Canvas Interface to Style Transfer

Mahika Dubey
Computational Media Department
University of California, Santa Cruz
mahika@ucsc.edu

Jasmine Otto
Computational Media Department
University of California, Santa Cruz
jtotto@ucsc.edu

ABSTRACT

This paper introduces in-browser applications for the application of style-transfer brushes onto an image. We present two distinct approaches to the creation of an application that invites 'casual creators,' and other nontechnical users to interact with pre-trained deep convolutional neural networks to co-create customized art. In the first approach, called Magic Markers, we give the users an experience that mimics painting with a brush on a canvas, such that they are able to 'paint' a style onto parts of their image through intuitive mouse selection and dragging over a canvas object. The second approach, Compositing Stamps, uses a real-time transfer method for applying style 'filters' to selected rectangular portions of an image. This process reveals to users some interesting features of the style transfer functions such as border artifacts, spatial stability, and multi-layering of different styles. The two applications provide new perspectives on a well-known algorithmic process, and enhances intuition for its expressive range, or lacks therein.

KEYWORDS

Style Transfer, Web Applications, Neural Networks, Tools for Artists

ACM Reference Format:
Mahika Dubey and Jasmine Otto. 2019. Canvas Interface to Style Transfer. In *Proceedings of ACM Conference (Conference'17).* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

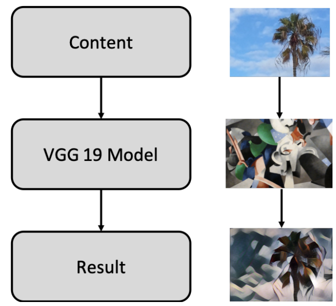

Figure 1: Flowchart of style transfer algorithm. Given a content image (or image selection), ML5 applies a trained VGG19 model to transform the content image to match as closely as possible the selected style. Here we have an example of a palm tree image rendered in the Udnie style.

1 INTRODUCTION

Neural style transfer is a technique for rendering arbitrary images in particular 'artistic styles', using an image transformation function encoding strokes, palette choices, and other perceptual properties of a work possessing the desired 'style'.

Our system gives users the experience to evaluate trained style transfer networks on their own images without requiring any technical or programming background. The interaction works off a playful tension between what defines explainable AI and creative tools for art.

2 BACKGROUND

2.1 Real-Time Style Transfer

Style transfer is a now commonly used method of applying style features of an image to the content of a new image to generate a similar result. Training is accomplished through a large Neural Network, built on the inspiration of the biological neuron connections in the human brain, to mimic visual recognition and understanding [7].

Though there have been many implementations of style transfer networks published, the technical details of the network presented by Gatys, Ecker, and Bethge [7] are most similar to the one used in our ml5.js backend [1]. The pre-trained models in our implementation utilize a VGG 19 network comprising of 3 convolutional layers (representing kernels), 5 pairs of convolutional layers (representing a difference between kernels), 2 transpose layers (representing stamp-like patterns of application of kernels), 1 more convolutional layer, and finally the activation function and normalization steps.

The loss function used for training takes the error between the Gram matrices (spatial autocorrelation) of the reconstructed image's $f$eature representation, and the training image's feature representation [8]. Johnson et. al. emphasize that these learned representations (in terms of spatial kernels) are not the original pixel data, and selecting on their closeness results in a better-trained network. Minimizing this 'perceptual loss' results in a complicated set of kernels (acting like motifs: brush strokes and contours) and spatially-dependent weights on them, that manage to encode the style image in a reduced number of dimensions. Running another image through this network likewise attempts to simplify its representation, and reconstruct. But because the reconstruction can only represent the learned style, it ends up being applied.

'Fast' VGG architectures for deep style transfer do not train on the content image, resulting in a style transfer network which works even on unseen images. Some implementations have even resulted in stable application of style transfer to frames of live video [5]. Therefore, we consider the technology mature enough to support a casual creator experience.

Conference'17, July 2017, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM.
https://doi.org/10.1145/nnnnnnn.nnnnnnn

2.2 Casual Creators

Casual creators, as identified by Compton and Mateas [6], are autotelic (self-motivating) tools to support creativity. Rather than the production of any particular result, such tools are designed to support the experience of using them.

Certain design patterns are known to reinforce these aims. Our system emphasizes alteration (users can upload their own image, allowing them to personalize it), annotation (users are learning a common set of trained 'style transfer' functions), and improvisation. Creative paralysis due to 'blank slate' is discouraged by the filter-like nature of our style transfer brushes: responding to content, yet being point-and-click.

