Using Computer Vision to Increase the Research Potential of Photo Archives

by John Resig

Journal of Digital Humanities, Vol 3 Issue 2, Summer 2014

Using Computer Vision to Increase the Research Potential of Photo Archives

John Resig

In art history research, photographs of art are the lifeblood of study. Since it's usually impossible for a scholar to travel the globe and visit an artwork as need arises, there is substantial demand for archives of photographs of artworks for reference and study. There are photo archives around the world with millions of photographs in them, including the prestigious Frick Photoarchive. These archives aggregate photographs from many institutions and private collections. It is their job to make sure the photos are maintained and the works of art they document cataloged, changes in attribution or ownership updated, and that they have properly identified and merged duplicate photographs and entries relating to a single artwork.

This process of finding duplicate artworks can be breathtakingly time-consuming. Many professional researchers spend years correcting and merging entries in even a moderately-sized archive. For an archive with over a million photographs, that process becomes impossible. This says nothing of the difficulty of sharing images between institutions where cataloging standards or metadata may differ drastically.

Image similarity analysis is an exciting computer vision technique for matching photos whose image content is substantially or completely similar. Through image similarity analysis, it is highly likely images depicting the same object will be found and matched.

The application of computer vision to art photo archives has largely been unexplored up to this point. Lev Manovich has explored ways of analyzing images of artworks while looking for trends in an artists oeuvre or entire artistic movements. However, most institutions have used large scale image analysis primarily for cases of copyright enforcement, face detection, or color/composition analysis.

To explore what image similarity analysis was capable of, I completed an analysis of the digital images of Italian anonymous art at the Frick Photoarchive. The image similarity analysis, using TinEye's MatchEngine service, was automated using newly-developed tools. I further processed and dissected the data using custom tools. The analysis was able to confirm some of the existing relationships between photographs that were manually generated by researchers. The analysis was also able to discover a number of completely new relationships, including: works of art before and after conservation, copies of the same artwork, cropped detail shots of the same artwork, and cataloging errors.

The custom toolkit developed to analyze the Italian anonymous archive will be publicly released as a generic image similarity analysis tool. Comparable results could be easily achieved by other institutions for a minimal cost using these tools.

The results of the image similarity analysis of a photo archive are extremely exciting and could completely change how the process of cataloging images is completed. It could also make some impossible tasks, such as merging multimillion image archives, a reality.

The Frick Photoarchive

Started in 1920, The Frick Photoarchive has continually expanded over nearly a century and now contains over 1.2 million photographs of works of art. In addition to sponsoring original photography of art around the world, the Frick has benefited from photograph donations from both institutions and scholars. To this day the library still actively purchases photographs.

The Frick Art Reference Library recently contacted me when they saw image analysis work that I was doing with my Ukiyo-e.org: Japanese Woodblock Print Search and Database project (which deserves a separate essay). They were curious if image analysis could work for photographs of paintings, three-dimensional artworks (instead of prints), and their collections in particular. Additionally, they were interested in where image analysis could aide in the process of merging multiple photo archives

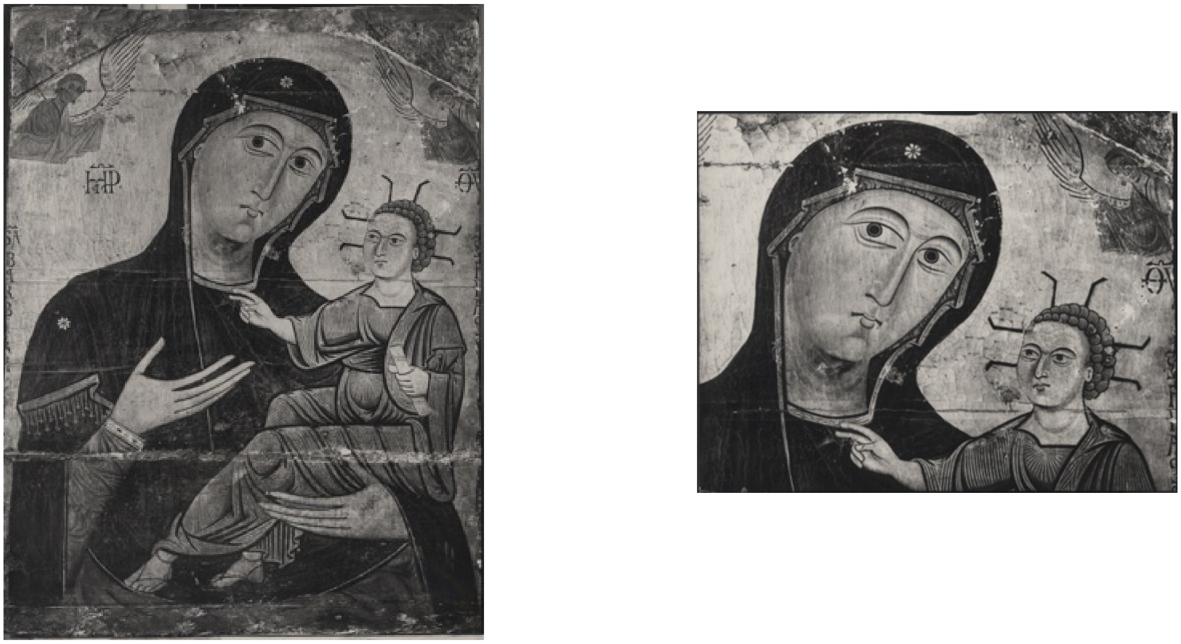
The Frick Library is a member of the newly-formed International Digital Photo Archive Initiative, a consortium of fourteen photo archives from Europe and the United States with an aggregate 31 million photos of art. Nearly all of these institutions are in the process of digitizing their photo archives. They see the tremendous power of sharing photos and photo metadata amongst institutions: the aggregated information can yield a better understanding of the artworks (works before and after conservation, works that have been stolen or are missing can be revealed, and provenance and general scholarship can be accelerated).

The Frick Library is still early on in the digitization of their collection. Thus far, they've digitized about 70,000 photographs. Their in-house digitization lab has just recently been set up and will allow for a far greater volume of photos and increases in metadata quality. They've also received grants to digitize their collection of 57,000 original negatives of artworks, most of which is already available online in the Frick Digital Image Archive.

Frick Italian Anonymous Digital Archive

The first digitization project undertaken by the Frick Photoarchive, sponsored by the Pernigotti S.p.A., Averna Group in Milan, was to digitize 18,548 photographic reproductions of 14,284 works of anonymous Italian art and turn it in to a digital photo archive. This photo archive is made available to researchers through the Frick Digital Image Archive. The digitization was undertaken by an outside lab long before the Frick Photoarchive had its in-house digitization lab set up.

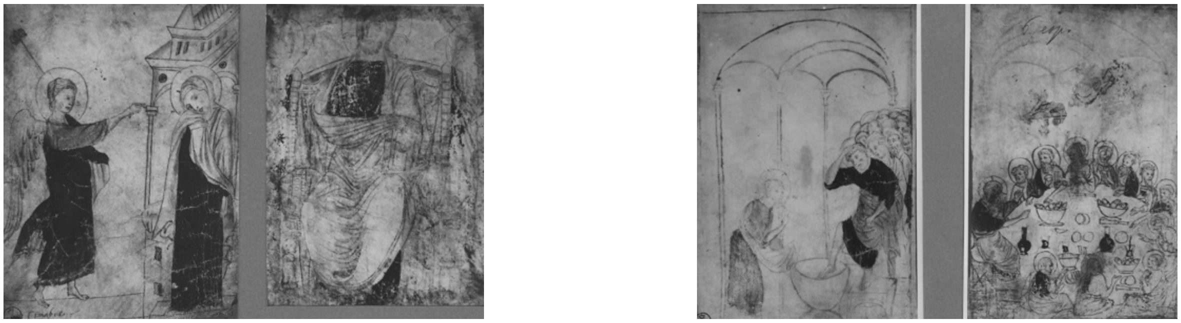
The artworks represented in the Italian anonymous archive are largely from around the time of the Renaissance and are either unattributed or considered to be anonymous. The archive is not limited to just two-dimensional paintings, but also includes frescos, drawings, prints, and sculpture. A representative example of the artworks and photos in the archive is shown below:



Madonna and Child, 13th century, La Chiesa di S. Eufrasia, Pisa.

In the case of this artwork, there are two separate photos representing the same piece: the full panel and a close-up detail shot. Note that the photos are in black-and-white: this is the case for nearly all the photos in this particular archive.

In the Italian anonymous digital archive, the photos are generally organized into groups with all photos from the same work of art clustered together under a single number (for example `10383a.jpg`, `10383b.jpg`, `10383c.jpg`, etc.). This clustering was done manually by the original digitization team using metadata associated with the photos in the archive. However just because the photos are of the same work of art does not guarantee that they'll be depict an image that is identifiably the same work of art. For example the following two photos depict different portions of the same work of art with no overlapping imagery:



Florentine, 13th century, Uffizi Museum in Florence

The Italian anonymous archive poses particular challenges to researchers at the Frick Photoarchive. Most of the photos in the photo archive are organized by attributed artist, making it easy to find duplicate, or alternate, photos of the same work of art. The fact that none of the works in this particular archive are attributed makes it extremely hard to guarantee that every alternate photo of an artwork will be grouped together.

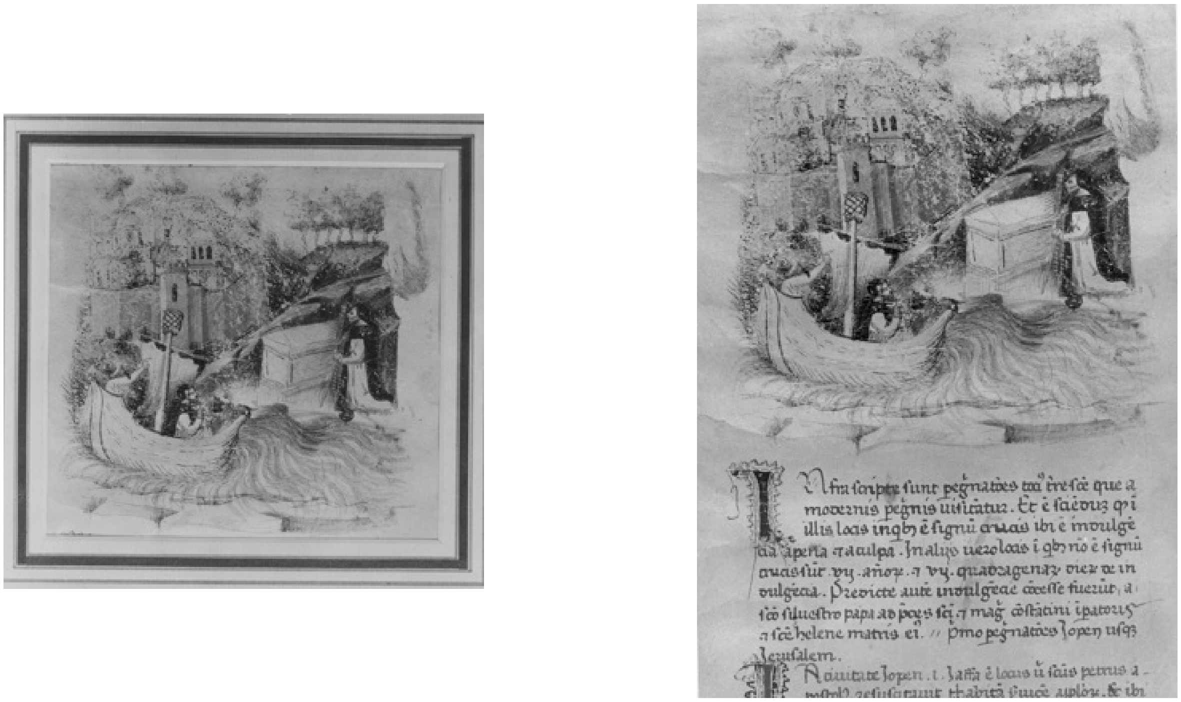
Interestingly, the problem of grouping related art photographs is actually quite similar to the problem of grouping images across multiple major (and sometimes international) photo archives. If one were given two sets of images, each with thousands (or millions!) of images in them, it would be physically impractical for humans to go through all of the entries for a particular artist and cluster every identical work of art. When faced with a problem of this magnitude the smart thing to do would be to turn to the metadata associated with the images to support the merging.

To appropriate a famous quote from the programmer Jamie Zawinski:

> Some people, when confronted with a problem, think
> "I know, I'll use metadata." Now they have two problems.

In theory good metadata attached to records should be able to solve most problems that come with merging or correcting problems in a collection (or between collections). However, in practice, it's very likely that institutions will have varying interpretations of quality, make mistakes in cataloging, and make mistakes in data entry. When merging multiple collections whose metadata is written in different languages or between collections that are missing critically important metadata (as is the case with the Italian anonymous archive's missing artist names), the challenge becomes even more difficult.

This is where the effectiveness of computer vision and using image analysis to correct archives becomes crucial. Accurately matching two images that have identical visual characteristics in two different collections can reveal missing or mistaken data. As a representative example two images found to be similar through the analysis are shown below: one is a photo from a Christie's auction catalog dating to 1936 and the other is a photo from the Harvard Art Museum in Cambridge.



Tuscan, 15th century, Harvard Art Museum.

Naturally the artist is unknown in both of these cases, but it's very possible to have found a match after the fact if the metadata was good enough. Unfortunately, for these two images that was not the case. For whatever reason, the Harvard Art Museum fails to mention that this piece came from an auction at Christie's (or that the owner who donated it had purchased it at Christie's). Given that there is no identifiable artist, title, or date of this piece, it thus makes it incredibly unlikely that a human would've been able to discover that these two images were of the same work of art.

Put simply: there is frequently not enough information for humans to intervene and make a connection between images in a scalable manner. Individual researchers can certainly hunt through photos that have been organized (hopefully correctly) by artist or national school and century and attempt to make the associations manually. However, this process is painstaking at best and does not work well across hundreds, or thousands, of artists and potentially millions of images.

If all metadata associated with an image is ignored, and only the contents of the image were analyzed, it becomes possible to find interesting image matches that were likely undiscoverable using raw human power. A computer vision image analysis algorithm that's capable of finding matches between images that have a set of identical content would be the perfect tool for performing the analysis. With such a tool any matches that occur would likely indicate that they are different images of the same artwork.

It's possible that some researchers may become skittish at the prospect of ignoring all the painstakingly-generated metadata that's been associated with their images (for the purpose of finding similar images, at least). However, it's important to note that images rarely lie. When they do, there's likely something interesting happening that would be a good area for further research (such as copies of the same work of art).

Image Similarity Analysis Implementations

Computer Science research into computer vision and image comparison techniques has been going on for decades. Research and implementation is finally at the point where image analysis can be performed against millions of images simultaneously (as can be seen in the services provided by Google, Yahoo, and Bing Image Search). The general availability of this technology however, has been mixed. There are some freely available, open source, tools such as imgSeek and libpuzzlea>, which bring rudimentary image comparison technology to a larger audience. There are also commercially-available tools that provide fast image analysis with a greater level of clarity, such as TinEye's MatchEngine.

