Paper Dreams: An Interactive Interface for Generative Visual Expression

Guillermo Bernal
Media Lab
MIT
Cambridge, MA
gbernal@mit.edu

Lily Zhou
EECS
MIT
Cambridge, MA
lilyzhou@mit.edu

Haripriya Mehta
EECS
MIT
Cambridge, MA
hmehta@mit.edu

Pattie Maes
Media Lab
MIT
Cambridge, MA
maes@mit.edu

Abstract

Paper Dreams focuses on how current state-of-the-art machine learning algorithms and multi-modal input technologies can be integrated to augment human creativity and empower a wide range of users. We present a web application that promotes a highly dynamic, interactive back-and-forth between a human artist and the Paper Dreams system, incorporating customized machine learning models and representations to assist the user's visual expression. The Paper Dreams system processes the sketch the user generates, "recognizes" what has been drawn, and creates personalized suggestions for new elements and colors. Papers Dreams aims to inspire and empower individuals who believe they are not "the creative type", are experiencing an artistic block, or are looking to express themselves in new ways.

1 Introduction

In today's highly visual world, it can be frustrating for those who experience reduced motor or artistic skills, such as children and the elderly, to easily create drawings and illustrations. While art therapy has been proven to be a beneficial activity for many of these groups[1][2][3], there are relatively few tools available to support and encourage this medium.

Paper Dreams focuses on the exploration of creativity support tools, its intersection with modern technology, and the implications of what it means for people to have a tool to support them during their creative endeavours. The system encourages creative exploration by simulating the experience of the artist engaging in a dialogue or creative brainstorm session with a fellow artist, suggesting novel visual elements and color associations to the user based on the current landscape of their sketch.

2 System Design

Paper Dreams is a web-based canvas for sketching and storyboarding, with a multi-modal user interface integrated with a texturizer conditional GAN (Pix2PixHD[5]), an eight-layer sketch recognition CNN (inspired by sketch-A-Net[6]), and a natural language processing library (sense2vec[7]) for determining relations between objects. The Paper Dreams system currently augments the drawing experience in three ways: adding textures/colors, suggesting other elements/drawings for the scene, and introducing serendipity.

2.1 Recognizing Sketches

Paper Dreams responds to user stroke inputs by rapidly performing sketch recognition in addition to customized coloring/texturizing of the image. The current architecture of the sketch recognition

33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.

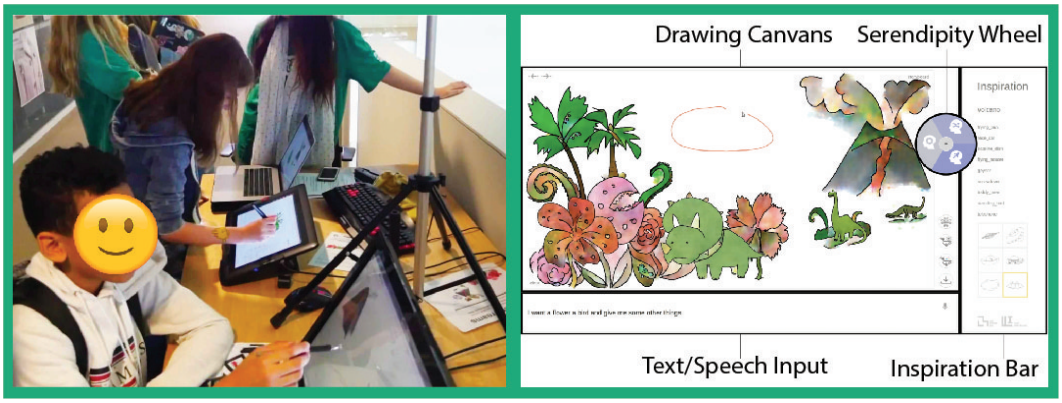

Figure 1: a) Paper Dreams setup during pilot studies with middle school students. b) Overall user interface, with important subsections marked.

model is based on that of Sketch-a-Net, a deep neural network model that beats human recognition performance by 1.8% on the TU-Berlin dataset (a large-scale benchmark dataset of human-sketched images). We trained our convolutional neural network on an aggregate dataset comprised of the TU-Berlin dataset, the Sketchy Dataset, and our own collected dataset of sketches, augmented with Augmentor[10]. We found that we had a 75% accuracy rate for completed drawings across those 125 classes; however, while we can use our architecture for recognizing incomplete or "partial" sketches, the resulting labels are often incorrect until enough defining features are drawn.

