Application of Image Recognition in the Field of Museum Digital Content

Yixuan Li

Media, Arts and Technology, UCSB

School of Data Science, Fudan University

This is the English Translation Version of the original paper

Table of Contents

Abstract

In the era when everything enters artificial intelligence, the museum industry has gradually begun to apply artificial intelligence technology for museum management, curating, and guided tours. This article hopes to provide a certain degree of image recognition technology reference and application exploration for the museum staff from the technical perspective of the intersection of computer and museum. This article takes the most widely used image recognition technology in artificial intelligence technology as an example to analyze the development law and multiple application scenarios of image recognition technology in the museum industry. Combining the latest technology in the field of artistic image recognition, as well as the subject and many domestic and foreign image recognition projects that have been implemented, it summarizes the three advantages of "small data", "big data" and interactive modes of image recognition technology in digital content curation. Meanwhile, the paper demonstrates the egoism and altruism of the application of image recognition technology in museums, analyzes the challenges and development directions that still exist in the development of this technology in the museum field, and puts forward feasibility suggestions for the current stage of museum image recognition project construction.

Keywords: Artificial Intelligence, Image Recognition, Museum Informatics, Museum Collections

Application of Image Recognition in the Field of Museum Digital Content

The development of Computer Vision and Image Recognition

If vision is the most important channel for humans to obtain information, then computer vision analysis is one of the most important ways for computers to help humans process information. Compared with other forms of information (such as text information), the storage method of image information is more complex, the information complexity is higher, the connotation is richer, the application scenarios are wider, and the business value is higher. Therefore, computer vision is the most invested and fastest developing branch of artificial intelligence in recent years. Image recognition is one of the most basic and popular, and also the direction with the highest achievement landing rate.

Image recognition is a process in which the computer distinguishes and locates the objects according to the observed images and makes a meaningful judgment. Since the 1950s, image recognition has experienced three stages: character recognition, digital image processing, and recognition, and object recognition (Zhang, 2010). It has been widely used in traffic, medicine, and other research and commercial fields. Inspired by the famous Hubel-Wiesel biological vision model, most of the current image recognition is based on the construction of artificial neural networks (Yu et. Al, 2013). Through the training of a large number of data sets with labeled information and the application of deep learning frameworks, a model suitable for business scenarios is obtained.

How can Image Recognition Be Applied in the Field of Museums?

Behind a collection, there is often multi-level knowledge covering art, culture, history, science, etc., and there is often a logical relationship such as evolution and reference with other collections. With the advancement of museum informatization, digital collection resources

including photographs and images of collections, 3D models of cultural relics, and digital exhibits continue to accumulate and accumulate. These massive digital collection resources have greatly changed the way museums manage and display the original collections (Qin, 2015). As the channels for obtaining information are becoming more and more extensive, the public has higher requirements for the quality and quantity of information that museums can provide. The mining of information, the creation of related networks, and the display of knowledge are far beyond the reach of human resources. Computer systems are urgently needed to provide more advanced data management and mining methods. In 2017, the American Museum Alliance pointed out: “Artificial intelligence will become an indispensable tool for museums to manage huge-scale data in the 21st century.” (Zhuang 2018) Image recognition technology exactly meets the research demands of museums.

The initial stage of image recognition application -- visitor guide and venue management

Although the research of image recognition in the field of computer is quite mature, the introduction of image recognition is still a relatively new idea in the field of the museum. In recent years, scholars have made many attempts in the field of the museum for image recognition. In the early stage, image recognition is applied to the visitors' guide. Visitors can scan exhibits directly by using mobile devices in the browsing process of museums, to obtain more information about exhibits. In the early stage, the app was an off-line stand-alone mode. Visitors used the mobile navigation device developed by the museum to take photos during the visit, so as to obtain further information about the exhibits (Bay et. Al, 2006).

With the development of computer technology, in the medium-term stage, more museums choose to develop their app programs or cooperate with visual service providers to establish image scanning software based on personal mobile terminals. Users can scan the artworks in the

museum with mobile phones to get more information about the images. In this stage, there are many independent applications in each museum, that is, the application is used by the monomer museum and only provides the art information in the museum, and does not communicate with other museums. For example, the Australian National Portrait Gallery cooperated with catchoom, a visual technology provider, to create an interactive application "Headhunt!"

