who-searches-out-search-engines-bias-against-women-and-minorities/.

3. "How AI Can Enhance the Value of Research Libraries."

Figure 1. Portrait of Helen Clay Frick in her office at the Frick Art Reference Library, 1939, photographer unknown. Frick Family Photographs. Courtesy of The Frick Collection/Frick Art Reference Library Archives.

in 1920 by the philanthropist Helen Clay Frick (1888–1984) (Figure 1) as a memorial to her father, the industrialist Henry Clay Frick, who in addition to his professional activities assembled one of North America's finest private collections of European painting, sculpture, and decorative arts.[4] Henry Clay Frick shared his passion for the Old Masters with his daughter. Through study and travel, Helen gained a strong familiarity with European art, particularly Italian and Spanish painting, and often researched the works her father acquired or considered acquiring, documenting their provenance.[5]

4. For information on Henry Clay Frick's collection, see Martha Frick Symington Sanger, *Henry Clay Frick: An Intimate Portrait* (New York: Abbeville Press Publishers, 1998) and Colin B. Bailey, *Building the Frick Collection: An Introduction to the House and Its Collections* (New York: The Frick Collection in association with Scala, 2006).

5. For information on Helen Clay Frick, see Martha Frick Symington Sanger, *Helen Clay Frick: Bittersweet Heiress* (Pittsburgh: University of Pittsburgh Press, 2008).

Confident that his daughter would safeguard his legacy, in his will Henry Clay Frick appointed her one of the trustees of the museum he planned to leave the city of New York, naming her the sole heir of the collection should the city refuse his bequest. After his death on December 2, 1919, Helen immediately began preparations for the foundation of the art gallery that would bear the family name. Throughout the early months of 1920, she sought to advance her understanding of the museum sciences and expand her network of contacts by writing to gallery directors across Europe and the United States for their advice, and in the summer she embarked on a tour of Europe to obtain new information regarding the works of art in her father's collection and scout for future acquisitions for the new public gallery.

While abroad she visited London, where she paid a call to Sir Robert Witt, the historian, art collector, and co-founder of both the National Art-Collections Fund (now the Art Fund) and the Courtauld Institute of Art, University of London.[6] With his wife, the former Mary Helene Marten, Sir Robert Witt had formed a comprehensive study collection of photographs and print reproductions of Western paintings and drawings from the twelfth century onwards, a resource that eventually comprised approximately 750,000 items. This research photograph collection, assembled at a time before books on art historical subjects were extensively illustrated, was invaluable to the couple's collecting activities. In 1944, Sir Robert Witt deeded the archive to the University of London, where it became known as the Witt Library.[7] When Helen Clay Frick toured the Witts' photograph collection in the summer of 1920, she immediately recognized its importance for the study of art and resolved to found a similar archive in New York City, a resource that would be freely accessible to the public and would honor her father's legacy by advancing the field of art history in North America.[8]

Throughout the remainder of her tour of Europe, Frick laid the groundwork for the foundation of her own art research library, obtaining photographs and research materials and contacting scholars for information and advice. During her visit to Paris, she engaged the art historian and librarian Clotilde Misme, later Brière-Misme (Figure 2), a specialist in Dutch painting of the Golden Age, to act as her agent in France and Holland and to secure photographs, books, and auction catalogs for the new institution.[9] Frick returned to New York in November and immediately founded her research photograph collection as the Frick Art Reference Library.[10] The archive was originally housed in the billiard room and adjacent bowling alley located in the subbasement of the Frick residence at 1 East 70th Street (Figure 3). To accommodate the library's growing collections as well as a reading room for the public, an elegant one-story building designed by the architectural firm Carrère and Hastings (the firm that had designed the Frick mansion) was constructed at 6 East 71st Street four years later

6. Sanger, *Helen Clay Frick*, 132; Katharine McCook Knox, *The Story of the Frick Art Reference Library: The Early Years* (New York: The Library, 1979), 5–6.

7. "Sir Robert Witt," Dictionary of Art Historians, https://arthistorians.info/wittr.

8. Knox, *The Story of the Frick Art Reference Library*, 8–9.

9. Knox, *The Story of the Frick Art Reference Library*, 11.

10. Knox, *The Story of the Frick Art Reference Library*, 15.

Figure 2. Clotilde Brière-Misme in Fontenay-en-Parisis (France), July 1960, photographer unknown. Frick Family Photographs. Courtesy of The Frick Collection/Frick Art Reference Library Archives.

(Figure 4).[11] In the early 1930s, as the Frick residence underwent its conversion from a private home into a public museum, the 1924 building was demolished to make room for additional public galleries. A third—and final—home for the library was built on a lot adjacent to the museum at 10 East 71st Street. The thirteen-story building designed by John Russell Pope opened to the public on January 14, 1935 (Figure 5).[12] By that time, however, the library had been extant for fifteen years.

THE ORGANIZATION OF THE PHOTOARCHIVE

Within weeks of the founding of the library, Frick hired two staff members to help her plan and organize her research photograph collection. Although modeled on the Witts' library, FARL's Photoarchive expands on its prototype in three ways. First, the collection's scope extends to Western sculpture and architectural ornament. Second, documentation for each work of art represented in the archive is more comprehensive. Finally, the Photoarchive is organized according to a numerical system.

To prepare images for inclusion in the archive, library staff mounted black-and-white photographs or reproductions cut from sales catalogs on nine-by-twelve-inch sheets of archival-quality gray cardboard. The artist or, in the event the artist was not identified, the national school, title or subject, collection, medium, and dimensions of the work of art were noted on the front of the sheets, just as in the Witts' archive. Frick and

11. Knox, *The Story of the Frick Art Reference Library*, 26–7.
12. Knox, *The Story of the Frick Art Reference Library*, 33.