2.2 Coloring with GANs

Paper Dreams currently implements pix2pixHD's architecture for the texturizing network, with minimal changes to NVIDIA's original algorithmic design. The Paper Dreams system has fifteen distinct trained texturization models, each encompassing a relevant subset of the classes from Creative Commons illustrations, downloaded using the library Beautiful Soup [8]. The sketch-identified label (e.g., "hedgehog") is associated with a model (e.g., "animal"), and then the active user sketch is processed by that model to return an appropriate texture. Examples of other models include "plants", "buildings", "transportation", and "fruit."

2.3 Making connections

Paper Dreams uses sense2vec [9] to extract multiple possible meanings ("senses") and calculate cosine similarity values for each pair-wise combination of the classes in our dataset. The serendipity dial then uses these similarity values to influence the system's suggestions- if the user turns up the serendipity, then the suggested colors and concepts would be more unrelated to the current concept. The user can then include pre-made sketches from a suggested concept to their canvas, with full control of the size and location of each concept. The serendipity wheel aims to encourage the user towards further divergent thinking.

3 Results and Discussion

We have run two pilot studies on approximately 50 individuals (children and adults) and found that users have generally found Paper Dreams very engaging and interactive. Nearly all users gave positive feedback regarding the automatic coloring and related class suggestions, and many users (approximately 75%) wanted to know whether the application was available for their own devices. We found that young children quickly began to understand that different stroke patterns and shapes brought different colors from the texturizing aid, and that children noticed this association as early on as the second trial of sketching. Adults expressed that the level of suggestions provided by Paper Dreams allowed them to create connections that wouldn't have occurred naturally for them.

References

[1]Mihailidis, A., Blunsden, S., Boger, J., Richards, B., Zutis, K., Young, L., & Hoey, J. (2010). Towards the development of a technology for art therapy and dementia: Definition of needs and design constraints. The Arts in Psychotherapy, 37(4), 293-300.

[2]Lazar, A., Feuston, J. L., Edasis, C., & Piper, A. M. (2018, April). Making as expression: Informing design with people with complex communication needs through art therapy. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (p. 351). ACM.

[3] Johnson, C. M., & Sullivan-Marx, E. M. (2006). Art therapy: Using the creative process for healing and hope among African American older adults. Geriatric nursing, 27(5), 309-316.

[4]Magno, C. (2009). Explaining the Creative Mind. International Journal of Research & Review, 3.

[5]Wang, T. C., Liu, M. Y., Zhu, J. Y., Tao, A., Kautz, J., & Catanzaro, B. pix2pixHD: High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs.

[6]Yu, Q., Yang, Y., Song, Y. Z., Xiang, T., & Hospedales, T. (2015). Sketch-a-net that beats humans. arXiv preprint arXiv:1501.07873.

[7]Lu, Y. H. (2017). A sentence similarity calculation method based on Word2Vector and edit distance. Comput. Knowl. Technol, 13(5), 146-147.

[8]Richardson, L. (2013). Beautiful soup. Crummy: The Site.

[9]Trask, A., Michalak, P., & Liu, J. (2015). sense2vec-a fast and accurate method for word sense disambiguation in neural word embeddings. arXiv preprint arXiv:1511.06388.

[10]Bloice, M. D., Roth, P. M., Holzinger, A. (2019). Biomedical image augmentation using Augmentor. Bioinformatics.

4 Additional Results

During our pilot studies participants used the web-app to create visual work that range from stories, collages and abstract art. Below are a few examples of content generated by users during the pilot studies.

