The Burgeoning Computer-Art Symbiosis

Computers help us understand art. Art helps us teach computers.

By Shiry Ginosar, Xi Shen, Karan Dwivedi, Elizabeth Honig, and Mathieu Aubry
DOI: 10.1145/3186655

During the Renaissance, artists often employed the use of standard templates, or pattern drawings, and models to aid in the production of their art. Such drawings could be used for many parts of an image: the intricate design on a brocade cloth, the full figure of a flying angel, or the commonly occurring horse and cart on a roadway. Drawings would be used and reused within the workshop—both by the master and by his assistants—thereby standardizing the look of the workshop's output. The drawings could be traced, pricked with pinholes for transfer, or simply visually copied. In any case, the forms produced were very consistent. Patrons were aware of this reproductive practice, and occasionally complained about it. For example, Pietro Perugino, Raphael's teacher, was criticized for using the exact same angels in several altarpieces. But once artworks were being produced for an open market, or for patrons at great distances from one another, repetition was less likely to be noticed. Repetition slowly became a more attractive option for painters, who were sometimes producing many similar works for a broad art market. One such painter was Jan Brueghel. His elite patrons lived all over Europe, and he also had buyers among the ordinary people in Northern European cities.

Tracing the patterns used by artists like Brueghel is of interest to art historians in several ways. It gives us insight into production methods used in an artist's studio. It also lets us see how forms could be used over time, and in different contexts, revealing the development of the artist's visual thinking. On a more concrete level, it can help us determine which works were produced within the workshop, and which are copies made later by outsiders—ironically, exact repetition is a likely sign of authenticity. For an artist like Brueghel, to whom literally thousands of works have been—often wrongly—attributed, going through his whole oeuvre to find identical visual elements presents an enormous challenge. Moreover, the human eye quickly sees identity where only similarity exists. Two angels that look virtually the same may, in fact, have subtle differences in outline. A cart may be drawn on 20 different roads, but at ever-so-slightly different angles. Such minor disparities rule out the

Detail of "The Entry of the Animals into Noah's Ark" by Jan Brueghel the Elder.

Image by Everett - Art / Shutterstock.com

use of templated patterns, but it's hard to detect them. It is difficult to differentiate precise repetition from mere similarity with the naked eye.

Computer vision algorithms, however, excel at precisely the task of identifying repetitive patterns in data. Knowing this, we used the latest toolset in machine learning—convolutional neural networks (CNNs)—to develop a simple method to perform this task automatically. At its core, a CNN is nothing more than a succession of simple operations—such as convolutions, non-linearities, and local summations—that are applied to an image to perform, what is very often, a classification task. The parameters of these operations are learned from thousands of training examples. CNNs provide excellent results on most computer vision tasks, and their intermediate representations have shown a high degree of generality, making them easy to use.

Our method starts with an image region selected by the user. We compute the representation of this region using an AlexNet, a CNN that was originally trained to classify a set of one thousand objects in natural images. We then compare this representation to image patches in the entire dataset, and return the most similar ones. This comparison can be done quite efficiently using a sliding-window technique. The results usually yield false positives. However, it is much faster, and easier, for a human to verify several hundred matches proposed by a computer, than to manually go through thousands of paintings in enough detail to note minor features that could very easily

Figure 1. Asked to find windmills like the one annotated in blue in the left image, a human annotator, especially if asked to look through thousands of images, could easily miss the one in the central image, which was identified by our method.

be missed. Our method of repetitive region proposal can easily be coupled with human annotations, or applied in an active setting where matches that are selected as correct by the art historian are used for query expansion (see Figures 1 and 2).

Current computer vision models are sensitive to low-level image statistics like line width, color, and texture, and are only rarely able to detect repetitions of objects across artistic mediums. Humans, on the other hand, can easily identify the same object whether it appears in an oil painting or a pencil drawing. The next stage of our project seeks to use annotations by art historians to make our model invariant to these low-level statistics of object representation.