Finding the right tool that would work for the print images that were collected from the various institutions was especially tricky. The features needed for an effective print image search are:

- The process of adding in a new image, and performing a search with an image, must be fast. (If searches and comparison are too slow it'll be too hard to use effectively.)
- The engine should be capable of scaling up to hundreds of thousands, if not millions, of images.
- The engine should be able to find exact matches (cases where an artwork is definitively contained within an image). Inexact matches tend to confuse the results and make the matches hard to discern.
- The engine must be able to ignore differences in color, even differences between a color photograph and a black-and-white photograph. (Many institutions provide images only in black-and-white. Comparing those images with color matches at other institutions would be very useful.)
- It must be possible for an image of an artwork detail (part of a larger artwork) to match an image of the complete artwork.
- Images that have watermarks or other invasive imagery should still be matched (and not only match other images that also have watermarks).

Initially, imgSeek was explored because it did direct image comparison, worked quickly, and was open source. However, there were many difficulties in its practical use. imgSeek only analyzes pieces of an image (the colors and where those colors are located in the image), which causes similarly-composed images to appear as matches, even though they may be entirely different. For example, an image of blue sky with green grass would match all images that were blue at the top and green at the bottom, rather than just images of sky and grass. Additionally, it's unable to effectively find images that are in black-and-white or match details to a complete image of an artwork.

The MatchEngine tool, while a commercial service, is much better suited for finding images that are exact matches of one another or even details embedded inside a larger image. In all of the testing, MatchEngine outperformed the imgSeek service in quality. MatchEngine was much better at finding exact matches, ignoring differences in color, and finding details inside images.[1]

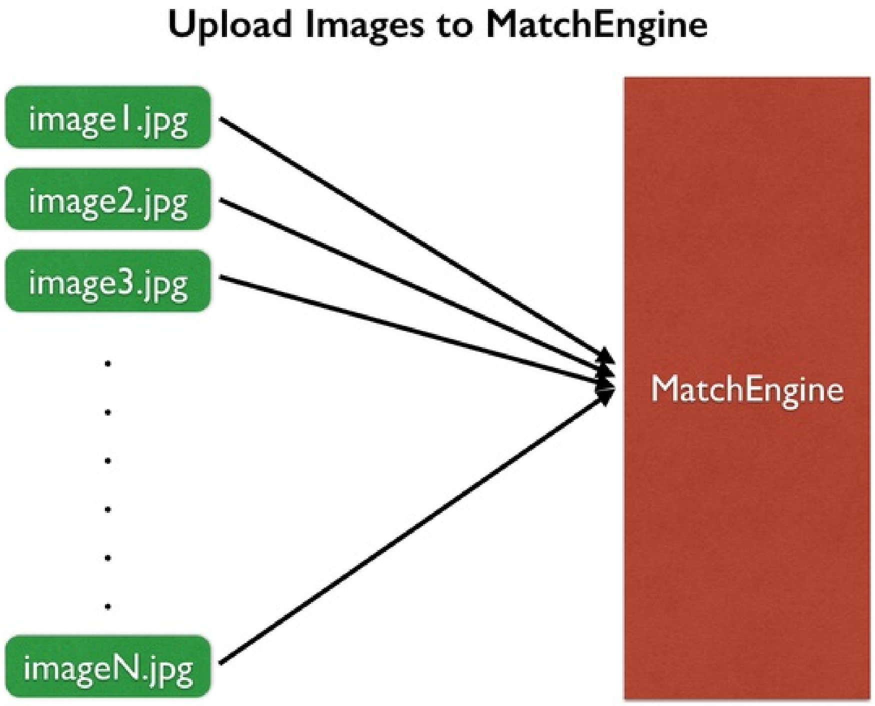
Implementation

With an image analysis utility in place, it is now possible to create a tool for automatically finding interesting new matches, correcting cataloging mistakes, and validating some of our existing matches.

The Frick Photoarchive provided an export of the 18,548 images in the Italian anonymous archive. MatchEngine will automatically scale down any image that is over 300 pixels tall or wide. Thus, to simplify the transfer, the Frick Photoarchive reduced the size of all the images before passing them along. In total, the size of these images was about 2 Gigabytes. Additionally, the Frick Photoarchive provided a CSV dump of all of the metadata associated with the images.

A number of tools were developed to perform the image analysis, collect the data, and analyze the results of the analysis.

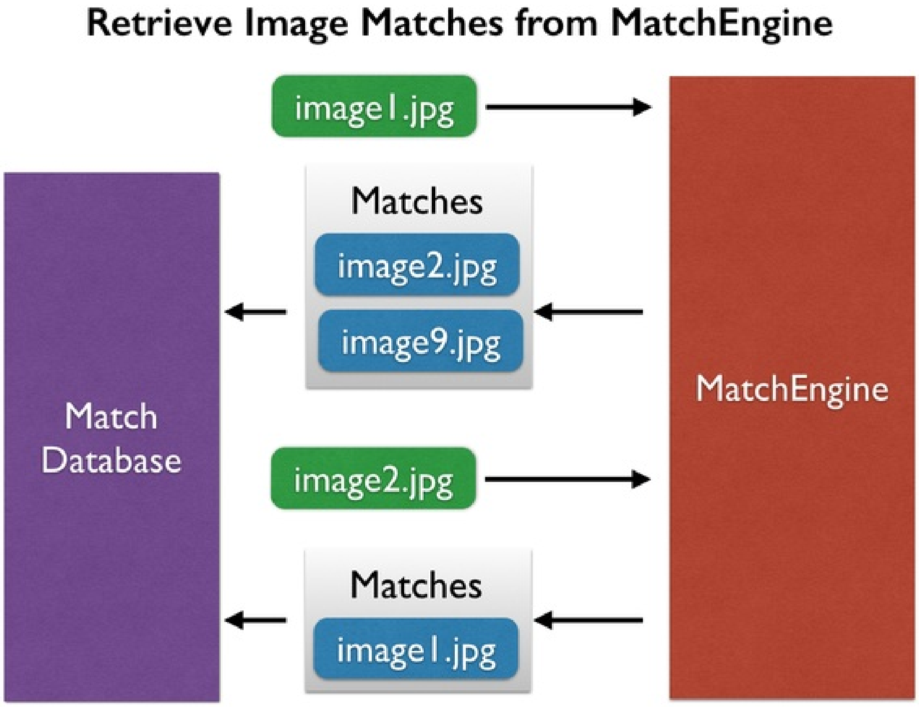
The first tool was a utility for uploading all of the images to the MatchEngine service through their private REST API.



The MatchEngine API supports uploading up to 1,000 images simultaneously. While the uploading is occurring, no other operations can be performed with the API. For the 18,548 Italian anonymous images, it took about 3 hours to complete over a standard home cable Internet connection.

Conventionally the MatchEngine service is used for two purposes: 1) providing a list of similar images for every uploaded image and 2) allowing a user to search images by uploading a photograph. Normally most users of MatchEngine keep images in the service over a long period of time to handle user search queries. For the analysis performed on the Italian anonymous archive, there was no need to keep the images in the MatchEngine service for any significant duration: only a bulk list of the similarities between the uploaded images was needed.

Another tool was then built to query MatchEngine for every previously-uploaded image to determine if any similar images had been found. MatchEngine's indexing of the images was performed immediately upon upload and was made available for querying. Thus every single uploaded image could be queried and a full relationship graph could be downloaded.



The MatchEngine results for an image may look something like this:

```
&quot;frick-anon-italian/13291.jpg&quot;: [
    {
        &quot;score&quot;: &quot;27.80&quot;,
        &quot;target_overlap_percent&quot;: &quot;100.00&q
        &quot;overlay&quot;: &quot;...&quot;,
        &quot;query_overlap_percent&quot;: &quot;47.18&quo
        &quot;filepath&quot;: &quot;frick-anon-italian/132
    },
    {
        &quot;score&quot;: &quot;12.50&quot;,
        &quot;target_overlap_percent&quot;: &quot;100.00&q
        &quot;overlay&quot;: &quot;...&quot;,
        &quot;query_overlap_percent&quot;: &quot;20.93&quo
```

```
            &quot;filepath&quot;: &quot;frick-anon-italian/132
        }
]
```

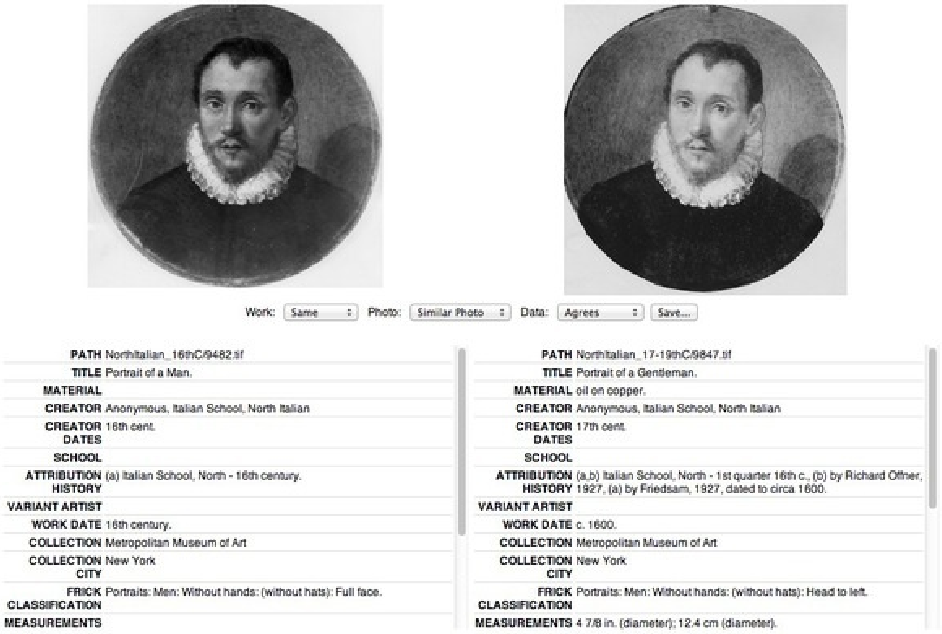
In this case, a query with the image `13291.jpg` received matches for the images `13291b.jpg` and `13291a.jpg` (I anticipated this result: all of these images were previously cataloged as being the same work of art depicted in alternate photographs or detail shots). The results show the “score” of the result, as specified by MatchEngine. The score represents how closely two images are deemed to be related. In practice, even very low-scoring images still appear to be the same work of art. MatchEngine also provides data regarding how much of the images were overlapping and provides some details on how to line up the images with one another; however, none of that is needed for this particular analysis.

The MatchEngine similarity data can be downloaded in parallel (using up to four simultaneous API connections). On a home cable Internet connection it took about an hour to retrieve all of the image similarity data for the entire Italian anonymous archive. All of the similarity data was then cached in a local JSON file for later retrieval. At this point the MatchEngine service was no longer needed or used. All of the images could then be deleted, using the API, from the MatchEngine servers.

Once all the image similarity matches have been downloaded to a local data store, the next step is to review all of the results and categorize the newly-matched results (this step is only performed for any previously unknown matches). The categorization of the matches isn’t completely necessary: the matches could be passed off directly to researchers and catalogers instead. However, performing a basic organization of the results could help optimize researcher effort and focus attention on particular results or problem areas.

With a result categorization tool I was able to easily categorize all of the image matches. This could easily be achieved by other non-experts, or at least by people who have a basic familiarity with the subject matter being depicted in the images.

The categorization tool provides the user with a view of the two images that were matched by MatchEngine paired together with the raw data provided in the CSV data dump.



Work: Same   Photo: Similar Photo   Data: Agrees   Save...

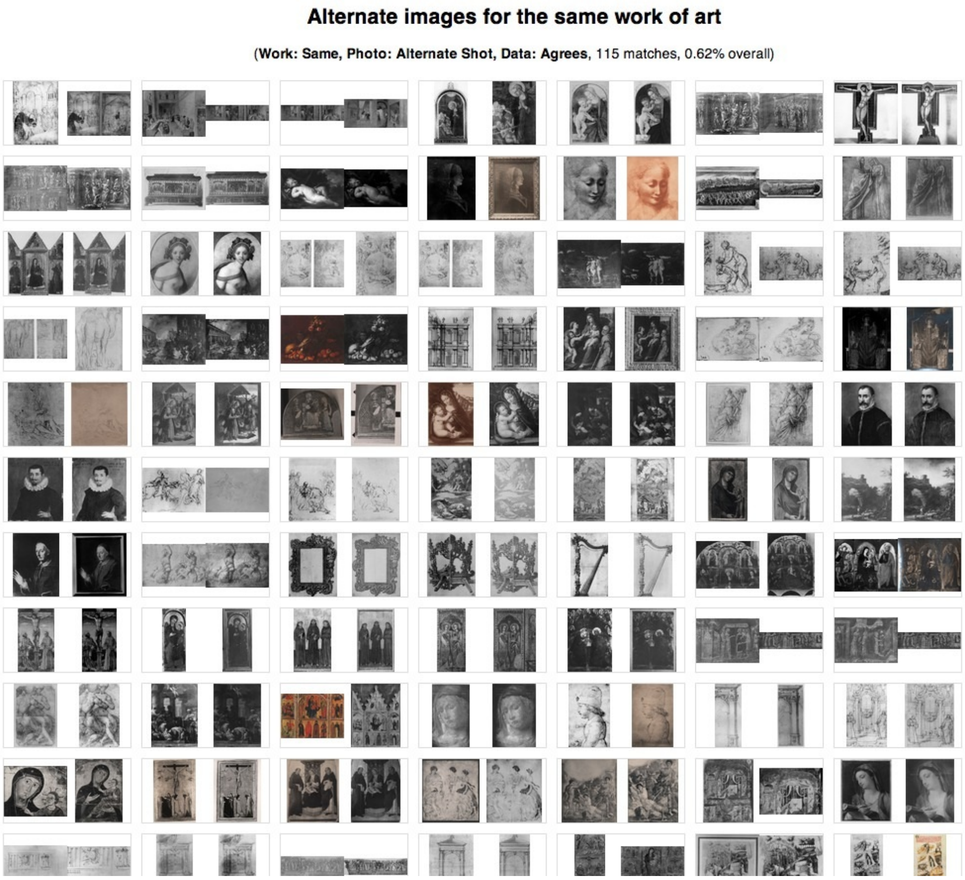
PATH	NorthItalian_16thC/9482.tif
TITLE	Portrait of a Man.
MATERIAL	
CREATOR	Anonymous, Italian School, North Italian
CREATOR DATES	16th cent.
SCHOOL	
ATTRIBUTION HISTORY	(a) Italian School, North - 16th century.
VARIANT ARTIST	
WORK DATE	16th century.
COLLECTION	Metropolitan Museum of Art
COLLECTION CITY	New York
FRICK CLASSIFICATION	Portraits: Men: Without hands: (without hats): Full face.
MEASUREMENTS	

PATH	NorthItalian_17-19thC/9847.tif
TITLE	Portrait of a Gentleman.
MATERIAL	oil on copper.
CREATOR	Anonymous, Italian School, North Italian
CREATOR DATES	17th cent.
SCHOOL	
ATTRIBUTION HISTORY	(a,b) Italian School, North - 1st quarter 16th c., (b) by Richard Offner, 1927, (a) by Friedsam, 1927, dated to circa 1600.
VARIANT ARTIST	
WORK DATE	c. 1600.
COLLECTION	Metropolitan Museum of Art
COLLECTION CITY	New York
FRICK CLASSIFICATION	Portraits: Men: Without hands: (without hats): Head to left.
MEASUREMENTS	4 7/8 in. (diameter); 12.4 cm (diameter).