Another more advanced and universal solution for museums is to encourage visitors to use independent third-party applications, such as Smartify App. The application has established a cooperative relationship with famous museums such as the Louvre and the Metropolitan Museum. It can obtain collection information from different institutions, and users can access the digital collection resources of the organization. Users can simultaneously browse and compare the collections from various exhibition halls. At the same time, this cross-agency information platform can also provide visitors and researchers with more comprehensive exhibition knowledge, and save a lot of development and maintenance costs for the museums (Hu, 2020).

Image recognition technology also provides a further guarantee for the security management of the museum. Through face recognition, venue monitoring, and other ways, the abnormal situation in the monitoring area can be identified and the network alarm can be given, so as to provide a further guarantee to the museum with intelligent security (Zhou, 2002).

Collection Research and Digital Content Curation

With the gradual maturity of the image recognition technology in the field of artworks and the gradual development of disciplines such as digital humanity, computational aesthetics, etc., more humanity scholars cooperate with experts in the computer field, the application value of this technology continues to increase, and more scholars use the technology as a new approach and new perspective for collection research and hope to gain further understanding of collections

through computer vision research methods. Research projects that have been carried out include collection pattern recognition, collection authentication, painting analysis, and the like (Zhuang, 2018). In 2015, Elgammal et. al used image recognition and analysis technologies to establish a model to analyze the similarity and citation logic between artworks and calculate the most creative artworks in the art history, which provided new ideas and methods for researches of how to evaluate the creativity of artworks (2015).

In recent years, more museum institutions have further opened up the digital information copyrights of their exhibits, and the digital content display has also become an important part of museum exhibitions. Image recognition technology is widely used in exhibition planning and digital content presentation due to the following characteristics.

Aspect 1. The image recognition technology can provide micro-cognition and carry out focusing and comparison in terms of the "small data" of a certain element (such as texture, pattern, etc.) across works, time, and space. For example, at the Shanghai Museum Dong Qichang Calligraphy and Painting Art Exhibition, through image recognition, mountains, water, and other elements in the author's calligraphy and painting works are located (as shown in Figure 1), and the audience can search for a certain element to obtain all the images that include the element so that more visual and profound understanding of the writer's portrayal style in a certain detail can be achieved (Liu, 2019).

Aspect 2. The image recognition technology can provide macro-cognition and carry out a general display of artworks in terms of different styles and different eras so that audience can have a more macroscopic cognition of the art history. With the continuous advancement of museum digitization, the number of online museums on the market has grown rapidly, but the number and quality of "big data"-level interpretation projects for such a large number of online

digital resources are still not promising. The greatest charm of the image recognition technology is to abstract valuable macro-law information from massive image data and realize the comparison of horizontal and vertical image information in any dimension and depth. For example, Google used to identify the main colors of works and calculate the similarity of the works. Through the t-SNE algorithm, according to the similarity of images, Google integrated and arranged 6 million images of artworks in its database to form a visually beautiful artwork display space so that audience can get a macro understanding of these 6 million images of artworks in a short time, understand the evolution of the color scheme of the artworks and compare the differences between different art styles. Figure 2 shows the interactive display results of 6 million pictures based on color similarity.

Aspect 3. The image recognition technology can provide rich interactive modes. Because of the fast calculation speed of the image recognition technology, the simple acquisition of required data (images are sufficient), and the convenience of online or mobile device operation, the technology can provide museums with rich interactive activities and improve the sense of participation of museum visitors online and offline. For example, in the “Jinshan Farmer Paintings” area of the “Meeting Shanghai” section of the CIIE, the audience can draw figures, houses, chickens, and other images in simple strokes according to prompts, through identifying the composition and outlines of the drawn strokes, similar elements and composition frames in the database are matched, and finally, a "Jinshan Farmer Painting" that conforms to artistic logic (Figure 3) is generated within 0.5 second calculation time. Another example is Art Selfie, a popular application launched by Google where users can take personal selfies and upload them, the application uses image recognition and similarity matching technologies to search for images according to their selfies, matches the most similar artistic images with the uploaded images in

the database of more than 10,000 artworks and sends back to users, which allows users to further explore the artworks by "finding the artworks most like you" (Figure 4). This kind of interactive activity can enhance the audience's participation and interest in the exhibition. At the same time, these activities are not limited to the offline exhibition form, extend the actual exhibition participation time of the audience, and are the embodiment of the advantages of the image recognition technology.