Figure 3. Photograph of the original home of the Frick Art Reference Library in the sub-basement of the Frick residence at 1 East 70th Street, New York, before 1924, photographer unknown. Courtesy of The Frick Collection/Frick Art Reference Library Archives.

her team, however, sought to provide more complete documentation for each work of art in the collection and also included information about the object's date of execution, attribution history, exhibition history, conservation history, provenance, and physical characteristics on the reverse of the sheet. Supplementary data, such as the source of the mounted photograph or reproduction, a record of other sources of reproductions, and a bibliography, was also added. Following the Witts' "system of arrangement" within their archive, these sheets or "photo study mounts" were grouped by national school and filed alphabetically by artist and subdivided by subject.[13] If the artist was unknown, the work was filed under the national school and subject only. Yet as noted above, Photo-archive staff expanded on the Witts' arrangement and applied a numerical classification system to the subject categories, one that incorporated the artist's national school. Thus, a full-length portrait of a man by the British artist Thomas Gainsborough was not simply filed away under the artist's name in a folder labeled "portraits" as it was in the Witts' archive but cataloged under the artist's name with the classification heading "221–1" (Figures 6 and 7). The breakdown of this heading is as follows: the first "2" in the series designates the British School; the following "21" denotes a portrait of a man;

13. Knox, *The Story of the Frick Art Reference Library*, 17.

Figure 4. Photograph of the first library building at 6 East 71st Street, ca. 1924, photographer unknown. Courtesy of The Frick Collection/Frick Art Reference Library Archives.

and the "1" included after the dash indicates a full-length standing subject (as opposed to one mounted on a horse, which would be indicated with a "3"). A group portrait by Gainsborough of a man and woman, such as the famous double portrait of *Mr. and Mrs. Andrews* in the National Gallery, London, is classified as "224–5" ("British School: Portrait Groups: Men and Women"); however, Gainsborough's intimate oil sketch of his two daughters also in the collection of the National Gallery is classified as "224–4" ("British School: Portrait Groups: Children").[14]

This in-house classification system based on subject matter affords a fixed filing position for each photo study mount, serving as an access point for retrieval as well as a discovery point for research since it allows users to narrow their search through the archive, targeting specific themes rather than having to sift through hundreds, if not thousands, of images associated with one artist or national school.[15] The system as designed by library staff originally encompassed approximately 500 headings; additional terms were added as needed and to date, staff employ 673 headings. Although the system is quite detailed, it has proven remarkably effective because of the flexibility it offers the archivists as they update existing records. For example, if a work of art is

14. Accession numbers NG6301 and NG3812, respectively.

15. Knox, *The Story of the Frick Art Reference Library*, 17.

Figure 5. Construction of the new library building at 10 East 71st Street, looking east; the first library building can be seen to the right of the construction site, August 2, 1934, photographed by Alfred Cook. Frick Collection and Frick Art Reference Library Construction Photographs, 1933–1935. Copyright The Frick Collection Archives.

deaccessioned or changes hands, staff simply notes that information on the back of the mount, adding it to the provenance history of the object: no new or updated classification heading is necessary. An added advantage of this system is that each file for an identified artist—of which the library has approximately 44,000 examples—essentially functions as an unbound catalogue raisonné, providing images and documentation for each work by the artist as collected by the library.

Unfortunately, producing photo study mounts is an extremely time-consuming process. By 1923, only three years after the establishment of the library, a significant backlog had developed. Advised by Sir Robert and the art historian and collector Dan Fellows Platt, who had become a frequent visitor to the library beginning in 1922,