This view gives a user, theoretically, everything that they need in order to determine what this newly-discovered match is and how these two images are related. The match was categorized on three axes:

1. **Work:** whether the artwork being depicted was the same work, a different work, or the same work but modified some how (e.g., before and after restoration).
2. **Photo:** whether the photograph was the same photo (100% identical), a similar photo (similar framing and composition with slight differences), or an alternate shot (such as a detail shot).
3. **Data:** whether the corresponding metadata of the two images agreed, disagreed, or was ambiguous. (When looking at the data it was only marked as 'agreed' if the data was obviously referring to the same artwork, typically held at the same institution.)

After I manually completed the categorization of all 446 matches between 815 images, the results were sorted into appropriate "bins" that denoted interesting trends.



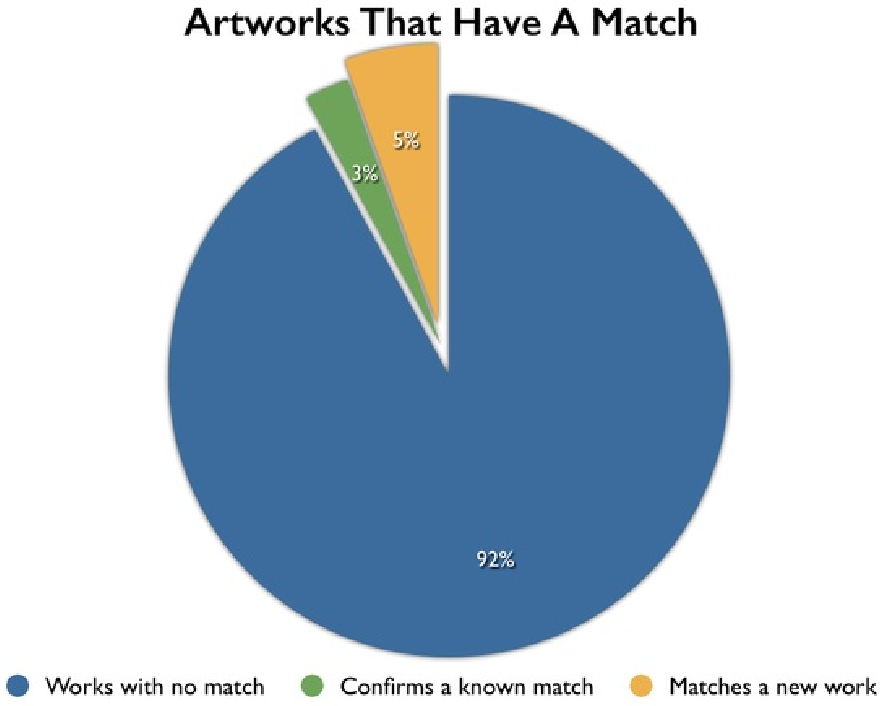
Alternate images for the same work of art

(**Work: Same, Photo: Alternate Shot, Data: Agrees**, 115 matches, 0.62% overall)

All of these binned matches were then passed on to researchers at the Frick Photoarchive for further analysis and record correction.

Results

The Italian anonymous photo archive was represented by 14,284 artworks. The image analysis found a match in 1,135 artworks (8%), including both newly-discovered matches and confirmations of existing relationships. Of those matched, 770 artworks (5%) had at least one new match with another distinct artwork, producing a total of 385 previously unknown inter-artwork relationships.



Out of the total 18,548 images, 1,187 images matched a known work of art and 446 new image pair matches were discovered. (An artwork can be represented by many individual images. In fact, one artwork alone had 152 photos associated with it.)

A complete examination of the image similarity analysis performed upon the Italian anonymous photo archive requires an understanding of three areas of results:

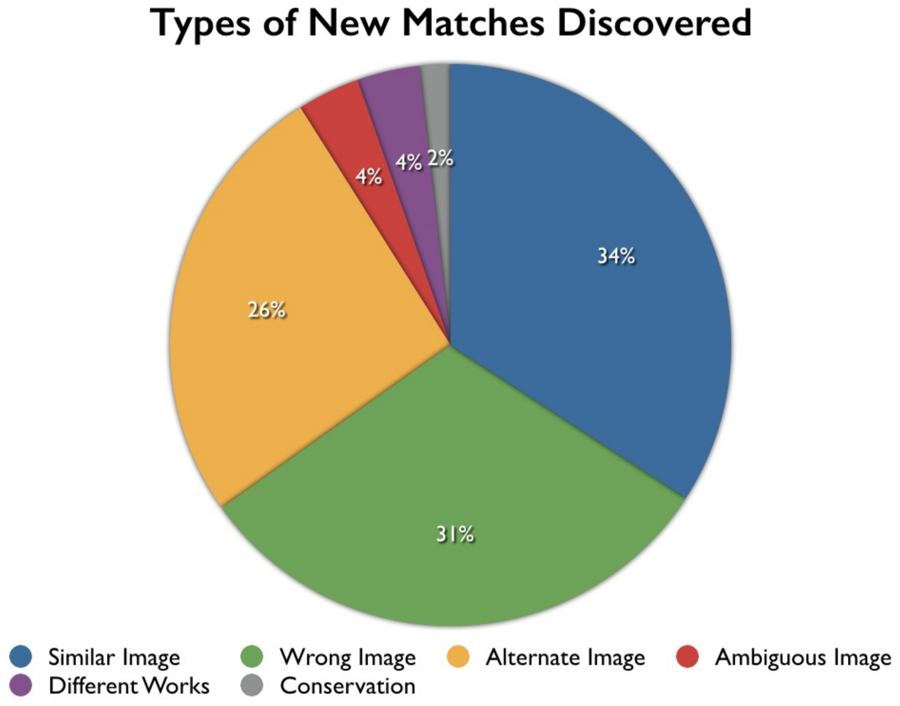
1. **New Matches:** completely new, previously un-cataloged, relationships between images discovered using the image similarity analysis.
2. **Confirmation of Known Matches:** confirming previously-cataloged relationships between images using the image similarity analysis.
3. **Unconfirmed Known Matches:** previously-cataloged relationships between images that the image similarity analysis failed to identify.

These studies were performed in order to look at all aspects of the image similarity analysis and determine what the analysis was capable of and what its limitations were. Learning that it was capable of confirming existing matches created by a researcher, as well as learning what matches it was unable to confirm, can help to set some expectations about how image similarity analysis can work for other photo archives.

New Matches

The new matches discovered by the image similarity analysis were certainly the most exciting for the researchers at the Frick Photoarchive. The analysis was able to accelerate their understanding and correction of the metadata associated with the digitized images.

The types of new matches broke down into a number of different areas:



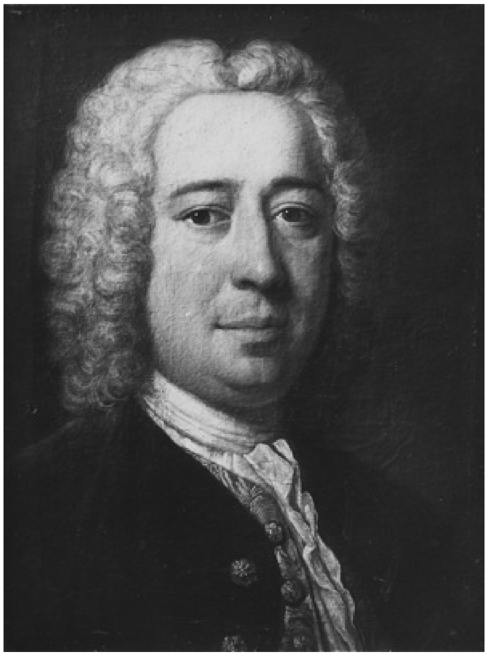
1. **Similar Images:** photographs that are highly similar (with the only differentiating factors being the difference in scan or lighting).
2. **Alternate Images:** matches where one photograph is an indirect, alternate, view of the same artwork (such as close-up of a detail or the same artwork viewed from an alternate angle).
3. **Conservation:** photographs of the same artwork most likely taken before and after conservation or during the process of conservation.
4. **Different Works:** photographs of two different artworks that are highly similar.li>
5. **Wrong Images:** the same, or similar, photograph but with the metadata in strong disagreement (likely resulting from a cataloging error).
6. **Ambiguous Images:** the same, or similar, photograph, but with ambiguous metadata (could be the same artwork but it's unclear).

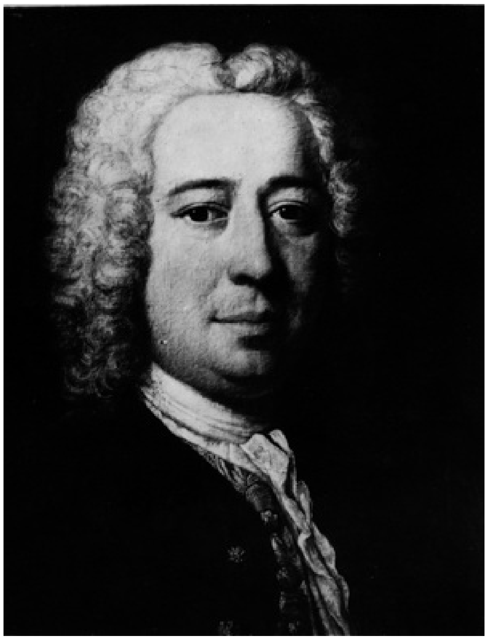
The majority of the new matches (65%) were legitimate new discoveries previously missed by researchers. The remaining 35% of the matches were potential cataloging errors (most of which likely happened during the digitization process of the images).

Similar Images

These are the same works that had a highly-similar photograph (of which there were 152 matches). This is the most obvious level of similarity: everything agrees (both the image and the data) in a very obvious way. Often times, these photographs would have similar cataloging details but were organized into different time periods or regions of Italy (thus making it more difficult for researchers to spot the discrepancy and correct it).

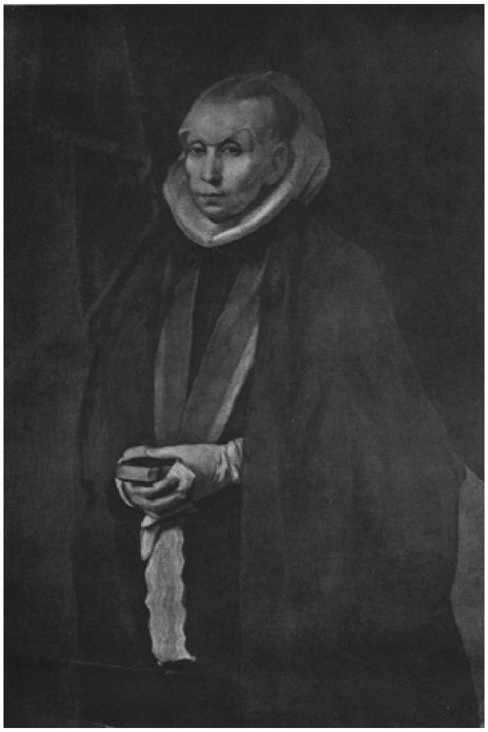
The first image shows the same work of art simply presented in two different, but similar, photographs. The only major difference is the lighting (obscuring a large portion of the painting). This was, by far, the most common type of similar image discovered through the analysis.

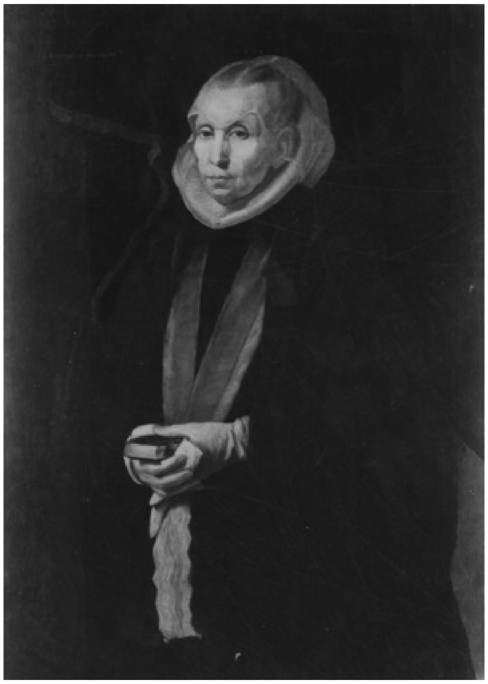




New Match: *different lighting, same work of art.*

Another similar pair of images was discovered in which virtually everything agreed except for a critical piece of cataloging: one was cataloged as a full-length portrait of a man, the other as a portrait of a lady.



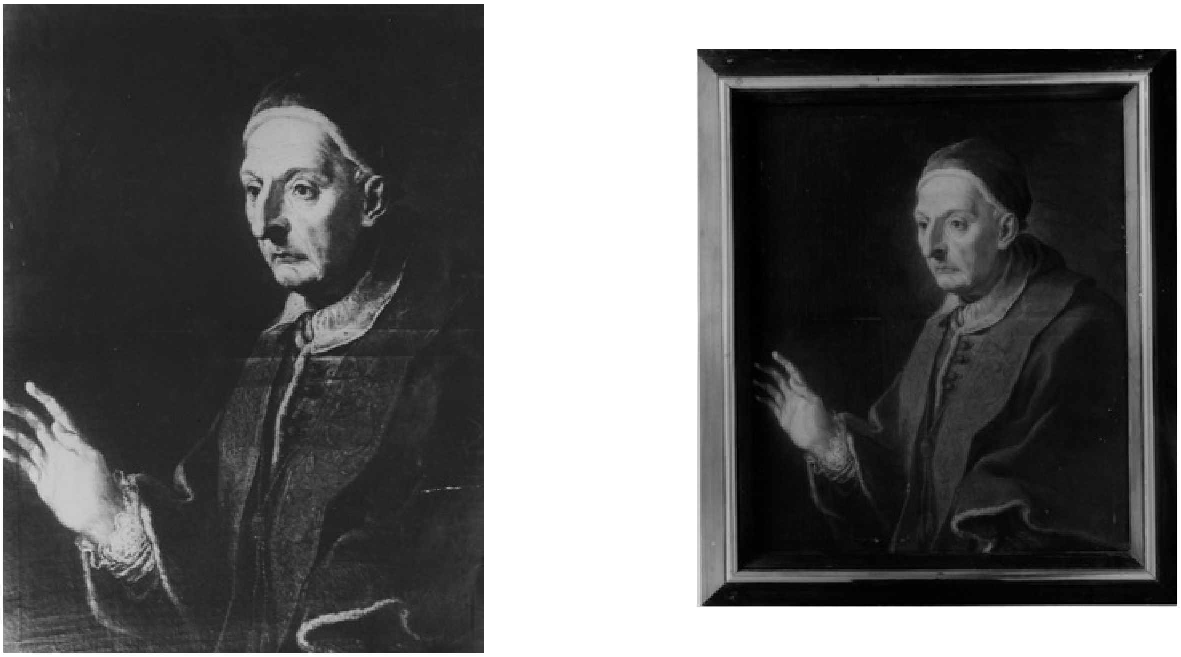


New Match: *different lighting, same work of art.*
(One categorized as a full-length portrait of a man, the other as a portrait of a lady.)