Based on the above three advantages, the image recognition technology provides the audience with rich interactive modes from the two dimensions of "small data" and "big data", is capable of becoming one of the important guiding techniques for museum curation, integrates resources and optimizes presentation methods so as to plan a digital exhibition which can give full play to the capabilities of the image recognition technology.

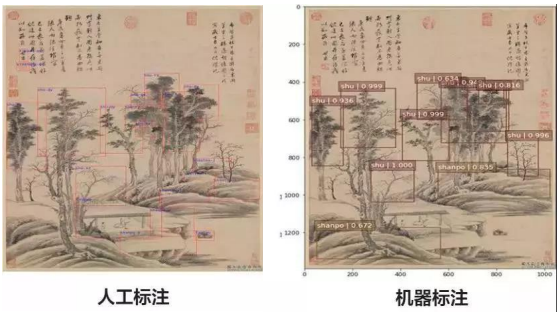

Figure 1. Object Detection of Dong Qichang Exhibition

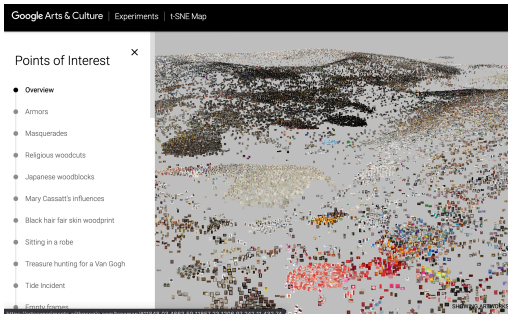

Figure 2. t-SNE algorithm result

Figure 3. Project Jinshan Farmer Painting

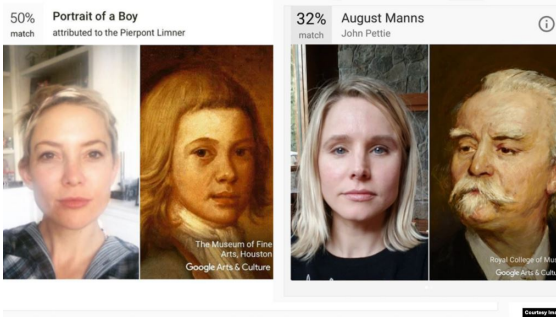

Figure 4. Art Selfie matching result

Image Recognition and Digital Image Collection Resource Management

The management of digital image collection resources relies on the following indicators and functions: standardized information labeling methods, database construction, and management, call and circulation between homogeneous and heterogeneous databases, efficient front-end information display interfaces, and the like. The International Image Interoperability Framework (IIIF) is a technical standard launched by the British Library, the Oxford University Library, and other institutions to define a set of standardized resource transmission interfaces (API), provide unified display and use for digital online resources with images as carriers and promote the exchange and sharing of image resources between heterogeneous platforms and resource libraries (Fu, 2020). However, the existing IIIF framework still relies on a large amount of human labeling, and in the image-level and object-level labeling work in the framework, the image recognition technology can play its auxiliary role to further meet the digital needs of museum management.