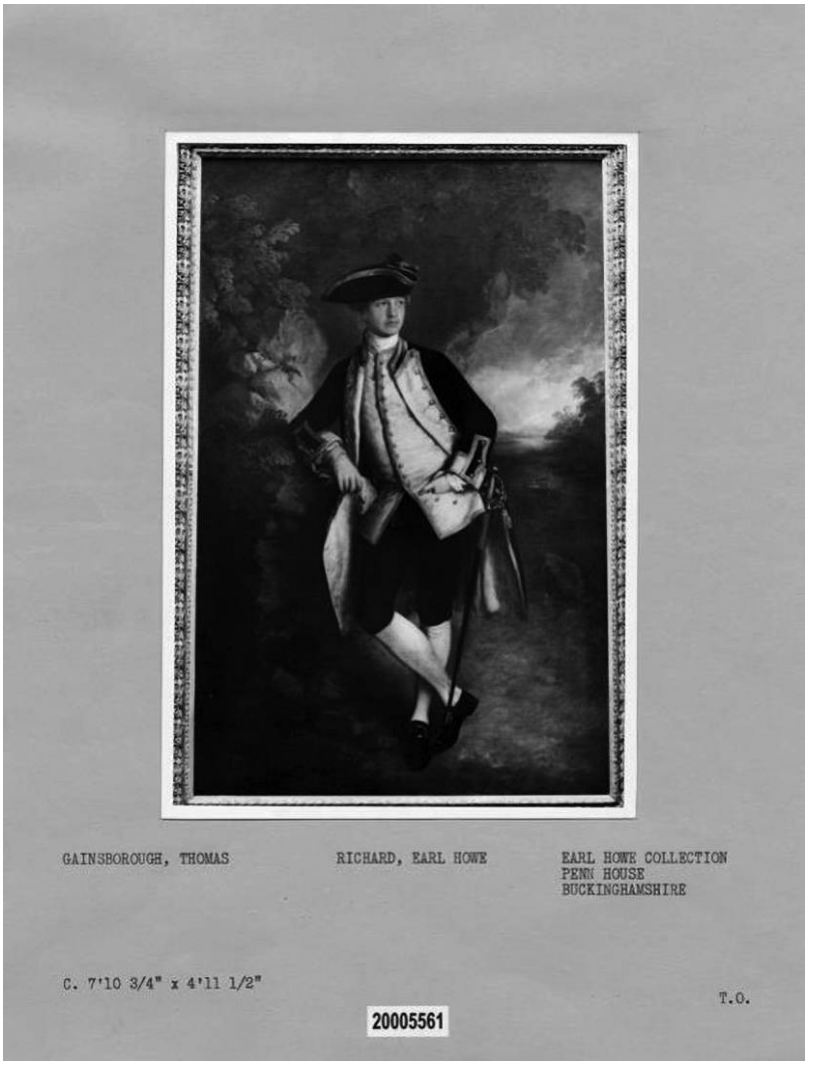
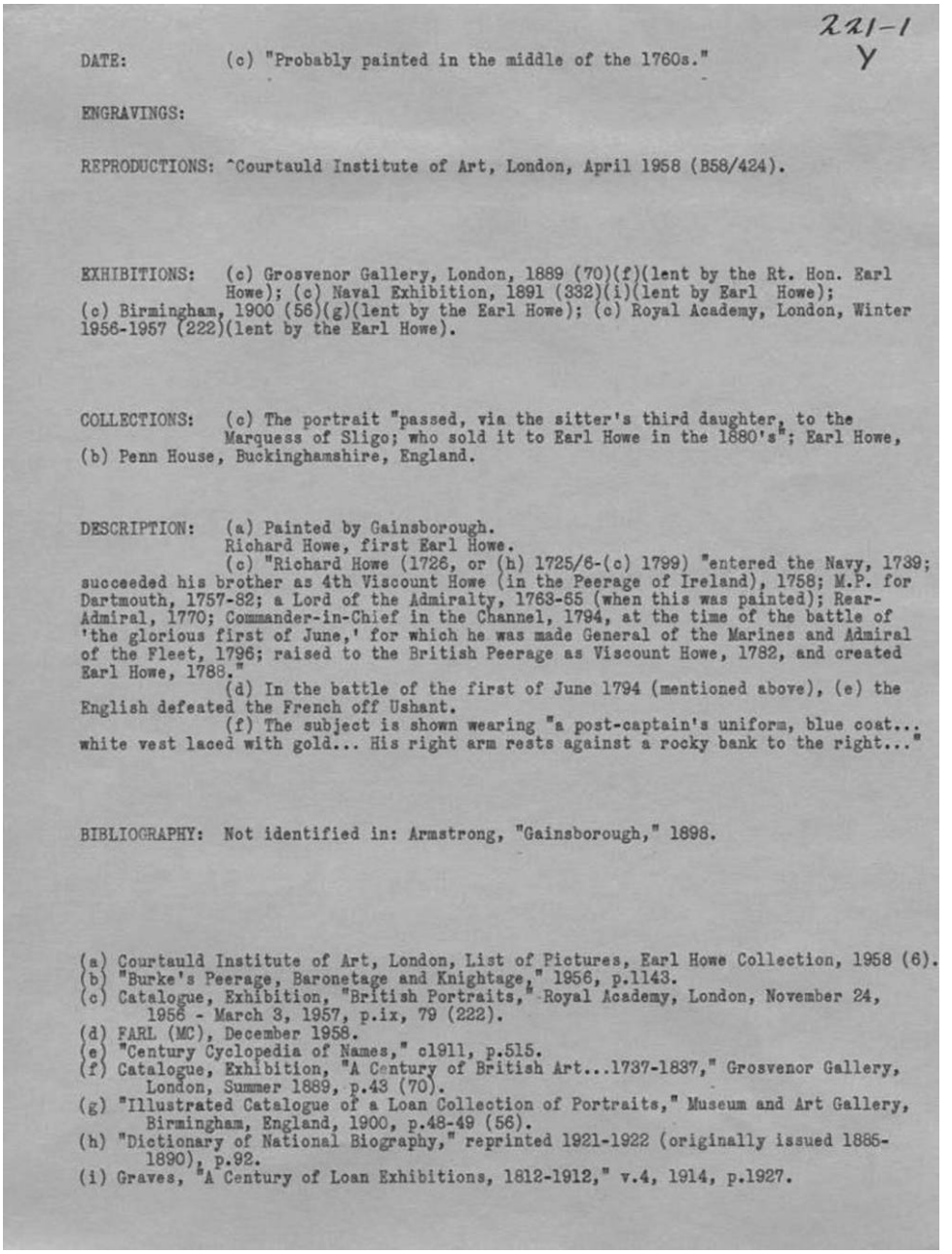


Figure 6. Photo study mount of Thomas Gainsborough's *Richard, Earl Howe* (ca. 1780s), front. Courtesy of the Frick Art Reference Library Photoarchive.

Photoarchive staff ultimately decided to release the backlogged photographs to researchers without complete documentation and classification headings, determining it was more beneficial to the library's public to release these images with minimal information rather than store them indefinitely.[16] Therefore, approximately 262,000 photographs and reproductions are mounted and classified according to the in-house system—these are known in the library as "classified mounts." The remainder have minimal documentation (i.e., author, title, collection, and date) noted on the reverse

16. Knox, *The Story of the Frick Art Reference Library*, 24–5. This minimal processing approach would later be advocated by Mark A. Greene and Dennis Meissner. See Greene and Meissner, "More Product, Less Process: Revamping Traditional Archival Processing," *American Archivist* 68 (Fall/Winter 2005): 208–63.



221-1
Y

DATE: (c) "Probably painted in the middle of the 1760s."

ENGRAVINGS:

REPRODUCTIONS: ^Courtauld Institute of Art, London, April 1958 (B58/424).

EXHIBITIONS: (c) Grosvenor Gallery, London, 1889 (70)(f)(lent by the Rt. Hon. Earl Howe); (c) Naval Exhibition, 1891 (332)(i)(lent by Earl Howe); (c) Birmingham, 1900 (56)(g)(lent by the Earl Howe); (c) Royal Academy, London, Winter 1956-1957 (222)(lent by the Earl Howe).