Alternate Images

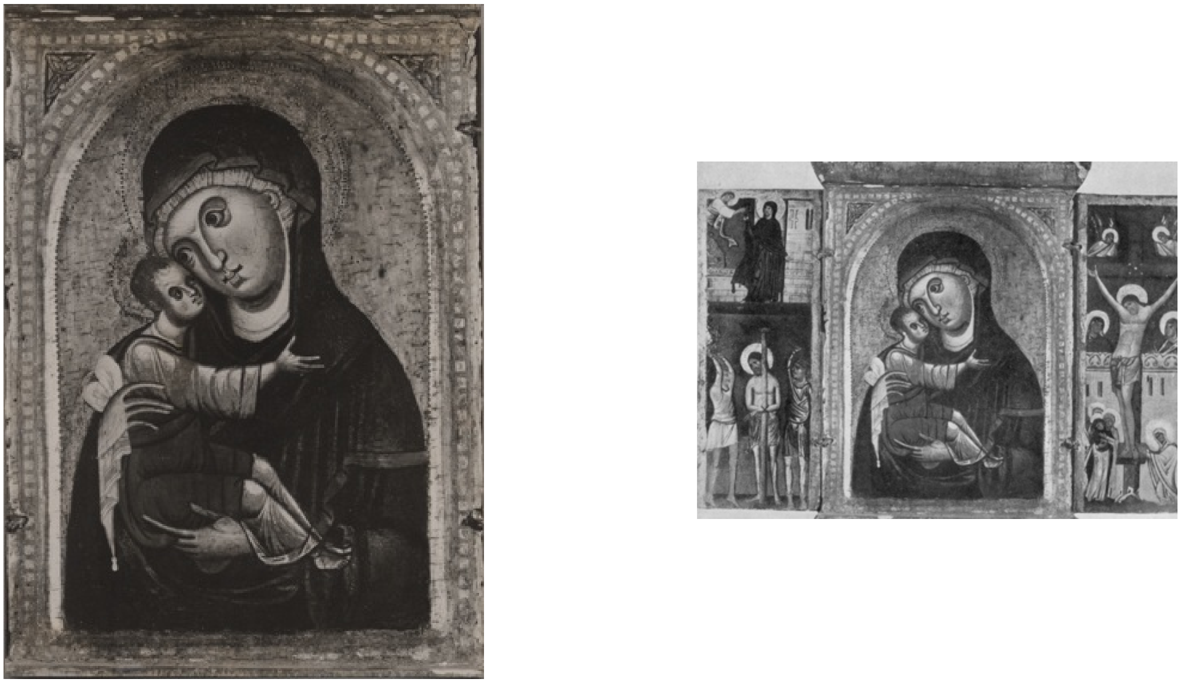
These matches were photos that both depicted the same work of art but showed alternate views (for a total of 115 matches). Frequently this was some sort of detail shot of the work. In all of these cases both the images and the data agreed. These matches were particularly interesting as finding a portion of an image inside another one can be quite technically challenging. Seeing the results provided by MatchEngine were quite heartening and suggested the possibility of finding many detail shots of a work of art.

The first work of art shows a dramatic difference in lighting as well as cropping. The photo on the right includes the frame of the work whereas on the left the image is cropped dramatically (into the painting itself).



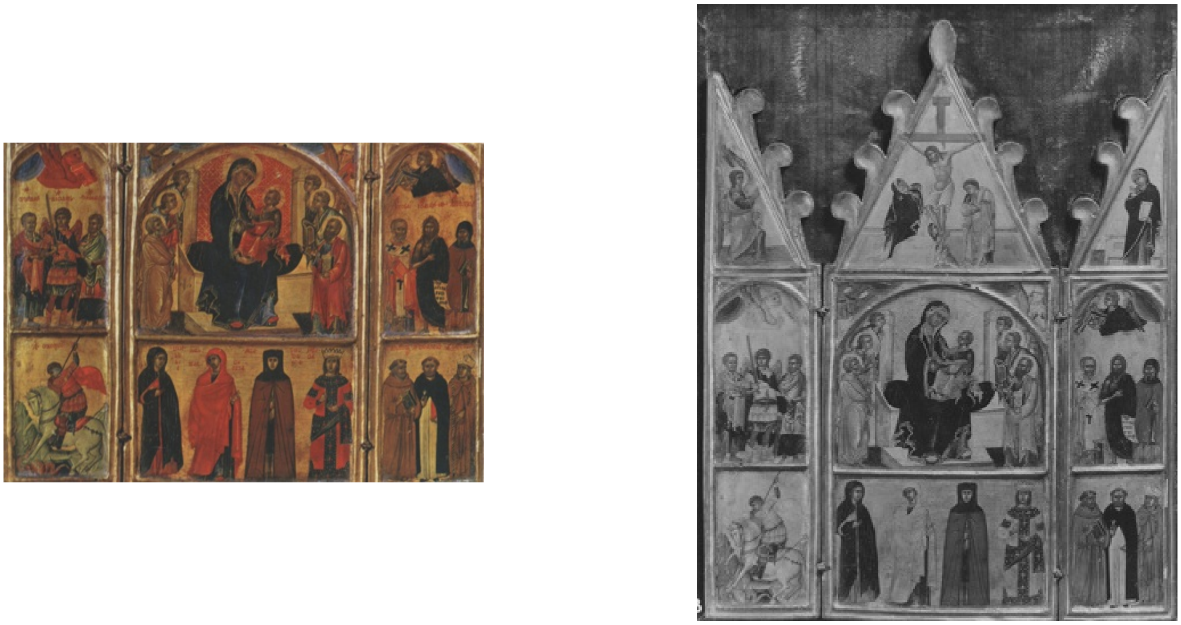
New Match: *different cropping and lighting, same work of art.*

The next work shows a close-up of the center portion of the work. Both photos are also in black-and-white.



New Match: *detail of the same work of art.*

This final representative on an alternate, match is both a close-up detail shot and in color, compared with the black-and-white full shot.

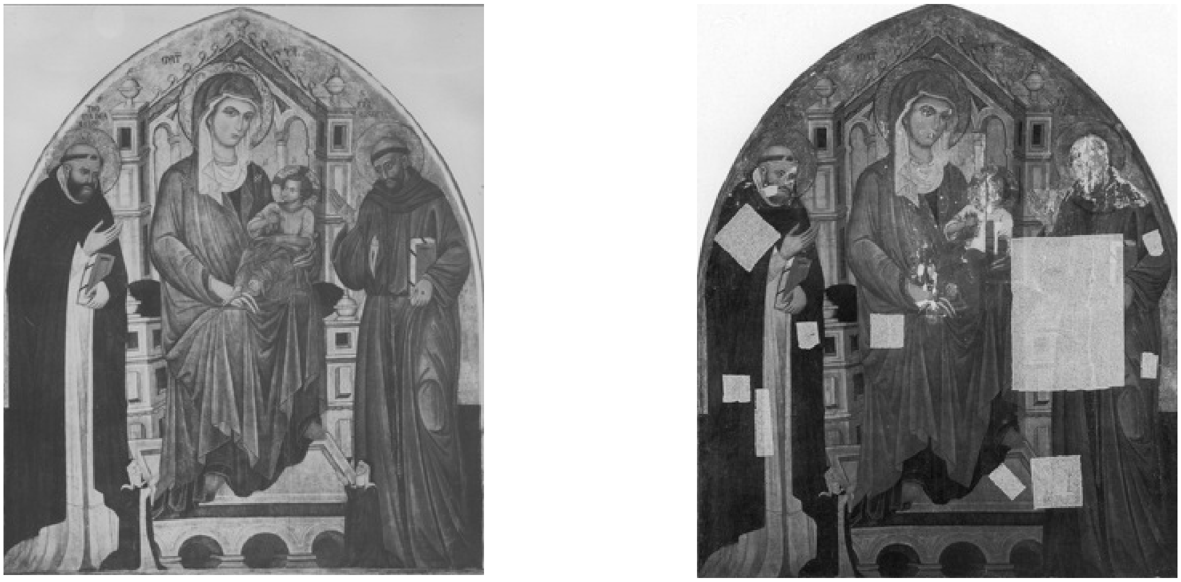


***New Match:** detail shot, color vs. black-and-white, same work of art.*

Modified/Conservation Works

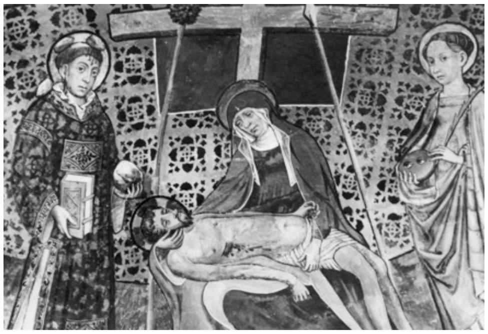
The same artwork before and after the conservation process were discovered during the image similarity analysis. Since the photos in the Frick's collection span many years, there are many instances where there are early photos of an artwork (from the early 1900s) together with photos from later in the century. Occasionally, an artwork will be in the process of restoration or will have undergone restoration at some point in the interim. Eight works were discovered in which possible restoration had been undertaken.

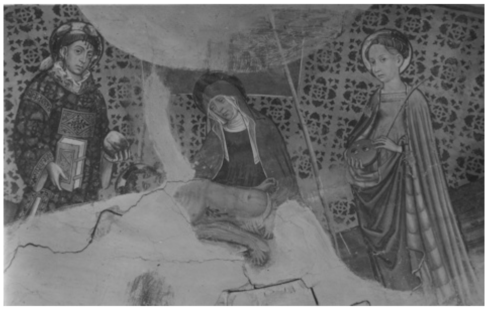
In the first work, restoration is in progress (seemingly an x-ray photography of the work):



***New Match:** same work of art, seemingly an x-ray or an in-progress restoration.*

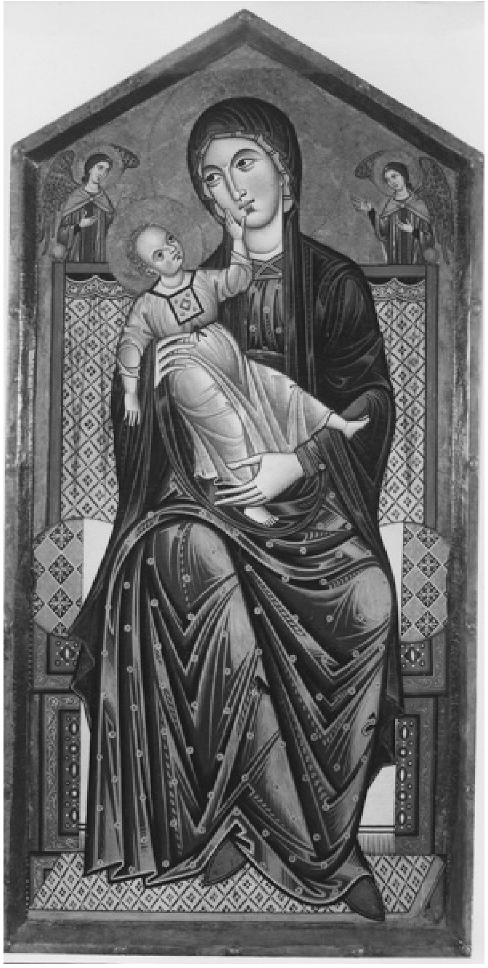
In another match, extensive restoration has been completed. Large portions of the fresco have been rebuilt and re-painted.

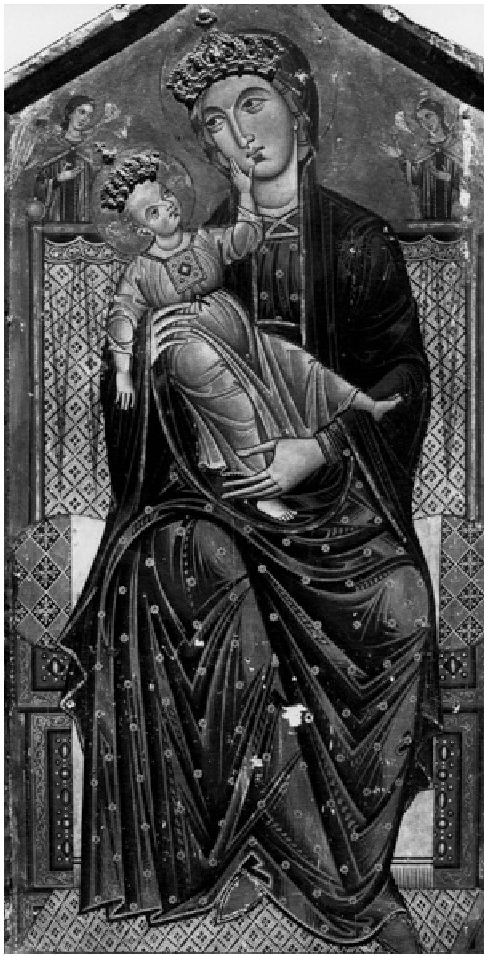




New Match: *same work of art, before and after restoration.*

Finally, a more subtle example: chipped paint has been repaired, the frame has been repaired, and seemingly extraneous crowns have been removed.





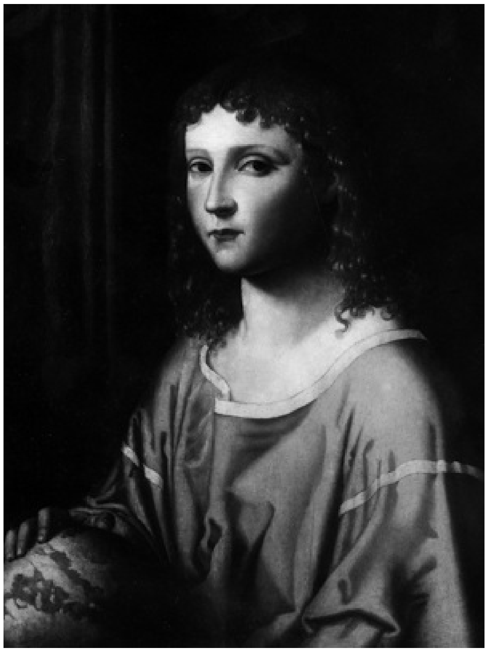
New Match: *same work of art, before and after restoration.*

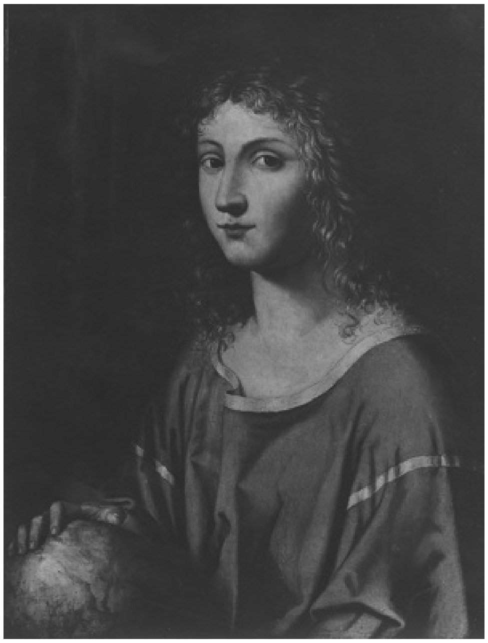
Copies

16 pairs of similar, but slightly different, artworks were discovered: the artworks were both copies of each other or of a third artwork. This discovery was especially interesting as it showed how potentially powerful MatchEngine's algorithm is. Even though the photographs aren't of the same work, it's still able to find the

strong similarities between the works and expose them as a strong match.