The Shanghai Library uses IIIF and the knowledge map to integrate image resources, uses image recognition technology to further retrieve images and discover knowledge, and proposes a set of new solutions for image resource integration and knowledge discovery (Zhang et.al, 2020). Organizing the current application of the image recognition module in the IIIF labeling framework can obtain the result as shown in Figure 5. Adding the creating style label to the first-level labeling level (picture level) and introducing the style classification model in image recognition to predict the image style, and the subjective style classification standard is objectively standardized to a model unified output result. At the second-level labeling level (element level), according to the NLP processing results of the style classification label and the work description, different pre-trained models are matched for object recognition and text

content recognition (OCR), and corresponding labels are generated to improve the accuracy of model prediction results. At the third-level labeling level (semantic level), through multi-angle processing of the object recognition and text recognition results, semantic level information is generated and forms the third-level labeling content together with the artificial semantic labeling results. The IIIF labeling framework added to the image recognition module objectifies the subjective evaluation method of the work style, which is conducive to the unification of classification standards across databases, and at the same time, the image recognition results of each step will be iterated to the next labeling level, thereby improving the uniformity, stability, and accuracy of the overall labeling results. At the same time, the labeling results and image recognition capability of "searching for images with images" also provide more convenience and possibilities for the query and management of digital collection resources.

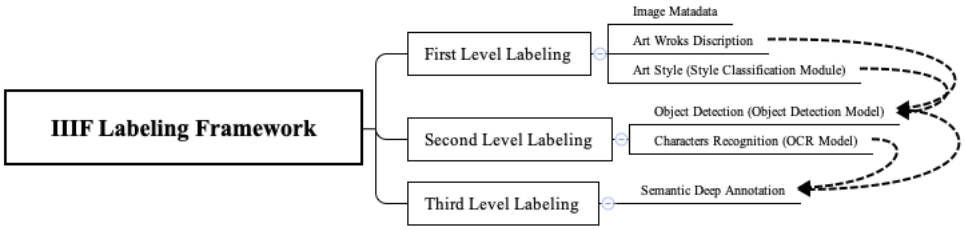

Figure 5. Combination mode of IIIF Labeling Framework and Image Recognition

It can be seen that image recognition is a technology that takes into account the egoism and altruism of museums, and the functionality and operability of image recognition are worthy of further development and exploration by the museum academic community.

The Challenge and Realization of Image Recognition in Artworks

Although the image recognition technology has gradually matured in the field of physical object recognition, there is still a lot of room for development in the accuracy of image recognition and the wide range of application in the field of artworks. Taking object recognition

in image recognition as an example, Figure 5 shows the results of the object recognition program for artworks. There are several reasons for the slow development of object recognition in artworks：

Aspect1

The training data sets which the generation of mature image recognition networks relies on (the industry-standard data sets are ImageNet, PASCAL VOC, etc.) are all natural images (a natural image which means photos of things that exist in life) such as cars, pedestrians, foods, animals, etc., and advanced models and algorithms are also optimized based on such networks. Therefore, network models that can be directly applied to identify elements (such as mountains, animals, Buddha statues, etc.) in artworks are very rare, and most of them are still in the development and debugging stage and are not capable of being applied in real life;

Aspect2

Although large image data sets such as ImageNet and PASCAL VOC contain a wide range of object categories and can cover most business backgrounds, the data sets do not include iconographic figures such as Christ and Buddha statues and the like which are often very important virtual characters in artworks like fine arts, so if a neural network that can recognize iconographic figures wants to be built, starting building a huge data set is needed, which will consume a lot of manpower and material resources.

Aspect3

As a creative art, the presentation of the contents of artworks varies greatly among different authors and different genres. For example, in the theme of “the Annunciation of the Lord”, Gabriel is sometimes portrayed as males and sometimes females (Madhu et. al, 2019). A more accurate recognition system requires several times more training data than the natural

object recognition system, which increases the difficulty of image recognition. At the same time, copyright and clarity issues are also needed to be resolved in the development of this type of recognition network.

In the field of computer researches, the recognition of artworks belongs to the field of cross-domain recognition. The ideas are roughly divided into the following categories: the first is to transfer the networks trained on natural images and fine-tune the last few layers of the neural networks to get the required prediction category; the second is to conduct style transfer on the original training datasets (ImageNet, PASCAL VOC, etc.) according to the painting style in the artwork dataset that one wants to predict, transform the original natural image into an image similar to the artwork through style transfer (as shown in Figure 6), and retrain the original network to obtain a model that can predict fine art images. More advanced algorithms in this industry are capable of identifying and detecting seven figures such as the Virgin Mary, the Crucifixion, angels, etc. with an accuracy of about 59.2% and a speed of 1000 images per minute (running on the GPU), but the result is still far from being applied for commercial use in real life (Gonthie et. al, 2019).