COLLECTIONS: (c) The portrait "passed, via the sitter's third daughter, to the Marquess of Sligo; who sold it to Earl Howe in the 1880's"; Earl Howe, (b) Penn House, Buckinghamshire, England.

DESCRIPTION: (a) Painted by Gainsborough.
Richard Howe, first Earl Howe.
(c) "Richard Howe (1726, or (h) 1725/6-(c) 1799) "entered the Navy, 1739; succeeded his brother as 4th Viscount Howe (in the Peerage of Ireland), 1758; M.P. for Dartmouth, 1757-82; a Lord of the Admiralty, 1763-65 (when this was painted); Rear-Admiral, 1770; Commander-in-Chief in the Channel, 1794, at the time of the battle of 'the glorious first of June,' for which he was made General of the Marines and Admiral of the Fleet, 1796; raised to the British Peerage as Viscount Howe, 1782, and created Earl Howe, 1788."
(d) In the battle of the first of June 1794 (mentioned above), (e) the English defeated the French off Ushant.
(f) The subject is shown wearing "a post-captain's uniform, blue coat... white vest laced with gold... His right arm rests against a rocky bank to the right..."

BIBLIOGRAPHY: Not identified in: Armstrong, "Gainsborough," 1898.

(a) Courtauld Institute of Art, London, List of Pictures, Earl Howe Collection, 1958 (6).
(b) "Burke's Peerage, Baronetage and Knightage," 1956, p.1143.
(c) Catalogue, Exhibition, "British Portraits," Royal Academy, London, November 24, 1956 - March 3, 1957, p.ix, 79 (222).
(d) FARL (MC), December 1958.
(e) "Century Cyclopedia of Names," c1911, p.515.
(f) Catalogue, Exhibition, "A Century of British Art...1737-1837," Grosvenor Gallery, London, Summer 1889, p.43 (70).
(g) "Illustrated Catalogue of a Loan Collection of Portraits," Museum and Art Gallery, Birmingham, England, 1900, p.48-49 (56).
(h) "Dictionary of National Biography," reprinted 1921-1922 (originally issued 1885-1890), p.92.
(i) Graves, "A Century of Loan Exhibitions, 1812-1912," v.4, 1914, p.1927.

Figure 7. Photo study mount of Thomas Gainsborough's *Richard, Earl Howe* (ca. 1780s), back, with classification number at upper right. Courtesy of the Frick Art Reference Library Photoarchive.

and are preserved in archival polyester film sleeves—these are known as "supply photographs." Yet as the Photoarchive is digitized, all images—including the supply photographs—will have to be classified to increase online access and discoverability. Thus, this unwelcome division within the collection will eventually be repaired.

THE IMPACT OF THE PHOTOARCHIVE

Through the establishment of her art research library, Frick provided the public with the opportunity to study reproductions of hundreds of thousands of works of art from across Europe and the Americas in one location. Not only were students, scholars, and art market professionals able to expand the scope of their research, but they were also able to interact with images in a manner that was not possible with print publications. They could move the photo study mounts around at will and sort and compare the images, reorganizing them into additional categories as their research questions evolved. This new way of working with images encouraged many art historians to shift their attention away from artists' biographies (the traditional focus of nineteenth-century art historians and critics) to comparative stylistic analysis, a development that stimulated in part the rise of formalism among North American art historians active in the first half of the twentieth century.

The Frick Art Reference Library's Photoarchive also motivated advancements in the study of North American painting. To expand the Photoarchive's holdings, Frick organized dozens of photograph expeditions between 1922 and 1967 to record significant and rarely reproduced works of art in collections located throughout the United States.[17] Hired photographers and library staff toured the country and visited the homes of private collectors to photograph the works of art in their care and obtain crucial information regarding their attribution and provenance and, in the case of portraits, the identity of the sitters. Thus, the Photoarchive contains much rich data about tens of thousands of collectors and portrait sitters and is an important resource for the study of North American genealogy. Photographers and staff members also visited various public institutions such as local libraries, schools, and municipal buildings to document their holdings. Since these collections are rarely cataloged, the Photoarchive offers a rich history of public art in the United States. The resulting collection of approximately 35,000 original negatives from these "field trips," which in many instances document works of art that have subsequently been altered, lost, or destroyed, has become one of the library's most treasured resources.

Frick sponsored a similar initiative in Italy. Between 1925 and 1951, she commissioned the Italian photographer Mario Sansoni to document paintings, frescoes, and sculpture located in remote churches and regional collections throughout the peninsula.[18] Sansoni produced approximately 8,800 original negatives, a collection supplemented by the more than 26,000 photographs he was able to acquire for the library

17. Knox, *The Story of the Frick Art Reference Library*, 39.

18. Knox, *The Story of the Frick Art Reference Library*, 30–1. For information on Mario Sansoni, see "Bencini e Sansoni," Istituto Centrale per il Catalogo e la Documentazione (ICCD), http://www.iccd.beniculturali.it/it/194/fondi-fotografici/3874/bencini-e-sansoni.

from other photographers and agents. As with the negatives resulting from the North American field trips, many of the negatives and prints produced or purchased by Sansoni are the only record of works of art that have since been destroyed or become inaccessible to the public.[19] A third campaign launched by Frick documented works of art sold at London auctions between 1921 and 1932; these reproductions record thousands of European paintings and drawings that remain in private hands, again providing information that is not widely available to students and specialists. Because the library owns the copyright to the negatives of these three research collections (approximately 57,000 items in total), they were the first resources to be digitized and made available on the Artstor Digital Library and The Frick Collection's online digital archive, The Frick Digital Collections.[20]