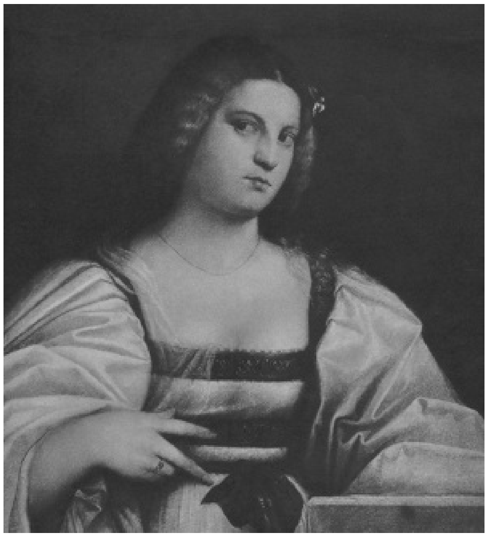
The first two works are both later copies of the same work by Leonardo Da Vinci. Note the differences in the faces and in the globe.

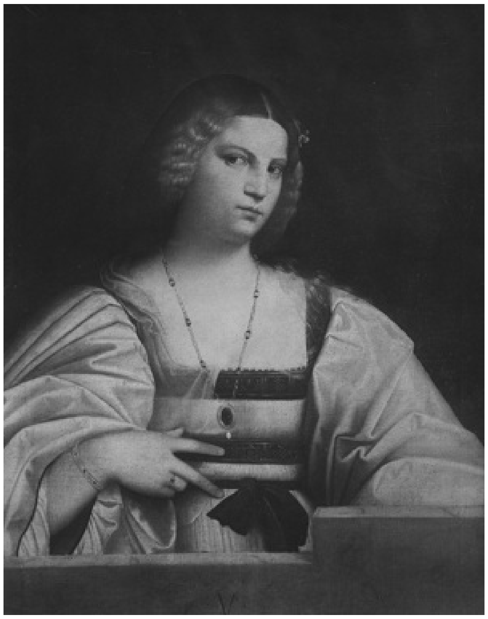




New Match: *different work of art. Note the different face and globe.*

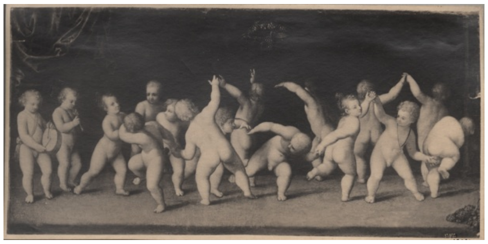
In another case, both works of art are copied from a third work (with slightly different faces and different necklaces).

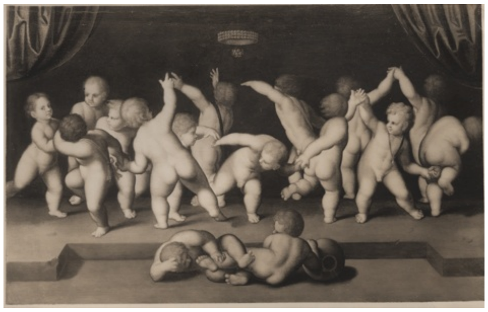




New Match: *different work of art. Note the different face and necklace.*

In this final case, both works are seemingly quite similar, with changes to the positioning of the children, the addition (or removal) of some children at the bottom of the work, and a change in the chandelier.



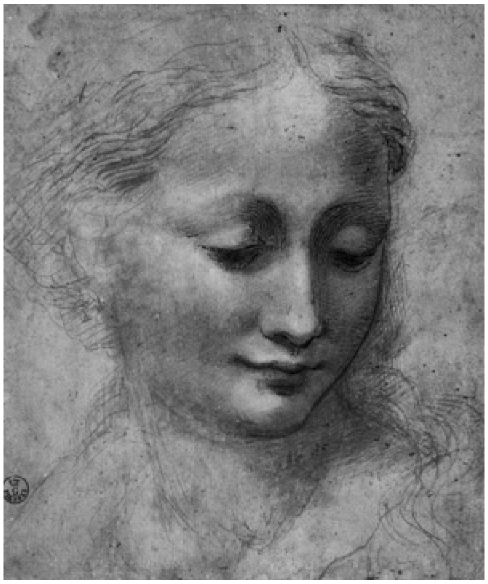


***New Match:** different work of art. Some children missing, added, changed.*

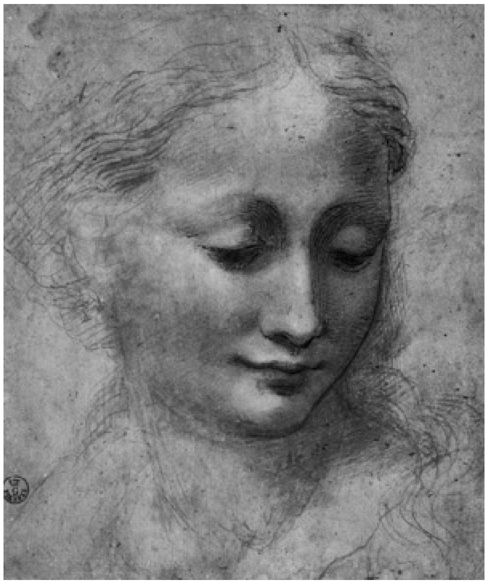
Digitization Errors

The image similarity analysis was also able to uncover 138 unexpected matches: cases of identical artworks with metadata in strong disagreement. These seemed to be the result of either the wrong image being uploaded for a work or the wrong metadata being used. Either way, it appears as if most of these problems occurred during the digitization process by the outside vendor because the Frick's internal physical records are still correct. Such discoveries are especially useful: the Frick Photoarchive has been able to correct the erroneous data and provide a better digital archive as a result.

The following works exemplify the kind of cataloging errors that were exposed. The images appear to be virtually identical yet have very different metadata. It's likely that the wrong image was paired with a metadata record, in this case:



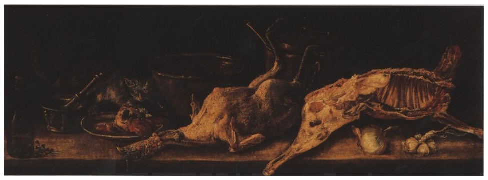
Arms with Folded Hands
Castello sforzesco, Milan.



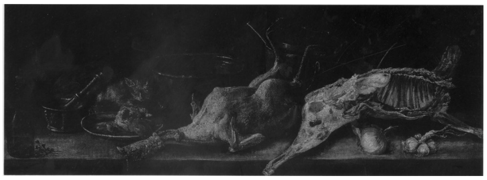
Female Head
Gabinetto disegni e stampe degli Uffizi, Florence.

First work doesn't match description, wrong cataloging.

Additionally, these photos are in color and black-and-white but disagree on the metadata. In this case, it's likely that the correct image was uploaded but the wrong metadata was used.



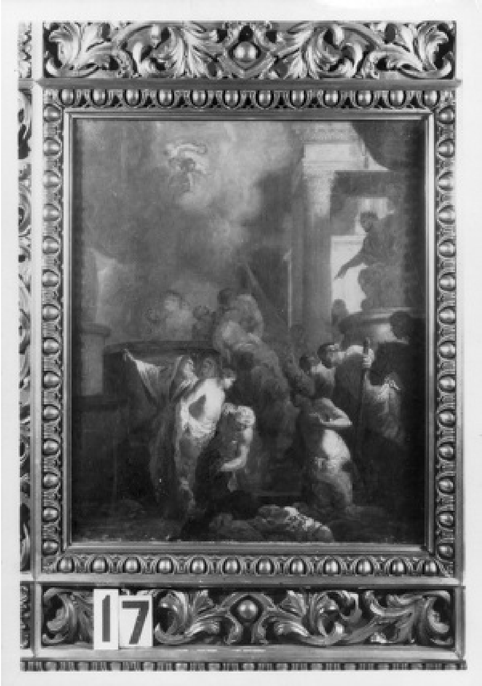
Still Life with a Bottle, a Plate, a Mortar and Pestle, a Bowl, a Pot, Game and a Cat on a Stone Ledge.



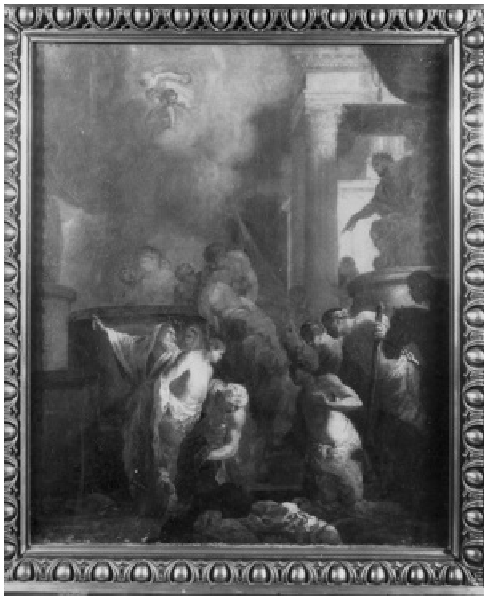
Virgin Entrhoned Nursing Christ, Between Two Saints.

Second work doesn't match description, wrong cataloging.

There were an additional 16 matches which may have been a cataloging mistake, or may actually be correct and require additional exploration by a researcher. The following match is such an example:



A Martyrdom
Accademia di San Luca, Rome.small>



The Corporal Works of Mercy
The Faringdon Collection Trust, Buscot Park.

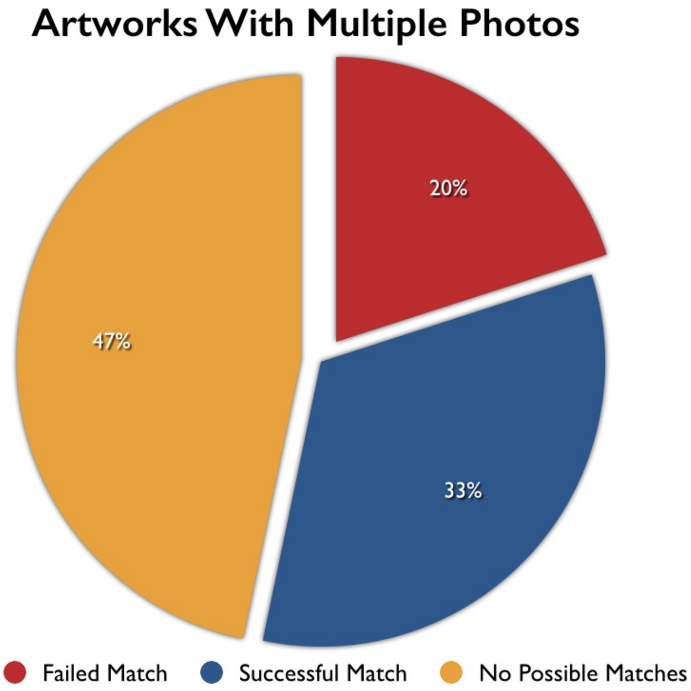
Same work, perhaps changed collections?

Measuring Image Analysis Efficacy

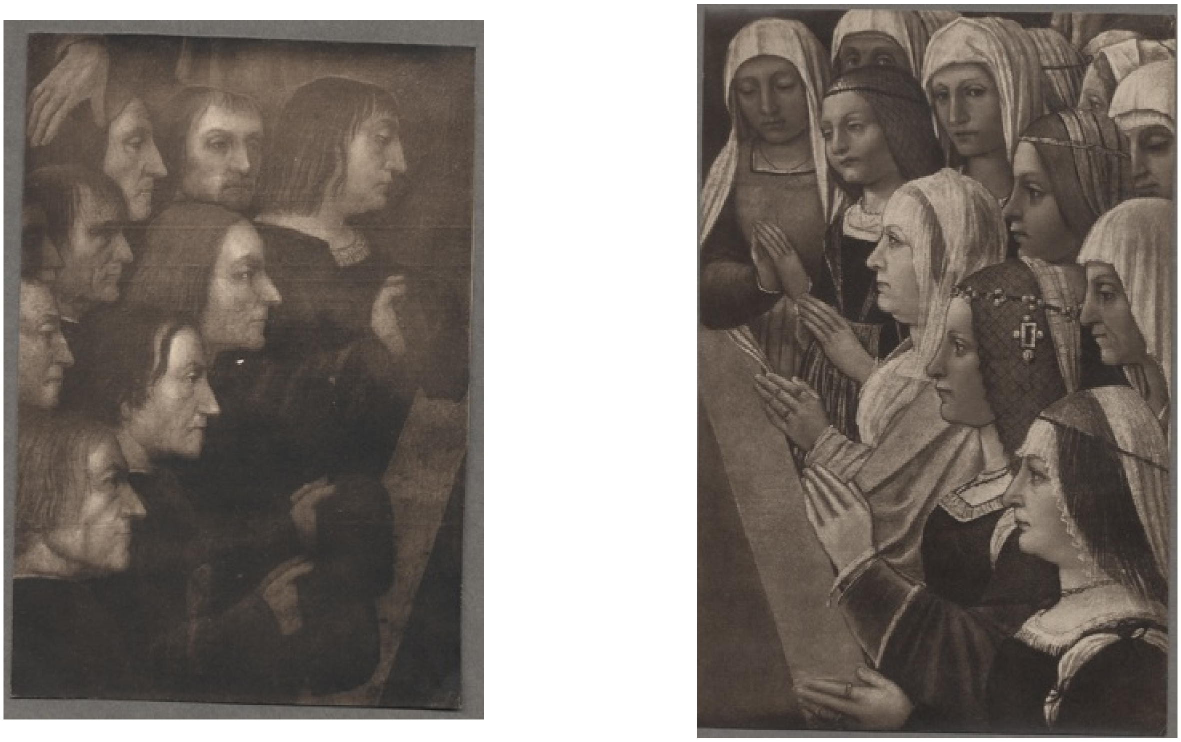
Even with all of these interesting new matches being discovered using image analysis it's important to attempt to understand how effective the MatchEngine algorithm is at finding matches. The best way to quantify this is by looking at images where a match should have occurred but did not.

Within the Italian Anonymous Art archive there are 1357 works of art associated with more than one photo. These photographs were manually grouped together by researchers at the Frick Photoarchive. The photos associated with a single artwork aren't always alternate views of the same work. Frequently, they are multiple photos of an artwork from different angle, the front and back of a work, or pictures of a three-dimensional artwork. Sometimes they are photos of different aspects of an artwork (for example three photos, each of a different panel in a triptych).