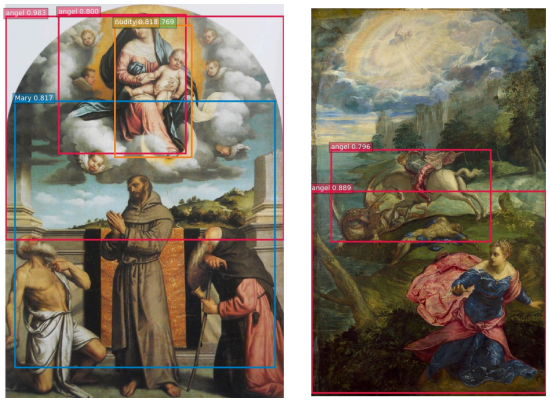

Figure 6. Object Detection Result

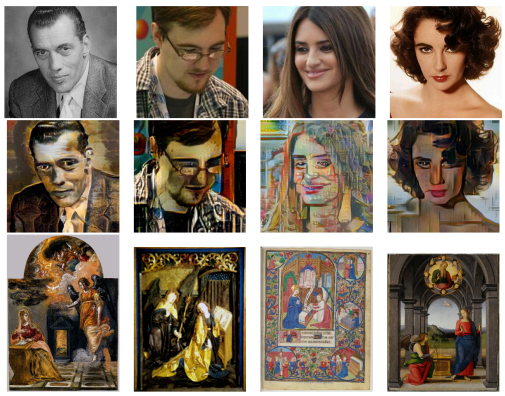

Figure 7. Face images in row 1 with their style-transferred counterparts in row 2 and their corresponding style images in row 3

Public API and its usage

In the field of practical application, since few museum institutions have full-time data scientists who understand the operation and management of museums, it is very unrealistic to rely on a monomer museum to establish a valuable object recognition network (Zhuang, 2018). The unified development is implemented by a mature technology platform, and an application programming interface (API) is provided to each museum. The business model of a monomer museum for network migration learning has become more valuable for development.

Take Google arts & culture as an example. As a non-profit art project, since 2011, the program has cooperated with more than 2000 museum organizations to display more than 6 million high-definition art images and videos from more than 80 countries. Based on the broader database and more powerful computing resources in the Google platform, Google has established Arts & Culture Experiments Platform. Engineers can use Google API and data resources to carry out a series of art exploration. In the platform, there are 6 landing art projects with more than 100 themes. For example, in the project of X Degrees of Separation, the evolution path between any two images selected by users can be found in more than 6 million images by means of image similarity analysis (as shown in Figure 7). Another example is the RUNWAY PALETTE project, which shows a variety of color schemes of designers through the extraction of main colors and similarity recognition of 140000 top fashion show clothing images (as shown in Figure 8).

The above experiments have stripped away the metadata of Artworks (such as the author, age, and other information), and simply use computer vision analysis method, based on Google's powerful database and algorithm, so that more valuable image recognition art projects have landing value and possibility. This kind of open platform also provides an API interface to museums, so that more museums have the ability to establish landing digital image projects

based on domain data and computer vision algorithms by the premise of lower cost. Museums should try to cooperate with international digital platforms to join the ranks of digital museums more conveniently.