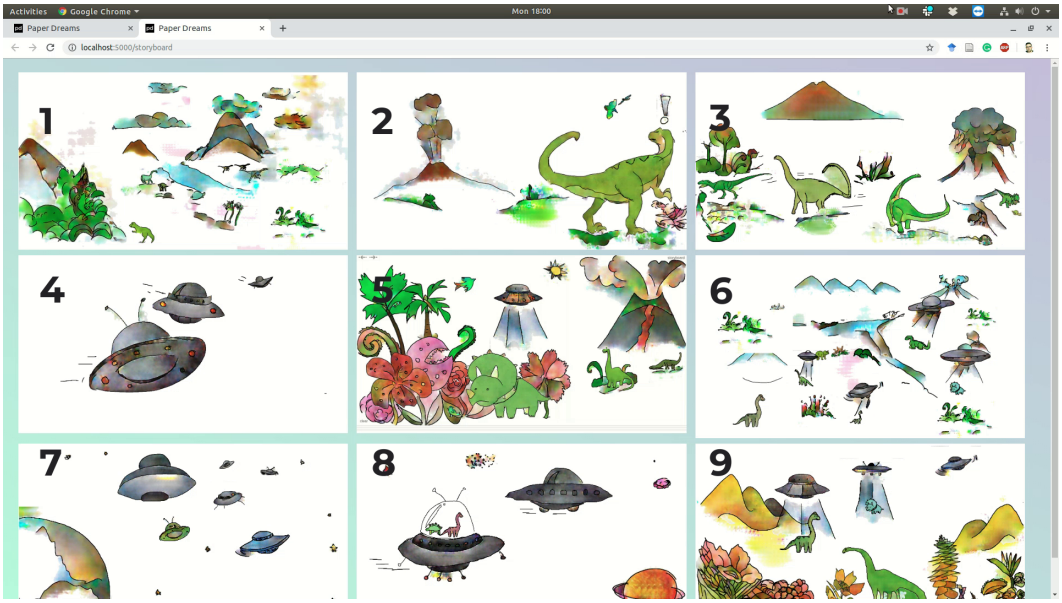

Figure 2: A screenshot from the storyboard feature as part of the Paper Dreams platform

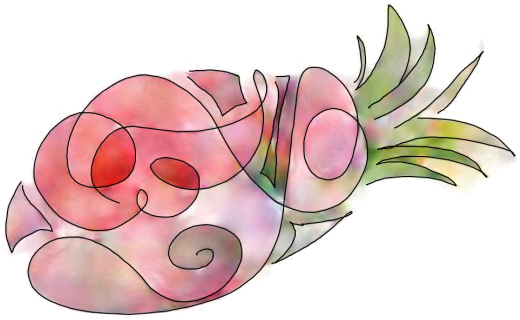

Figure 3: Pineapple Rose, named by author

Figure 4: Friendship, named by author

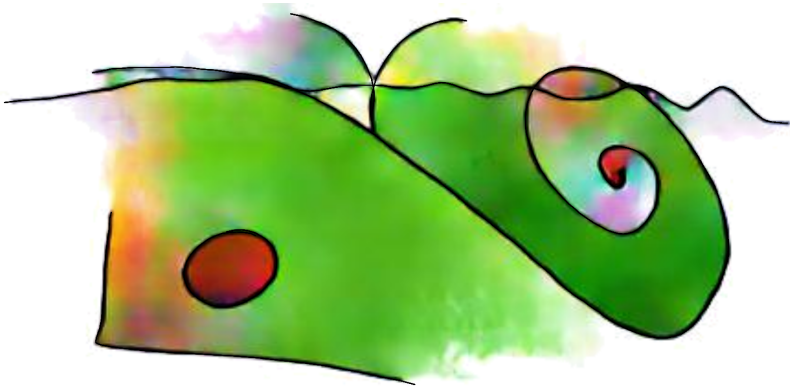

Figure 5: Do you see a whale, name by author

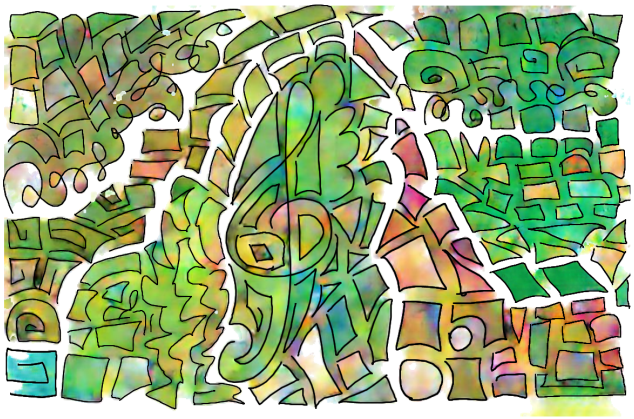

Figure 6: Jungle dream, named by author

Figure 7: Paper Dream's Origin of Species, named by author

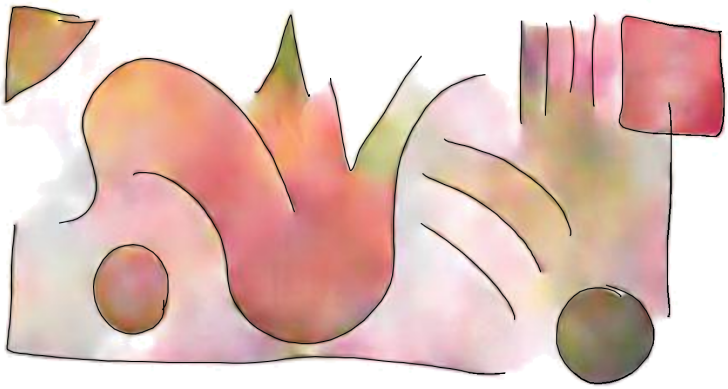

Figure 8: Untitled 2, named by author