To better understand the types of photographs that were available for the artworks, a full survey was done of all 906 artworks that have multiple photographs but were not explicitly matched by the MatchEngine algorithm. The artworks were broken down into two categories: artworks for which there was no obvious visual relationship between the presented photographs and artworks for which there was some strong visual similarity between two or more of the photographs.



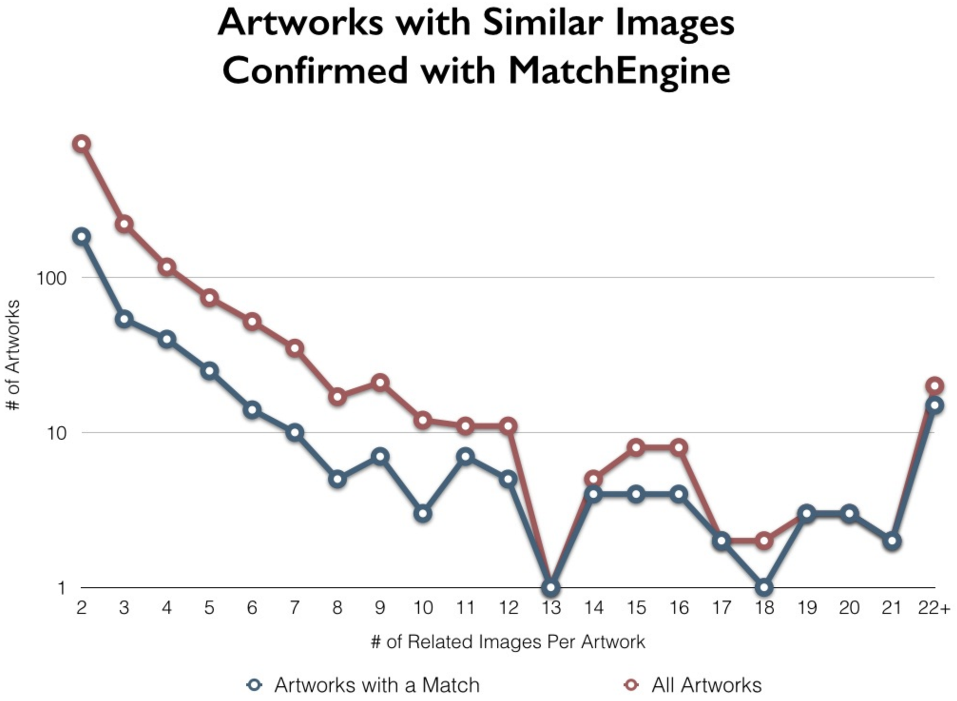
47% of all artworks with multiple photos had no two photos that were visually similar to each other. In those cases, the MatchEngine algorithm was incapable of finding any relationship: MatchEngine is only able to examine what is presented in the image itself. For example, the following artwork depicts two separate panels in the same piece:



em>Two different panels from the same artwork, no overlapping details.

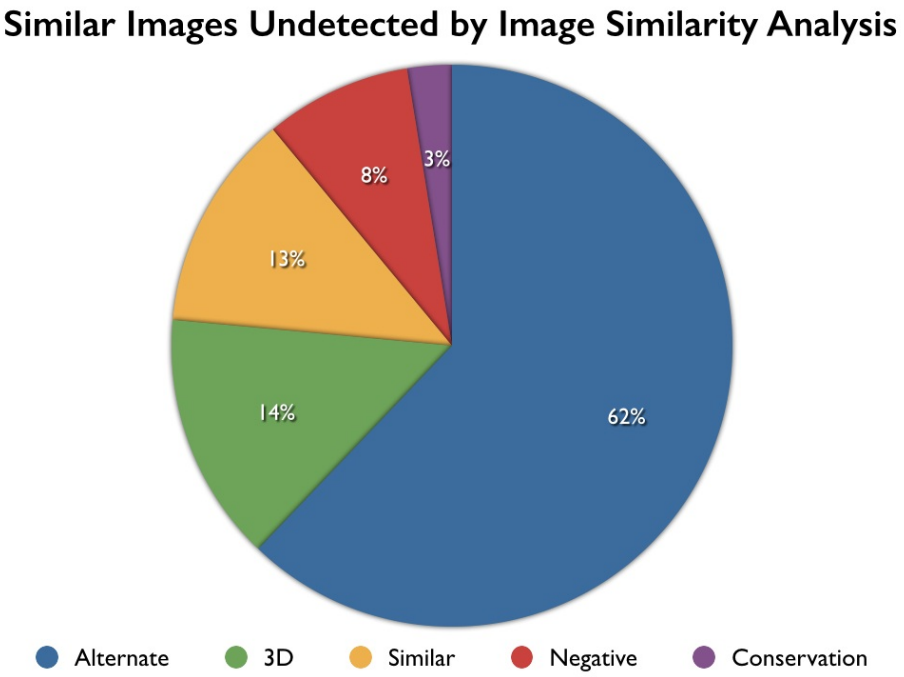
Of the remaining 53% of the images that did have a visual relationship between two or more of the images 33% were successfully matched and 20% were not.

Initially it was assumed that there might be a correlation between the number of photographs made available for an artwork and the likelihood of there being a confirmed match. An analysis was completed looking at artworks broken down by the number of photographs associated with the artwork:



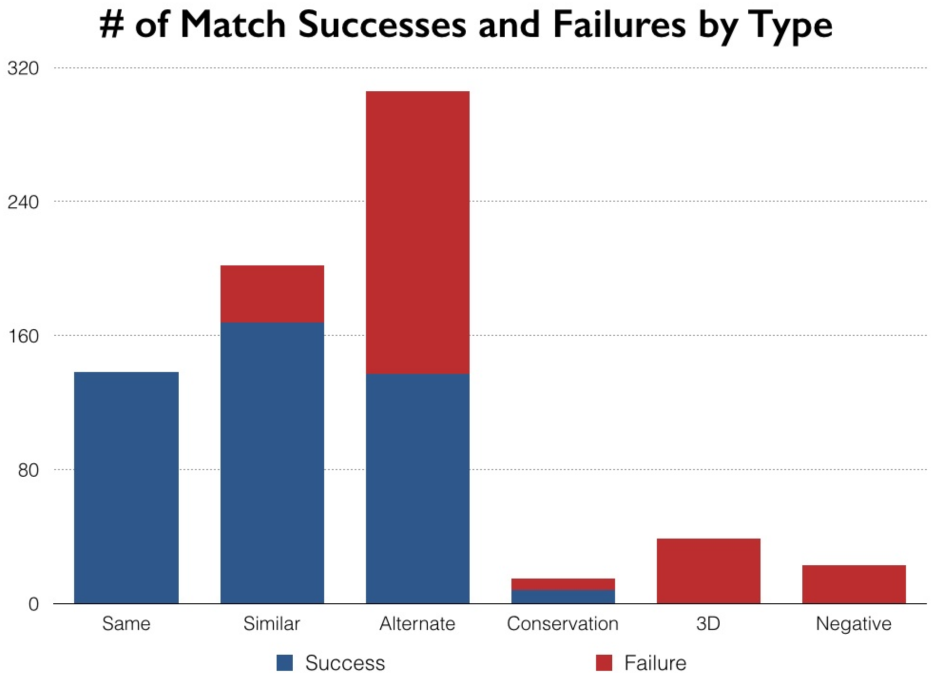
Looking at these numbers there does not appear to be a strong correlation between the number of photos associated with an artwork and the likelihood of there being a match. Only at the upper-end of the spectrum (for artworks associated with 19, or more, photographs) is there a strong correlation with a successful match occurring.

In order to understand where these matches come from and where the failings are, the images that were not matched need to be examined. This process will lead to a better understanding of the limitations of the MatchEngine technology and can help to set researcher expectations appropriately. A full breakdown of the types of images that weren't matched included:



To arrive at this breakdown, I performed a full survey of all the 272 artworks that have at least two photographs with a strong visual relationship. Where there were multiple potential matches between photographs, the best possible photograph pair was chosen to be representative for that artwork.

At first glance, the types of matches appear to be similar to the types of matches that MatchEngine successfully discovered. However, a final breakdown of the images that failed to match was compared to the images that successfully matched using MatchEngine:



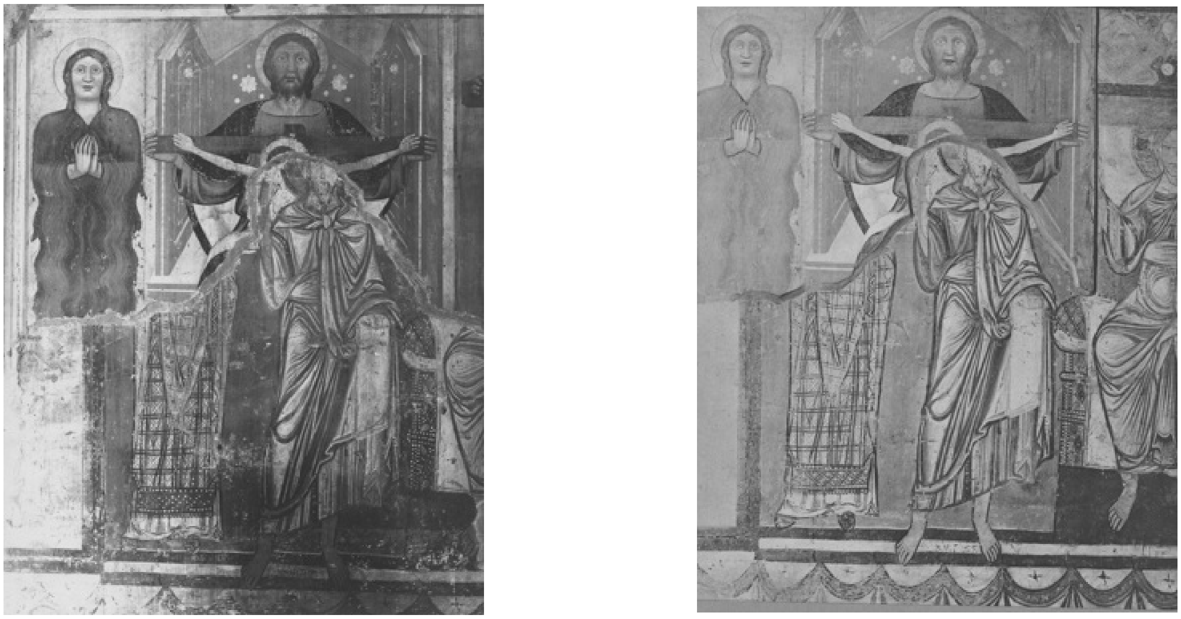
This is where the shortcomings of MatchEngine became apparent: every single three-dimensional and negative match failed in MatchEngine, as did the majority of alternate shots. Also note that there were comparatively very few failures where the photos were similar and no failures when the photos were near (or nearly) identical.

An analysis of all the individual types of match failures will help us to better understand what, specifically, MatchEngine struggled with in the identification of these images.

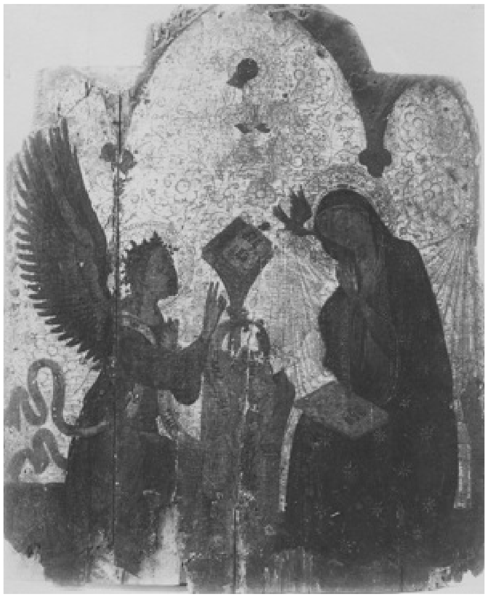
Similar Images

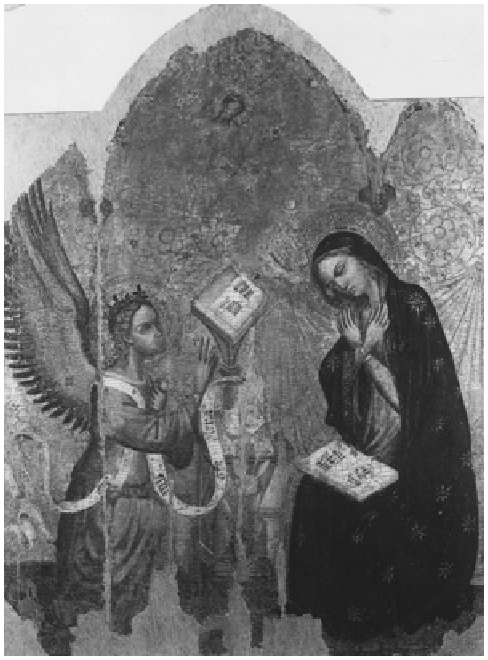
The following images are cases where the framing of the photographs are similar but the lighting between the shots is different. MatchEngine was able to successfully discover a number of these cases so it's a bit surprising that that it struggled with these results (there were 168 successful similar image matches and 34 unsuccessful matches – or about 17% of all similar image matches failed).

It's likely that the MatchEngine algorithm is looking at edges within the image, so with sufficiently different lighting some cases are no longer easy to pair. Below are some example of similar images that were not matched by MatchEngine:



Seemingly, the difference is between direct and raking lighting.



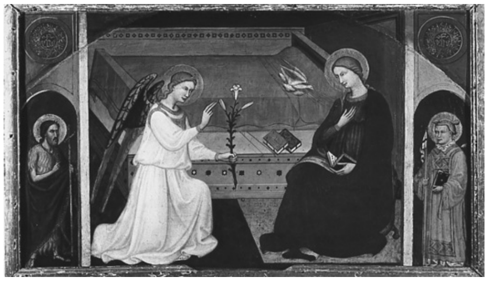


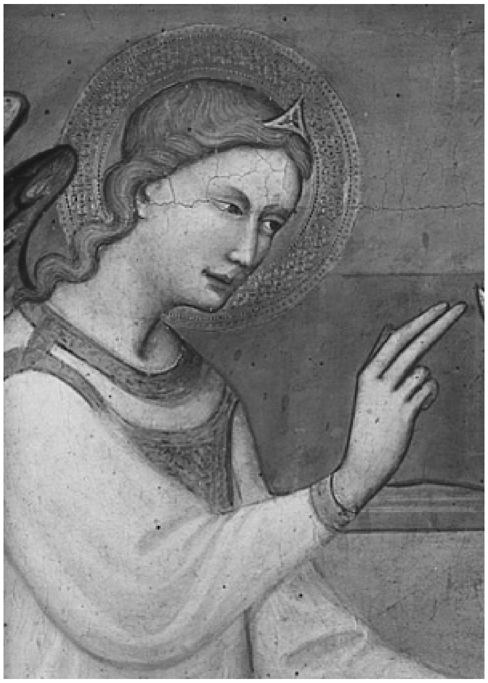
Very different lighting and exposure.