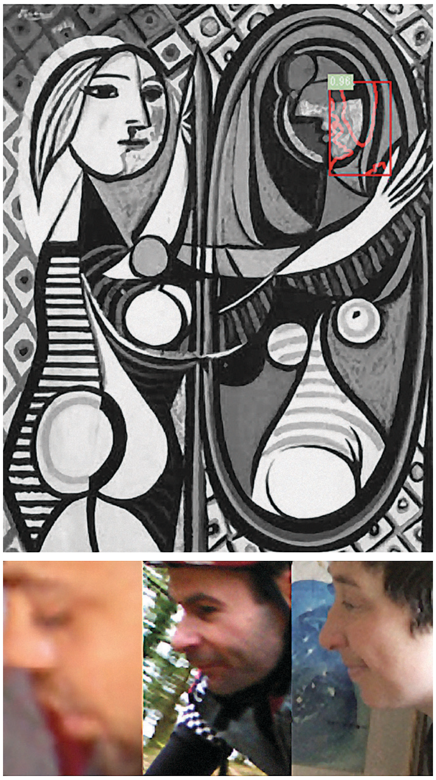
Invariance to low-level statistics, however, is only one type of generalization. Another type is the ability to recognize the same object depicted at different levels of abstraction—from the concrete and realistic, to the symbolic. Here again, humans hold the advantage over computers. It may even be argued that using abstract forms to communicate ideas visually is a direct consequence of the specifically human perceptive quality that allows us to extrapolate beyond the concrete. Thus, humans are uniquely adept at creating and understanding abstract forms. Computers, however, are not.

An exciting opportunity lies in this gap. If our goal is to model the human perceptive system, then abstract art can be viewed as the ideal tool for understanding the capabilities of perception and as a benchmark for testing our artificial models.

To demonstrate the use of abstract art as a benchmark for the generalization capabilities of computer vision algorithms, we tested several object-detection methods on Pablo Picasso paintings [1]. We collected a set of the maestro's paintings depicting human figures, ranging from realistic to abstract. We then ran out-of-the-box, human-detection systems to see if they could correctly identify all hu-

Figure 2. Starting from a query patch (shown in the leftmost column of each row) the top seven matches from our method show similar content across different styles and in different contexts.

Figure 3. By identifying discriminative parts, like faces, object detection methods based on part-based representations can correctly detect human figures in Picasso paintings (bottom) [5].



man forms shown in the paintings. Not unexpectedly, we found while all systems worked well in the realistic case, only methods that allowed part-based representations of objects were successful with the more abstract cubist examples (see Figure 3). We concluded that less rigid representations are more suitable for modeling humans' ability to generalize.

Our results have already inspired others to design computer vision algorithms that try to mimic the generalization capabilities of human perception, testing their systems, not only on natural images, but also on our Picasso painting dataset and other abstract works of art [2]. These developments show the interaction between computer science and art can prove fruitful in both directions.

Computer vision provides a tool for analyzing works of art. Works of art can act as a tool to analyze the workings of computer vision algorithms. While we have illustrated this idea with examples of pattern and object recognition tasks, the same holds true for other tasks. For instance, researchers have shown how automatic methods can help historians reason about the influence of color on the economic value of a piece of art [3]. On the other hand, by learning the statistics of current art, and producing a diverse set of images based on them, computer vision algorithms could help boost creativity [4]. They may even generate "art." We believe this mutually beneficial interaction to be one of the most exciting, and potentially fertile areas of research in both disciplines.

ACKNOWLEDGMENTS

This work was partly supported by National Science Foundation Graduate Research Fellowship under Grant No. DGE 1106400, by FBF grant #2018-0017 and ANR project EnHerit ANR-17-CE23-0008.

References

[1] Ginosar, S. Haas, D., Brown, T., and Malik, J. Detecting people in cubist art. In *Visart Workshop on Computer Vision for Art Analysis, European Conference on Computer Vision* (ECCV), 2014.

[2] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, 779-788.

[3] Pownall, R., and Graddy, K. Pricing color intensity and lightness in contemporary art auctions. *Research in Economics* 70, 3 (2016). 412-420.

[4] Elgammal, A., Bingchen, L., Mohamed, E., and Mazzone, M. CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms. In *Proceedings of the Eighth International Conference on Computational Creativity* (ICCC), 2017.

[5] Crowley, E. J., and Zisserman, A. The state of the art: Object retrieval in paintings using discriminative regions. In *British Machine Vision Conference* (BMVC), 2014.

Biographies

Shiry Ginosar is a Ph.D. student in the computer science department at U.C. Berkeley. She works in the intersection of computer vision, artificial intelligence and, sometimes, art.

Xi Shen is a Ph.D. student in the LIGM (UMR 8049) of École des Ponts ParisTech in France. His research interests include weakly supervised learning and unsupervised learning in computer vision.

Karan Dwivedi is an undergraduate computer science student at the Indian Institute of Technology, Delhi. He worked on instance matching and object discovery in artwork at École des Ponts ParisTech.

Elizabeth Honig is a professor in the History of Art department at U.C. Berkeley. She has published two books on Flemish painting, most recently *Jan Brueghel and the Senses of Scale* (Penn State, 2017), and is the creator and manager of a group of research websites about the Brueghel family, including janbrueghel.net.

Mathieu Aubry is a computer vision researcher at École des Ponts ParisTech, working in the joint LIGM (UMR 8049) lab with UPEM.

1528-4972/18/03 $15.00