DIGITIZING THE PHOTOARCHIVE

The library continues to acquire reproductions and has amassed more than 1.2 million reproductions since its founding. In the late 1990s, staff began digitizing the Photoarchive, and this project will continue through 2025. By September 2020, digital images and documentation for more than 317,500 works of art had been made freely available—that is, without fees or copyright restrictions—online for consultation and download through The Frick Digital Collections. By December 2022, approximately 230,000 supply images are scheduled to be uploaded. Thus, in the next two years, The Frick Digital Collections will host images and metadata representing more than 547,500 works of art. Cataloging and metadata creation, however, are not keeping pace with digitization. Another backlog—of images that have been digitized but are not yet cataloged—is growing rapidly. Library staff agreed that the accessibility and discoverability of the Photoarchive is the priority; thus, it has accepted that all digital images will be released online with minimal documentation, at least initially. Photoarchive staff plans to enhance the online catalog once the entire research collection has been digitized. Yet even applying minimal documentation—the artist, current title, date, and classification heading for each work of art—is time-consuming. Photoarchivists considered crowd-sourcing as one means to increase the rate of metadata creation, but preliminary experiments were unsatisfactory. Applying the in-house classification system described above proved too restrictive for many volunteers, who resorted to tagging the images in their own manner. Tagging certainly increases the discoverability of images in the digital realm, but it does not allow for standardization of search results, a necessity for researchers who understandably need all examples of a certain subject or theme, not a random selection. Fortunately, the library came to the attention of a team of deep learning researchers who were intrigued by its digitized research photograph collection as well

19. "Agents for the Frick Art Reference Library," Frick Art Reference Library, https://www.frick.org/research/photoarchive/acquisitions/agents.

20. Accessible at https://digitalcollections.frick.org/. The Frick Digital Collections incorporates the standards and technologies of the International Image Interoperability Framework (IIIF); thus, both images and metadata are provided to benefit comparative art historical study.

as the challenges of maintaining this collection and were willing to explore ways to improve the staff's workflow.[21]

THE COLLABORATION

Dr. David Donoho, a professor at Stanford University,[22] and his students Vardan Papyan (formerly a post-doctoral researcher at Stanford and presently an assistant professor at the University of Toronto) and X.Y. Han (a PhD student) are researching deep learning networks and are always on the hunt for new image sets for their work. Until their collaboration with FARL, they had been using popular image databases such as ImageNet, which presently contains more than fourteen million images, or smaller, more specialized datasets such as CIFAR10, CIFAR100,[23] FashionMNIST,[24] MNIST,[25] SVHN,[26] and STL10.[27] Dr. Donoho and his students were introduced to the library's Photoarchive and its staff by Dr. C. Richard Johnson, Jr., a Cornell University professor[28] and pioneer in computational art history who has spent much of his career connecting scientific researchers with the art history community.[29] During their first meeting, which was arranged by Dr. Johnson, both parties identified the opportunities the Photoarchive's digitized image library could offer Dr. Donoho and his students. Thus began the partnership between the library and the deep learning researchers to develop automatic image classifiers for the Photoarchive's holdings. For the photoarchivists, the potential of having their workflow streamlined through the automation of a necessary but time-consuming aspect of cataloging was certainly appealing. For the Cornell/Toronto/Stanford team, the possibility of working with a unique dataset consisting of works of art—specifically, paintings and drawings—and labeled according to an in-house classification system

21. "Deep learning" refers to the sub-field of ML studying the deep neural network algorithm.

22. Dr. Donoho is Professor of Statistics and the Anne T. and Robert M. Bass Professor of Humanities and Sciences at Stanford University.

23. Alex Krizhevsky and Geoffrey Hinton, "Learning Multiple Layers of Features from Tiny Images," MSc thesis/technical report, University of Toronto (2009), https://www.cs.toronto.edu/~kriz/learning-features-2009-TR.pdf.

24. Han Xiao, Kashif Rasul, and Roland Vollgraf, "Fashion-MNIST: A Novel Image Dataset for Benchmarking Machine Learning Algorithms," *arXiv preprint arXiv:1708.07747* (2017), https://arxiv.org/abs/1708.07747.

25. Li Deng, "The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web]," *IEEE Signal Processing Magazine* 29, no. 6 (2012): 141–42.

26. Yuval Netzer, Tao Wang, Adam Coates, Alessandro Bissacco, Bo Wu, and Andrew Y. Ng, "The Street View House Numbers (SVHN) Dataset" (2019), http://ufldl.stanford.edu/housenumbers/.

27. Adam Coates, Andrew Y. Ng, and Honglak Lee, "An Analysis of Single-Layer Networks in Unsupervised Feature Learning," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, ed. Geoffrey Gordon, David Dunson, and Miroslav Dudík (2011), 215–23, http://proceedings.mlr.press/v15/coates11a/coates11a.pdf.

28. Dr. Johnson is professor of electrical and computer engineering and the Geoffrey S.M. Hedrick Sr. Professor of Engineering at Cornell University, and the senior research advisor to the Frick Art Reference Library.

29. See C. Richard Johnson, Jr., Ella Hendriks, Igor J. Berezhnoy, Eugene Brevdo, Shannon M. Hughes, Ingrid Daubechies, Jia Li, Eric Postma, and James Z. Wang, "Image Processing for Artist Identification," *IEEE Signal Processing Magazine* 25, no. 4 (2008): 37–48; C. Richard Johnson, Jr. and W.A. Sethares, "Canvas Weave Match Supports Designation of Vermeer's Geographer and Astronomer as a Pendant Pair," *Journal of Historians of Netherlandish Art* 9, no. 1 (2017); William A. Sethares, Margaret Holben Ellis, and C. Richard Johnson, Jr., "Computational Watermark Enhancement in Leonardo's Codex Leicester," *Journal of the American Institute for Conservation* 59, no. 2 (2020): 87–96.

rather than relying on stock images culled from the internet, offered new intellectual challenges, allowing them to expand the scope of their research.