Alternate Images

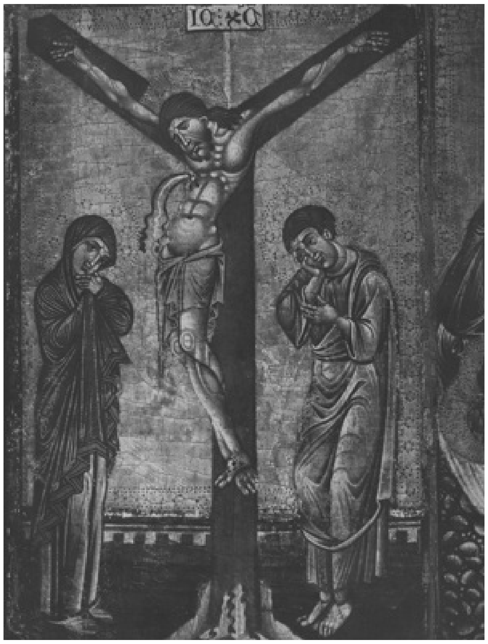
Alternate images of the same artwork produced the largest number of failed matches. In all of these cases a portion of an image contained within another image was not successfully matched. 137 alternate image pairs were successfully matched, whereas 169 alternate image pairs failed to match for success rate of only 45%.

Below are some examples of image pairs that failed to match, all of which were detail shots of small portions of the overall image:

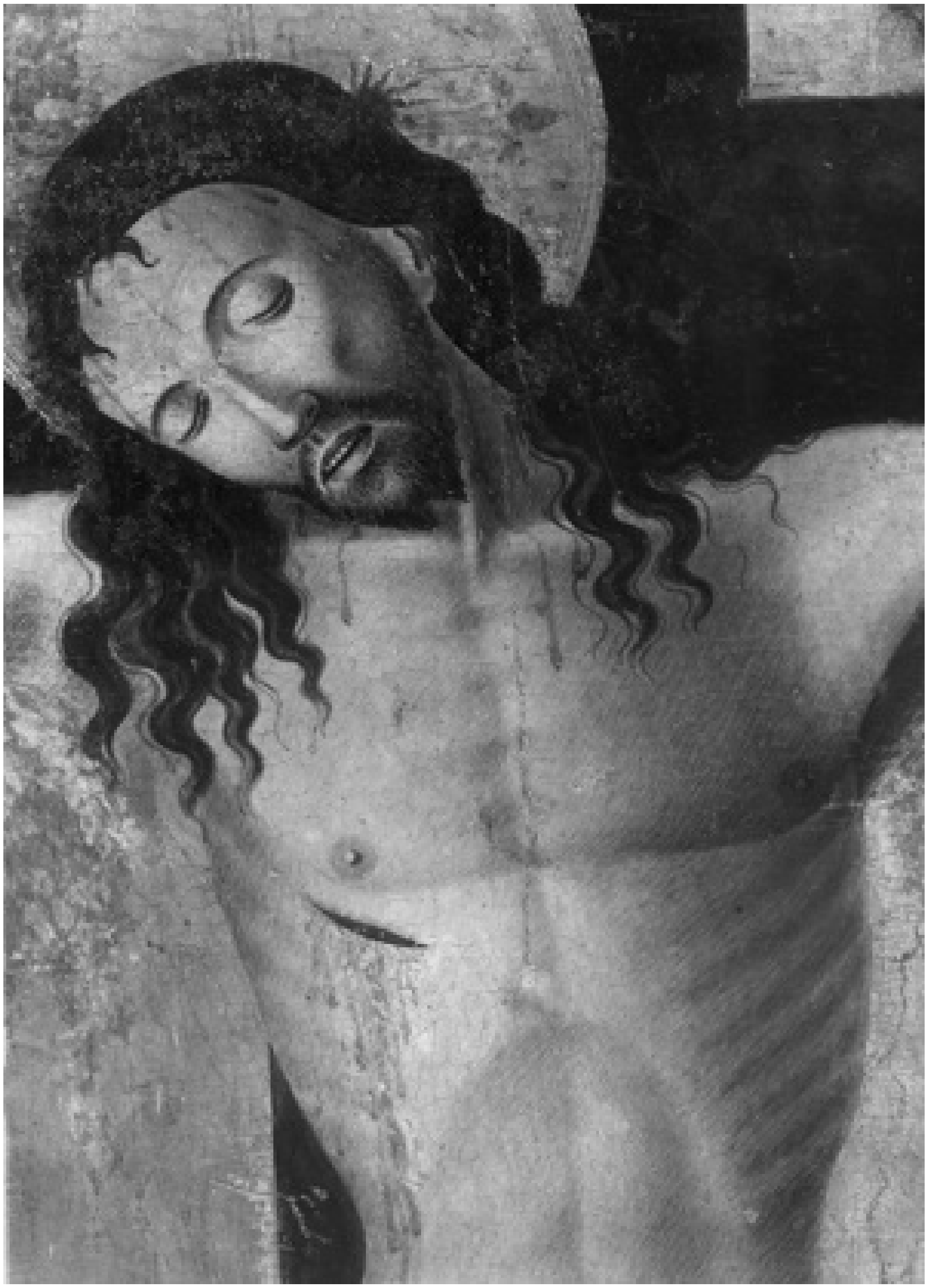




A small detail of the angel's head and arm.



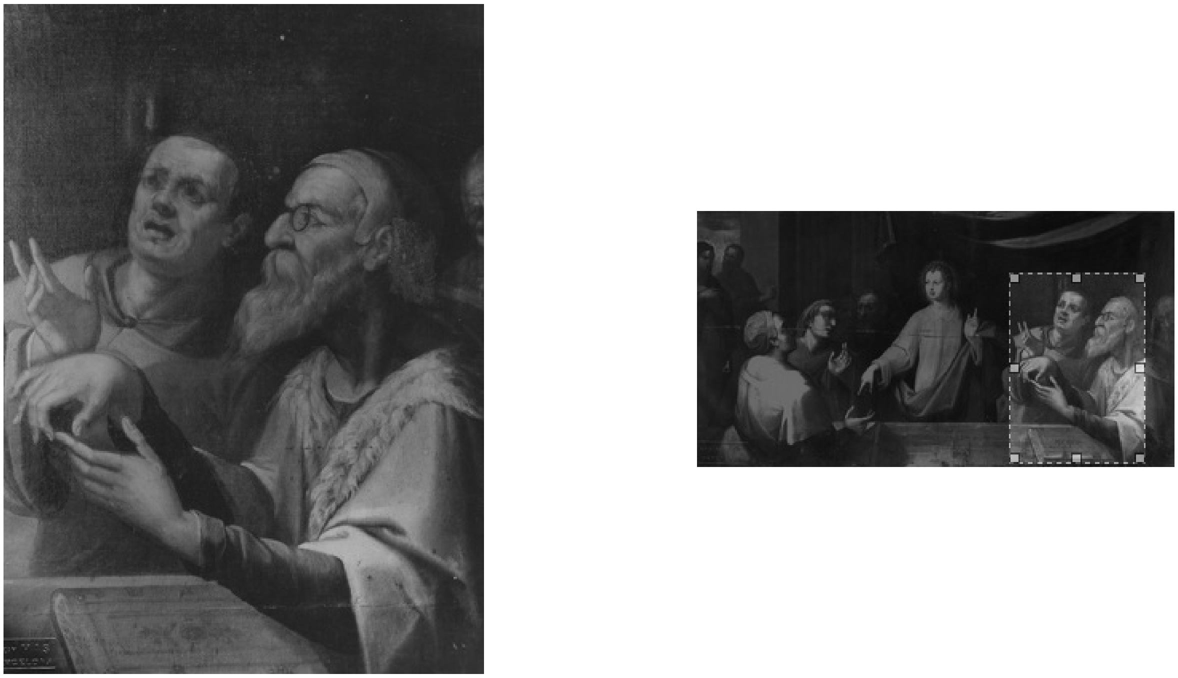
A tiny panel from the middle-right-hand side of the altar piece.



An extreme detail shot of the head of Jesus.

The poor results seemed particularly contradictory, as there was a large number of successful alternate image matches. However, one critical detail of the MatchEngine implementation is important to understand (this is also the case for most computer vision techniques): the image must be reduced in size before it can be successfully analyzed. In the case of MatchEngine, all images are reduced to 300 pixels in the smallest dimension before being processed. Taking this into account, and looking at the above example failures, an assumption can be made that there is a significant loss of detail during the processing of these images making a match difficult.

I also hypothesized that there is a correlation between the percentage of the image overlap between two images and the likelihood of there being a match between them (the larger the percentage the greater the likelihood of a match). To test this hypothesis all of the failed alternate image matches were analyzed. The overlapping portion of the image was manually selected to determine what percentage of the image was matching. A resulting selection would look something like this:

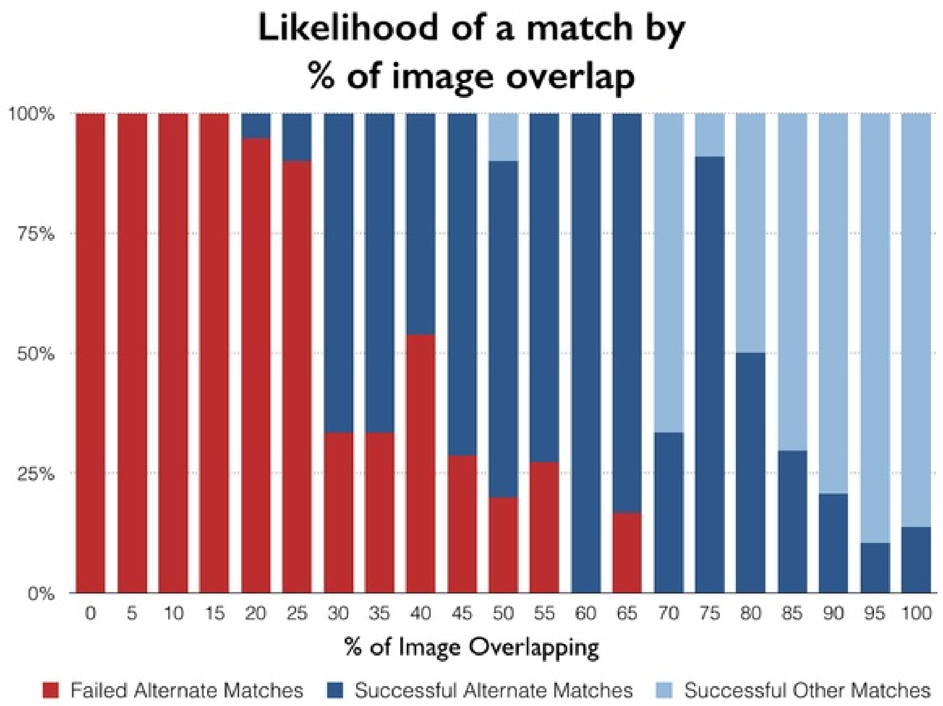


Manually selecting the portion of an artwork that overlaps with the corresponding alternate image.

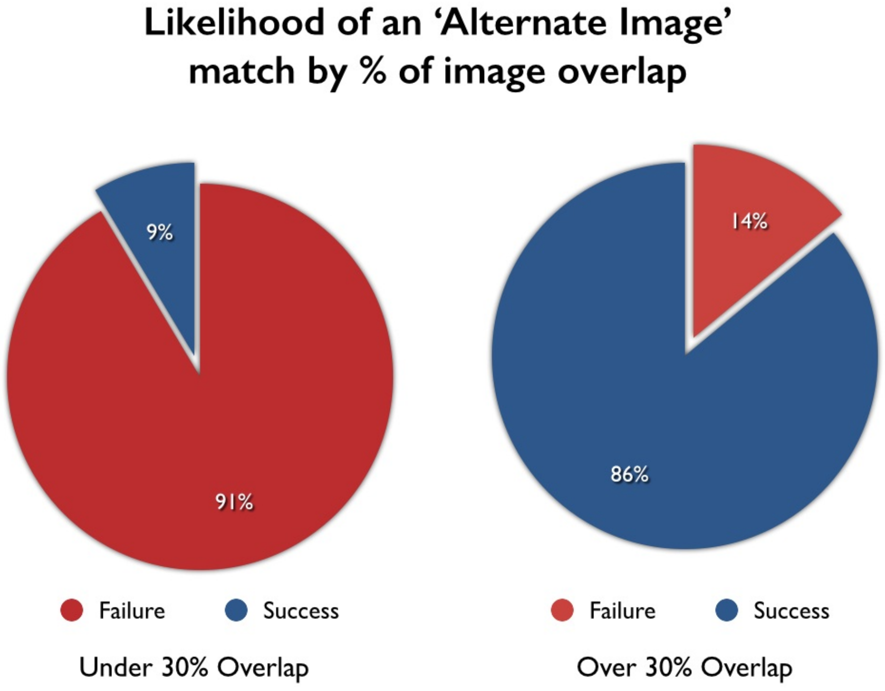
Thankfully, MatchEngine already provides the overlapping percentage for successful matches via their query API:

```
&quot;frick-anon-italian/13291.jpg&quot;: [
    {
        &quot;score&quot;: &quot;27.80&quot;,
        &quot;target_overlap_percent&quot;: &quot;100.00&q
        &quot;overlay&quot;: &quot;...&quot;,
        &quot;query_overlap_percent&quot;: &quot;47.18&quo
        &quot;filepath&quot;: &quot;frick-anon-italian/132
    }
]
```

All that was left was to plot out the alternate image match failures, the alternate image match successes, and the other successful matches.



Looking at these results, it becomes immediately apparent that there is a strong correlation between the percentage of the image overlapping and the likelihood of there being a successful match. Below 30% of the image overlap, there are almost no successful matches between images. If the results are broken down to show the matches with less than 30% overlap and the matches with over 30% overlap these striking results are generated:

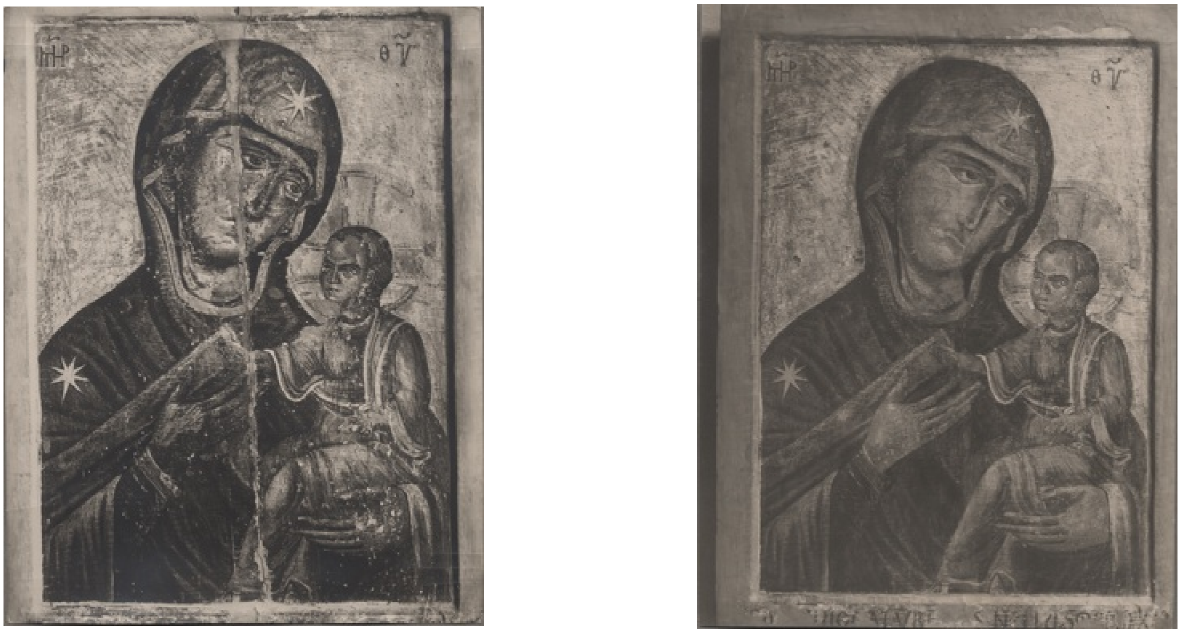


The results indicate that MatchEngine is not designed to adequately handle cases where there is less than 30% of the image overlapping. This is important to understand, as it can help catalogers better understand the limitations of computer vision systems such as MatchEngine. In many cases, when such a small fragment of the images overlap it is almost exactly like searching for a needle in an image haystack.