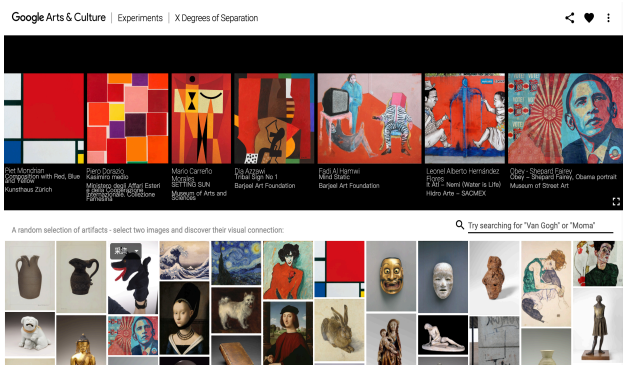

Figure 8. Project X-degree separation

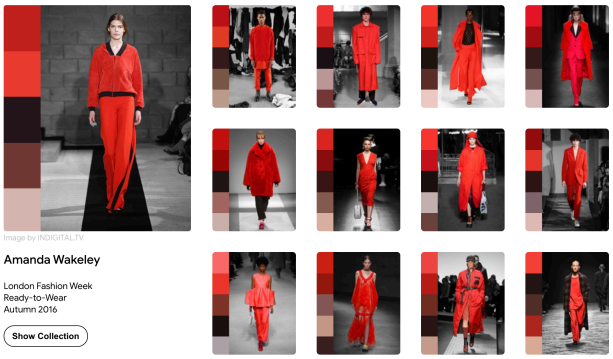

Figure 9. Project RUNWAY PALETTE

Conclusion

This paper first introduces the development of computer vision and image recognition technology to readers and then sorts the development of image recognition technology in museum applications. It mainly introduces the development situation and potential of image recognition technology in visitor guide, digital content curation, and collection management. Combined with several examples, it analyzes three key capabilities of image processing technology in digital content curation. The third part focuses on the problems and challenges encountered in the application of image processing technology in the museum art field, as well as two different solutions in the research field and the commercial field. Artificial intelligence and its image processing technology have brought a new wave of development to museums, and this year's special situation has also strongly promoted the development of museum informatization. Seizing the opportunity, daring to break through and innovate, actively opening up to new technologies and new challenges, are the keys to promote the digitization process of museums in China. At the same time, the promotion of museum digitization will further improve the technical

requirements of the combination of computer and museum field. To promote the development of interdisciplinary disciplines and develop interdisciplinary talents is also an urgent need in today's era of interconnection and intelligence of all things.

References

Bay, H., Fasel, B. & Van G.L. (2006). Interactive museum guide: Fast and robust recognition of museum objects.

Elgammal, A., & Saleh, B.(2015). Quantifying Creativity in Art Networks. Sixth International Conference on Computational Creativity (ICCC) June 29-July 2 2015, Park City, Utah, USA.

Fu Y. (2020) IIIF and its application in digital resource integration.Library Forum,2020,40(04):159-166.

Gonthier, N., Gousseau, Y., Ladjal, S., & Bonfait, O. (2019) Weakly Supervised Object Detection in Artworks. In: Leal-Taixé L., Roth S. (eds) Computer Vision – ECCV 2018 Workshops. ECCV 2018. Lecture Notes in Computer Science, vol 11130. Springer, Cham. https://doi.org/10.1007/978-3-030-11012-3_53

Hu, L. (2020) Image Recognition Technology Use In Museums: https://amt-lab.org/blog/2020/1/image-recognition-technology-in-museums

Liu, J.(2019). Exploration and Practice of Museum Data Visualization——Taking the Digital Construction of Shanghai Museum as an Example. Museum, 2019(02): 91-97

Madhu, P., Kosti, R., Mührenberg, L. Bell, P., Maier, A., & Christlein, V. (2019). Recognizing Characters in Art History Using Deep Learning. 15-22. 10.1145/3347317.3357242.

Yu K., Jia L., Chen Y., & Xu W. (2013). Yesterday, today and tomorrow of deep learning. Computer Research and Development,2013,50(09):1799-1804.

Zhang J., Liu W., Yu J., & Chen T. (2020). Research on Seal Resource Integration and Knowledge Discovery Based on IIIF and Semantic Knowledge Graph. Library and Information Service,2020,64(07):127-135.

Zhou, X. (2002). Intelligent security system for fire prevention and theft based on image recognition. Shanghai : Shanghai Jiaotong University,2002

Zhuang Y. (2018). Discussion on the potential applications of computer vision and machine learning in the research and display of museum collections. Southeast Culture,2018(S1):87-91.

Catchoom :https://catchoom.com/

Google arts & culture：https://artsandculture.google.com/

Tezign. CIIE. Jinshan farmers painting: https://www.tezign.com/designer/#/share/ys22rlan5/case