APPLYING AUTOMATIC IMAGE CLASSIFIERS

The cornerstone of the pilot project is built upon a deep neural network algorithm for image classification. Deep neural networks,[30] which were originally inspired by the human brain, pass images represented as numerical signals through layers, each consisting of numerous computer-implemented neurons, until it outputs a prediction of the category to which the image belongs. When a signal is passed between two neurons, it undergoes a mathematical transformation much like how biological neurons are electrostatically activated to different degrees in the human brain. The activations between neurons in a neural network are determined by a set of parameters associated with the network architecture, and these parameters number in the millions because, also like in the human brain, there are an exponentially large number of possible connections. Because of this, a prerequisite of the algorithm's efficacy is the existence of a large collection of pre-labeled training images that researchers can use to "tune" these parameters—and the larger the collection, the better. This tuning is what is meant when a deep learning researcher says a network is "being trained" or "learning." Training datasets typically contain at least a few tens of thousands of images even for the simplest tasks.

In domains in which such large training sets exist, deep neural networks have demonstrated near human-level performance that has led to significant technological leaps. For example, deep nets have been used to great success by doctors to identify cancer cells[31] and by physicists to discover new subatomic particles.[32] In this light, the Frick Art Reference Library's Photoarchive with its tens of thousands of already-labeled images (that is, the library's "classified mounts" discussed above) and hundreds of thousands of images that require labeling (the library's "supply photographs" described above) appeared to Dr. Donoho and his students to be perfectly positioned to benefit from these advancements.

Moreover, the Photoarchive presents an interesting intellectual challenge since its dataset differs from conventional deep learning classification tasks in two ways. First, each classification heading consists of a series of constituent descriptors. For example, the 1840 portrait of George Washington by Thomas Wilcocks Sully presently in The Metropolitan Museum of Art[33] is classified in the archive as "121–15" or "American School: Portraits: Men: Without hands: (without hats): Head to left." In contrast, most modern deep classification networks have been developed to identify single-word

30. Ian Goodfellow, Yoshua Bengio, and Aaron Courville, *Deep Learning* (Cambridge: MIT Press, 2016), http://www.deeplearningbook.org.

31. Dayong Wang, Aditya Khosla, Rishab Gargeya, Humayun Irshad, and Andrew H. Beck, "Deep Learning for Identifying Metastatic Breast Cancer," *arXiv preprint arXiv:1606.05718* (2016), https://arxiv.org/abs/1606.05718.

32. Pierre Baldi, Peter Sadowski, and Daniel Whiteson, "Searching for Exotic Particles in High-Energy Physics with Deep Learning," *Nature Communications* 5, no. 1 (2014): 1–9.

33. Accession number 54.74.

descriptors such as "man." Second, the Photoarchive classification headings follow a hierarchical structure. For example, the descriptor "Head to left" is only considered after the components "Portraits" and "Men" already apply. Adapting to this problem, the Cornell/Toronto/Stanford team re-engineered the network to incorporate the tree-like labelling system into the network's architecture itself as well as designed a new performance metric against which the network was tuned. This process required the synergy of two very different fields of expertise from the art historians at the library and the deep learning researchers. The photoarchivists developed the decision tree, which mimics how an art historian would read an image (i.e., "this is a bust-length portrait of a man without a hat") and the deep learning researchers mapped this decision tree to the neural network.

In the latest version, the Cornell/Toronto/Stanford team modified the popular ResNet152 network[34]—with 152 layers and six million parameters—and trained it on a collection of 45,857 classified reproductions of American paintings provided by the Frick Art Reference Library's Photoarchive. As noted above, these images derived from the North American photograph campaigns completed between 1922 and 1967 and were the first to be digitized by the library due to concerns regarding copyright. Thus, all classification headings in the test group began with "1," which denotes the American School. The algorithm, however, supplied the remainder of the heading, which is based on subject matter. Reflecting the origins of this collection of training images, the majority of these paintings were portraits and represented ninety-eight unique classification headings. The training took place on Stanford University's Sherlock supercomputer cluster during a period of three days.

After training, the network was fed images from the unlabeled portion of the library's Photoarchive, and the network predicted a classification heading for each one. These images were then annotated with the predictions and shown to the library's photoarchivists through an application developed using the popular crowd-sourcing platform, Zooniverse (Figure 8).[35] When developing this app, the Cornell/Toronto/Stanford team sought to produce an intuitive, even familiar interface that would allow archivists to review scores of images quickly and easily; thus, the program mimicked the notorious dating app Tinder. Library staff downloaded it on their computers or phones (Figure 9) and reviewed the algorithm's predictions. If the classification heading applied to an image was correct, it was considered a "match," and the staff member swiped right. If it was incorrect, staff swiped left and the image was sent to a folder for review. In testing the latest version of the network, library staff vetted 8,661 images in the year August 2019–August 2020 and agreed with the network on the classification headings applied to 5,829 (67 percent) of the images. Yet, even the incorrect predictions were, in general, "almost correct" with only one tag incorrect or missing. For example, the algorithm applied the heading "24–3: Portrait groups: Women" when it should have applied the heading "24–6: Portrait Groups: Women and Children." The Cornell/Toronto/Stanford team is confident

34. Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016): 770–78.