Conservation

As was the case with the successful matches there were a few cases where there were images of an artwork before and after the process of conservation. It was rather surprising that any images were able to match after conservation so it was unsurprising that nearly half of the conservation cases resulted in a failure to match (8 successful matches, 7 unsuccessful matches).

An example of a work, after conservation, that failed to match:



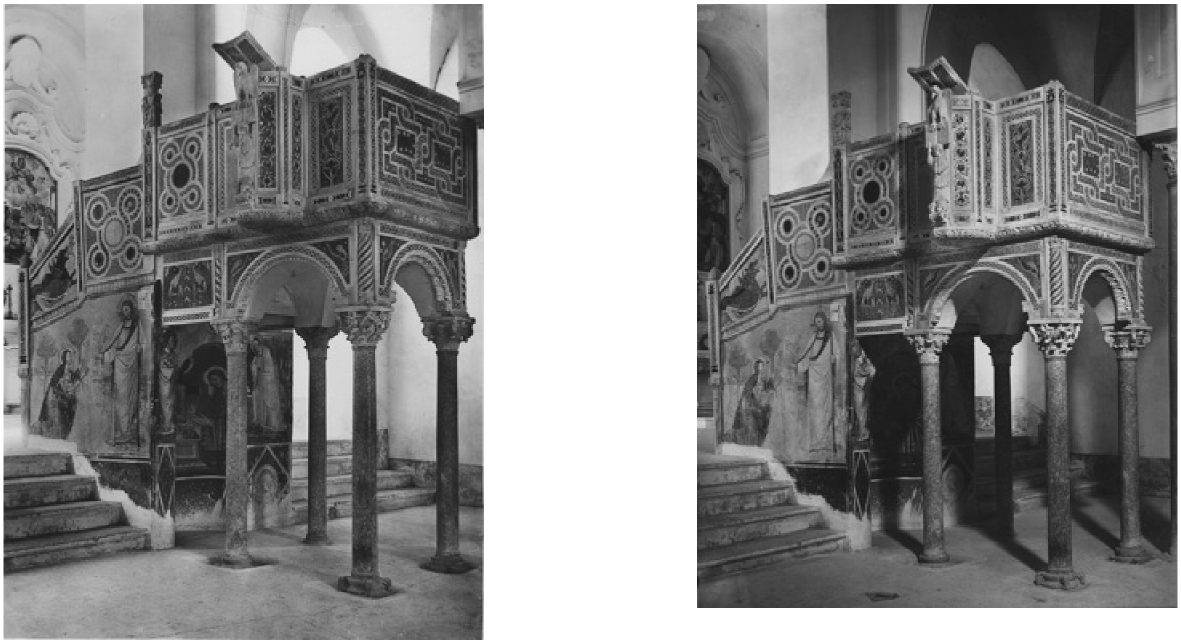
Work after conservation with different lighting.

Three-dimensional Works

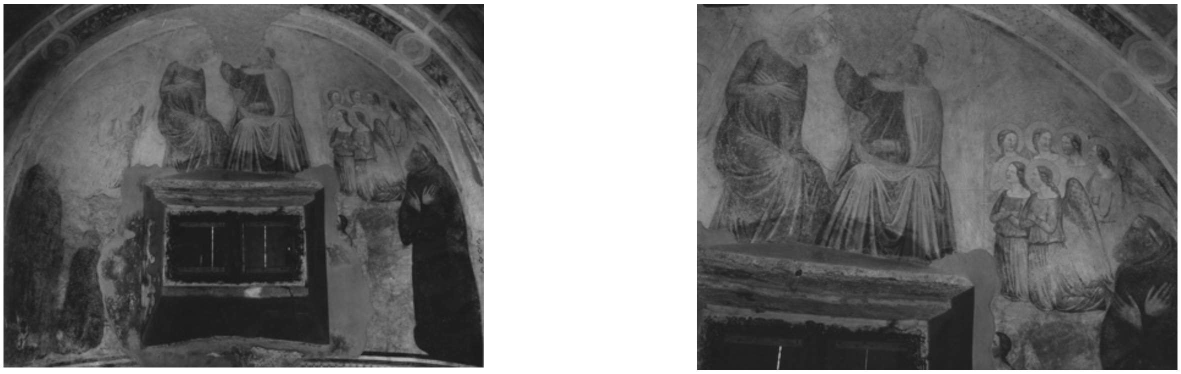
According to the MatchEngine web site, MatchEngine “cannot be used for identifying 3D objects.” Analyzing the failures tends to come to the same point of agreement: none of the 39 three-dimensional artwork images successfully matched each other.

Presumably, a different service would need to be used to find three-dimensional matches of this nature. Unfortunately, I am not aware of any services that provide this technology in a way that is able to gracefully scale to thousands of images in the way that MatchEngine can.

The results included the following incomplete matches:



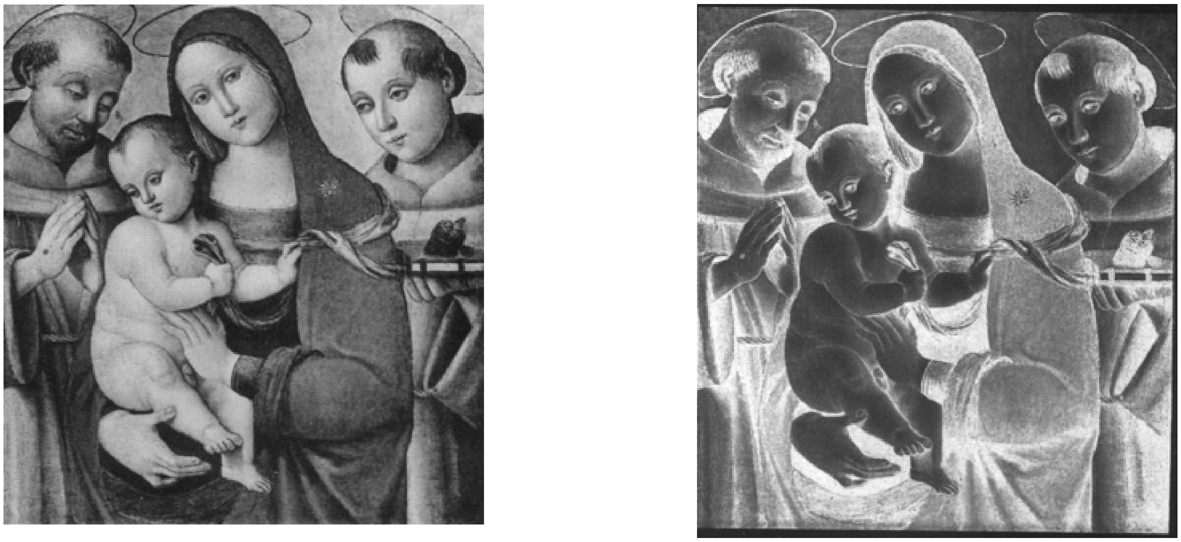
Same object with different lighting.



Same object at a different angle (even though it is a fresco, it's observed from different angles, causing a failure).

Negative Images

The anonymous Italian art archive contains 23 artworks whose only alternate image is a negative. In all 23 cases, MatchEngine failed to find a match between the primary image and the negative. Considering that MatchEngine never claimed to match these types of images, it is safe to assume that this is not a use case that MatchEngine was designed to handle.



The same artwork in normal and negative views.

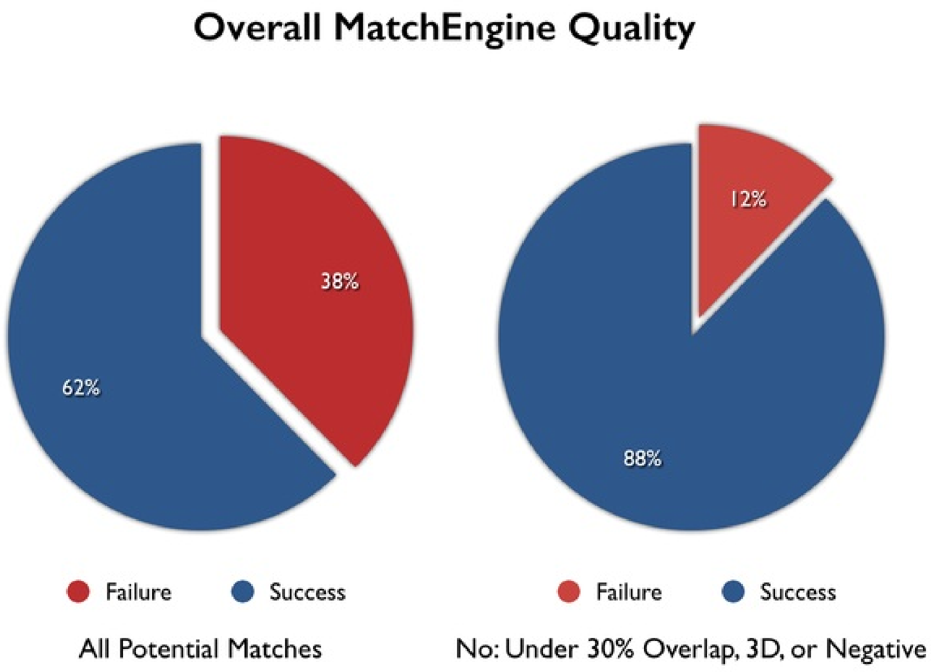
A proper match between the positive and negative forms of the image would be possible with MatchEngine if all negative images were first converted to their normal, positive, form. There would be extra work involved in making the match happen and since so few of the images in this particular archive fall under this criteria it was not deemed worthwhile to make this conversion.

Conclusion

This initial foray into using computer vision techniques to enhance the research potential of photo archives was exceedingly successful. A number of unknown relationships were discovered between images, digitization mistakes were detected, and corrections were offered. Additionally, the MatchEngine service used for performing the image analysis worked better than either the Frick Photoarchive or I could have hoped.

The potential of the MatchEngine service for the image set was fully explored: it works exceptionally well for images that are very similar, or for photographs that have moderate lighting changes, or for detail shots of the same artwork. However, MatchEngine is not a good tool for analyzing three-dimensional objects, detail shots with small amounts of overlap, and photographs with drastically different lighting.

Taking all of this into account, the overall quality of matches that MatchEngine provided within the anonymous Italian art archive was around 88%:



While there are limitations to computer vision techniques on the whole, these results are very promising. This high rate of match implies that there could be relatively few undiscovered new matches. Moreover, even after looking through all of the matches, MatchEngine never presented a single mistaken match. Every match had a high level of similarity between the two images and made sense to the catalogers.

It's important to note that this particular archive is likely one of the most challenging use cases for using computer vision techniques in general (other archives are likely to have a much higher rates of match). The fact that most of the images in this archive were black-and-white (lacking additional information about the colors of the work) was a major hindrance to improved matching. The less data that the analysis engine has to work with, the harder it is to make a successful match. Additionally, many of the photographs in the set had drastically different lighting between shots, making it very hard to do comparisons. Presumably, another archive that had consistent lighting would fare much better.

With this new, powerful image analysis, the real fun begins: looking for other ways in which this analysis can benefit archives. There are three areas in which this image analysis would have immediate impact:

1. **Analysis and Error Correction:** the case demonstrated in this paper. Analyzing an established archive and using image analysis to look for undiscovered connections and to correct potential cataloging mistakes.
2. **Digitization:** performing image analysis during the digitization process. This analysis would provide the digitizer with contextual information about the work they're processing and help them to spot possible duplication or errors before they update the catalog.
3. **Merging:** given two archives of photographs, detect similar images and automatically merge the metadata records for a photograph. At the moment, the only solution to merging two archives is to attempt to rectify all of the metadata (which can be especially challenging if the archives are in different languages). If image analysis was used then all of the troublesome metadata could be ignored and relationships would be discovered purely based upon the images themselves.

The potential for computer vision and image analysis to change how photographs and images are managed in archives, libraries, and museums is absolutely staggering. Tasks that previously were insurmountable (such as merging two million-photograph archives) are now in the realm of possibility. The implications of this technology are still being explored and are likely going to completely change photo archives as they currently exists.

Originally published by John Resig on February 10, 2014. Revised for *Journal of Digital Humanities* July 2014.

Thanks

I would like to thank the Frick Art Reference Library for their interest and collaboration in exploring the potential of image analysis for photo archives. I received tremendous encouragement from them to explore this research and I'm very excited about collaborating with them more.

The Tineye team have been a pleasure to work with. I've been extremely pleased with the quality and reliability of their MatchEngine API. A few years ago, I explained to them some of the projects that I wanted to work on and they were excited to support me in their development by providing me with free access to their MatchEngine service. They've asked for nothing in return but I feel duty-bound to point out how good the service is and why you should use them if you have similar image matching needs.

I would also like to thank the Kress Foundation for providing a grant to fund future collaboration with the Frick Art Reference Library in developing Open Source tools for art photo archives to perform image analysis on their collections.

[1] At the moment the pricing for MatchEngine only works on a monthly payment cycle and doesn't exactly match the use case outlined here. Presumably, this exact analysis could've been achieved by signing up for a "Basic" plan, which has a $500 one-time setup fee and a monthly cost of $500. It supports an image collection size up to 20,000 images and supports 30,000 searches – both of which would've been enough to perform the analysis outlined here. It's almost certain that the TinEye team will have better ideas on how to perform this analysis in the most cost-efficient manner possible. ↩

About John Resig

John Resig is the creator of the Ukiyo-e.org Japanese woodblock print database and search engine. He develops tools to aid in the research of Ukiyo-e and other art history subjects. A Visiting Researcher at Ritsumeikan University, he recently presented at the 2013 Japanese Association for Digital Humanities conference in Kyoto, the Japanese Art Society of America, and the Digital Humanities 2014 conference. Mr. Resig is the Head of Computer Science at Khan Academy and is a renowned computer programmer, having created the jQuery JavaScript library used by over two-thirds of all web sites. He has also published two books on JavaScript programming: *Pro JavaScript Techniques* and *Secrets of the JavaScript Ninja*.