35. Zooniverse, https://www.zooniverse.org/. As of January 2021, this project is not accessible to the public.

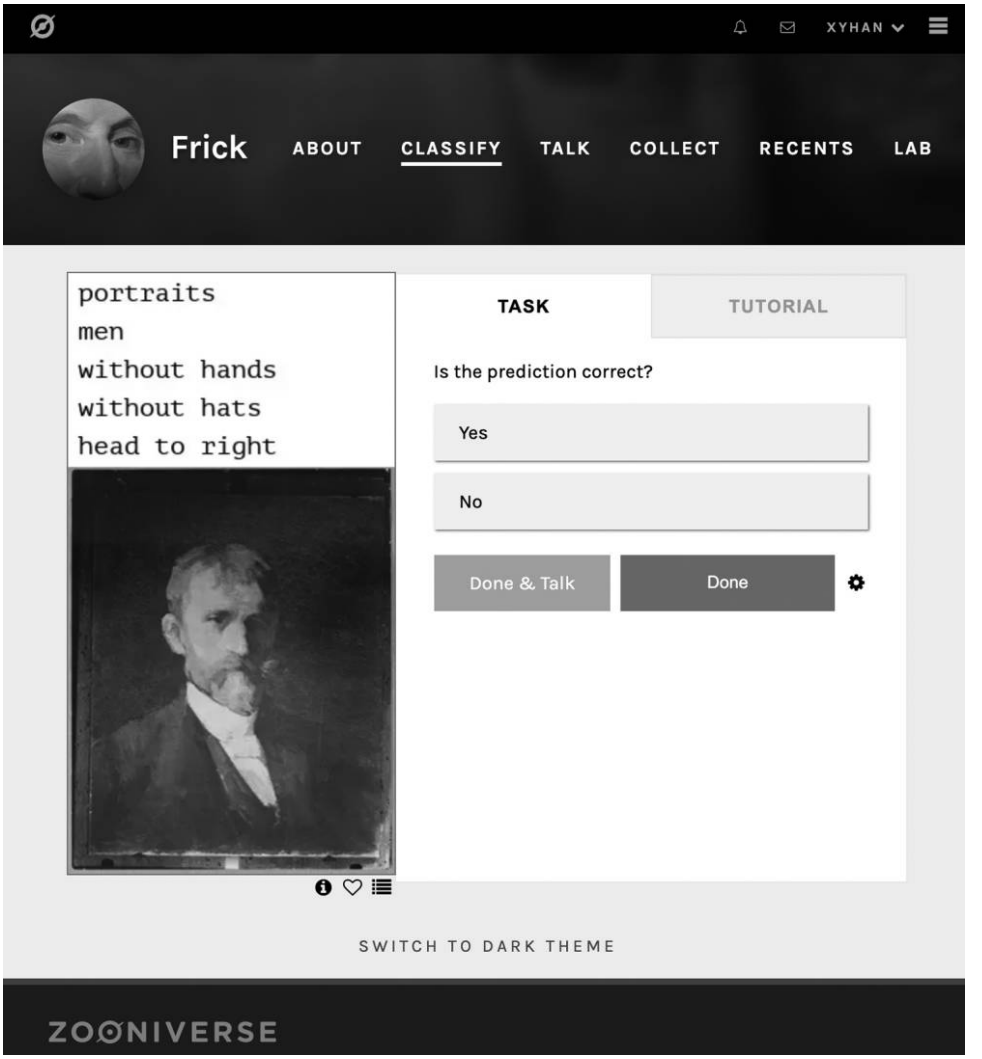


Figure 8. Screenshot of the application developed using the crowd-sourcing platform Zooniverse to vet the algorithm's predictions as it appears on a computer desktop. Courtesy of the authors. Please see the online edition of *Art Documentation* for a color version of this image.

that the network can learn to identify these cases when given more training examples of the headings on which it erred.

THE FUTURE OF THE COLLABORATION

In the spring of 2020, an additional dataset of approximately 210,000 raw image files (that is, the images only) was sent to the Cornell/Toronto/Stanford team for additional testing. The goal is to refine the algorithm so that it achieves 93 percent accuracy when applying the in-house classification headings. Weighing the computational cost of further training the deep neural network against the cost-of-human-labor of photoarchivists

Figure 9. Screenshot of the application developed using the crowd-sourcing platform Zooniverse to vet the algorithm's predictions as it appears on a smartphone. Courtesy of the authors. Please see the online edition of *Art Documentation* for a color version of this image.

to create training images, the Cornell/Toronto/Stanford team determined that this percentage is a "sweet spot" for maximum efficiency.

This percentage is sufficient for the library's needs as it mimics human labeling. Even when trained art historians apply classification headings, mistakes can occur. A photoarchivist might mistype a number, thereby transforming an equestrian portrait of a man ("21–3: Portraits: Men: Equestrian") into a portrait of a female literary character such as Hester Prynne ("22–3: Portraits: Women: Imaginary"), or temporarily forget that the library's classification system distinguishes between still-life scenes of game only ("16–1: Still life: Dead game and animals") and those game pieces that contain a live animal such as a salivating hound ("16–3: Still life: With living animals"). Certainly education, training, and personal preferences can complicate efficient labeling by humans. For example, there have been intense debates among library staff regarding the correct classification heading for a scene of fashionable young people enjoying an afternoon in a crowded park. There are at least three possible headings offered by the in-house system: "20–17: Genre: Fêtes"; "20–24: Genre: Outdoor scenes"; or, due to the crowd of peasants in the background, "20–37: Genre: Kermesses," and each one has its partisans. Although the algorithm will not be able to resolve iconographical disputes among staff, it does have the potential to motivate library archivists to explore the issue of bias in their recording practices, a possible future project for FARL and the Cornell/Toronto/Stanford team. Certainly, the algorithm can help to standardize the application of certain headings, thereby increasing the discoverability of works of art represented in the Photoarchive.

This standardization itself poses a set of philosophical and technological questions. In particular, the network learns not to *reason* about whether an image is "20–17: Genre: Fêtes" or "20–24: Genre: Outdoor scenes," but rather learns to *mimic* decisions it has seen in the training examples. Intuitively, the network effectively learns to answer the following questions: Of all the *previously annotated* examples it has seen, what are the most similar images to the current image-of-interest? And, among those, which label was applied? Such mimicry in ML is sometimes referred to as "human imitative AI" and raises two open questions.[36] First, while the network has the potential to standardize cataloging decisions, such decisions are not objective. They will reflect an aggregate of the views of the annotators of the training images—with more prolific annotators having more influence. Thus, creating a classifying network that is at the same time objective and disinterested requires additional collaboration between the library and the Cornell/Toronto/Stanford team to curate training datasets that contain a representative balance of trustworthy expert decisions. Second, this trait implies that the network's predictive ability will degrade when presented with images that are noticeably different from those in the training dataset. The Cornell/Toronto/Stanford team indeed found this to be the case. The neural network, trained on (generally) photo-realistic paintings from the library's collection of North American portraits tends to output lower-quality labels when applied to

36. See Michael I. Jordan, "Artificial Intelligence—The Revolution Hasn't Happened Yet," *Harvard Data Science Review* 1, no. 1 (2019), https://hdsr.mitpress.mit.edu/pub/wot7mkc1/release/8, and David L. Donoho's response, "Comments on Michael Jordan's Essay 'The AI Revolution Hasn't Happened Yet,'" *Harvard Data Science Review* 1, no. 1 (2019), https://hdsr.mitpress.mit.edu/pub/rim3pvdw/release/5.

more abstract drawings and sketches. Therefore, engineering the network to be resistant to shifts in the artistic style of its inputs is a current avenue of technical investigation for the Cornell/Toronto/Stanford team.

With these considerations in mind, the Cornell/Toronto/Stanford team continues to develop the network and improve its performance. In 2021, the Cornell/Toronto/Stanford team plans to put the algorithm to work and apply classification headings to the 155,000 supply photographs that are scheduled to be digitized by library staff during the calendar year. The team's ambition is that Photoarchive images will be labeled with the in-house classification system in tandem with image capturing and therefore can be released immediately to the public with minimal information. Although these images will not be accompanied by extensive documentation, researchers will be able to locate images in the digital archive through the word-searchable classification headings as well as by artist and title. Because of the information added to the back of each supply photograph—which is also digitized—they will also be able to determine the date of the work of art and its current location. Perhaps the library might once again explore crowdsourcing as a possible means to enhance the online catalog, developing a new project on Zooniverse that would allow volunteers to transpose the information recorded on the backs of the photographs to the online catalog. This would be a project focusing on transcription rather than tagging, and therefore the results would not require time-consuming vetting by library staff. Thus, by employing AI to streamline metadata creation, the Frick Art Reference Library will be able to capitalize on its in-house classification system to increase access to and discoverability of the library's resources.

CONCLUSION

The pilot project launched by FARL and the Cornell/Toronto/Stanford team provides one model for the effective use of AI in the digitized image archive, offering librarians and archivists a partial solution to managing their workloads while remaining sensitive to the ethical issues of applying automation in the formation of the catalog. It is critical to note that the algorithm developed by the Cornell/Toronto/Stanford team does not seek to improve upon or supplant the archive's existing classification system; rather, its purpose is to improve the efficiency of Photoarchive staff. The deep learning researchers worked closely with library staff at each stage of the pilot project and relied on the photoarchivists' expertise during the vetting process to ensure that the algorithm would mimic their decision-making process.[37] By producing useful metadata, the algorithm affords scholars and students increased access to the library's research collections and introduces additional possibilities for digital art history projects focusing on genre and iconography. Taking advantage of breakthroughs in AI and ML has thus made the Photoarchive a more valuable resource by expanding knowledge of its holdings and potential.

37. Significantly, the CMU team also recommends employing an "expert-in-the-loop" process to prevent error and bias as caused by the training set. Determining that "computer vision by itself, or even with non-expert human guidance, would actively impede the public use of a photo archive," the team concludes that the technology should be integrated with a digital asset management system so that "expert archivists and editors can strategically leverage machine learning models without making the collection beholden to them." See Corrin et al., "CAMPI," 3.

Interestingly, this is not the first time in its history that the library has courted controversy by exploring how technology might improve art historical research. The establishment of the Photoarchive itself was viewed with suspicion by many art historians who considered the use of photographs for art historical study problematic. For example, Richard Krautheimer censured art historians who maintained that "old or new photos, taken by friends, could replace actual inspection."[38] Krautheimer and other like-minded scholars questioned the objectivity of the photograph, lamenting the fact that issues of scale, context, and viewer engagement were elided by these deceptive documents that privilege the perspective of the photographer rather than the artist. Obviously, Helen Clay Frick did not share their reservations and embraced not only photography but also recent advances in the field of art conservation. She and her team sought out x-rays and early infrared photographs of works of art, exploiting these technological breakthroughs in conservation to expand understanding of materials, technique, and stylistic formation. Thus, the collaboration between FARL and the Cornell/Toronto/Stanford team should not be viewed as an innovation but instead, as the natural continuation of the library's legacy.

ACKNOWLEDGMENTS

The Cornell/Toronto/Stanford team thanks Anastasia Levadas, John McQuaid, and Louisa Wood Ruby for their warm welcome and acceptance of a group of mathematicians and engineers into their community. Specifically, we thank Anastasia for patiently answering the Cornell/Toronto/Stanford team's numerous questions regarding the Photoarchive dataset, often providing insightful art historic contexts; John for his enthusiastic help translating and developing the library's in-house classification system into tree-structure that is amenable to a neural network; Anastasia and John for organizing the testing of the Zooniverse app by library staff and interns, thus allowing us to collect almost 9,000 annotations as of August 2020; and, of those annotations, we specially acknowledge Sarah Bigler and Kerri Pfister, who contributed 3,947 and 1,819 entries, respectively, as of August 2020.

38. Quoted in Costanza Caraffa, "From 'Photo Libraries' to 'Photo Archives.' On the Epistemological Potential of Art-Historical Photo Collections," in *Photo Archives and the Photographic Memory of Art History*, ed. Costanza Caraffa (Berlin: Deutscher Kunstverlag, 2011), 16.