Yet this form of play supports the study of a black box, be it the neural network, or the user's own artistic process. Interactive exploration blurs the boundaries between the two, creating an emerging process that is potentially both entertaining and educational.

Art creation is often thought of in terms of a single author making decisions, perhaps within a genre of content, using a style which is distinctively their own. Yet unwanted artifacts generated by image macros and photo filters is also consumed like art - as conferring status to the 'creative director', as well as promulgating the tool used to make these ephemeral, usually digital artifacts. The artistic use of trained networks further complicates the matter, not necessarily of authorship, but of the relationship between curation and creation [10].

Style transfer requires feature annotation followed by texture synthesis. State-of-the-art algorithms for the latter, combined with manual annotation, achieve 'style transfer' as well [9]. Our system explores mixed-initiative feature annotation through its real-time graphical interface. By incorporating multiple blending modes, we also give the user a little control over texture synthesis, in exchange for greater power than unskilled use of traditional, sprite-based digital brushes could achieve.

3 STYLE TRANSFER BRUSHES

We create two interactive web applications that allow users to play with applying the styles of different famous paintings onto an image of their own choosing. Users can isolate parts of their input image and select a pre-trained 'brush' that transforms the chosen section to the style of the brush.

We implement the user interface in Javascript using layered HTML5 canvases and D3 brushes. D3 is a powerful data visualization library that provides a massive customizable range of tools for programmers to develop graphical representations and mouse interactions in a website [3]. We use the functionality of a 'brush' to capture mouse movements across the canvas for pixel manipulation.

Our backend uses the style transfer architecture of ml5.js [1]. Ml5 is a simple machine learning built for the web on top of TensorFlow.js, aimed to improve the availability of common machine learning algorithms for developers and creative professionals of varying backgrounds.

3.1 Magic Markers

The Magic Marker application uses natural brush interactions so that users can 'paint' on different styles through intuitive selection and dragging. Figure 2 shows the initial state of the application

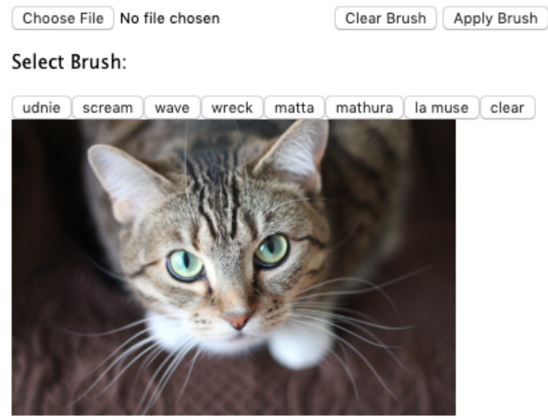

Figure 2: Magic Marker Application loaded with default image. Custom image upload button is available, as well as options to clear or apply a brush to the image. Brushes switch between a set of pre-trained style models based on user selection.

upon opening. Our default image of a cat can be used if the user does not wish to upload their own image [2].

Usage of the application is fairly simple and was built to mimic the physical space of a painter with a palette of 'colors' (replaced by styles) and the ability to undo recent changes before committing them to the canvas. A set of buttons lists the available pre-trained models, and clicking one begins the process of computing the styled image for that style and saving the information on a hidden layer underneath the main canvas image. This computation time is dependent on the size of the image, though the process is usually complete within a few seconds, as indicated by the change in the cursor from an arrow to the '+' symbol of an active brush. Rectangle shaped selections can then be made on the canvas, and dragged around or reshaped as needed to reveal the styled image in the 'painted' areas, creating a mask by removing pixels from the top layer.

Painted sections are permanently applied to the main canvas upon the click of the 'Apply Brush' button, which flatten the layered canvases such that the result of the masked style image is used to replace the underlying image content. If changes have not been applied yet, the 'Clear Brush' can be used to remove any un-applied 'paint' by resetting the top canvas layer through the reversal of the pixel removal. Figure 3 displays some interesting paintings made through the Magic Marker tool. Based on the varied size of the brush, users can change their level of details for different features, thereby adding an additional layer to the co-creation process.

3.2 Compositing Stamps

The Compositing Stamps application uses real time transfer of selected sections of an image to layer on patches of style to create interesting works of art resembling tiled mosaics or collages. Users can select different rectangular sections of their image of choice using a brush to get immediate transformation of content based

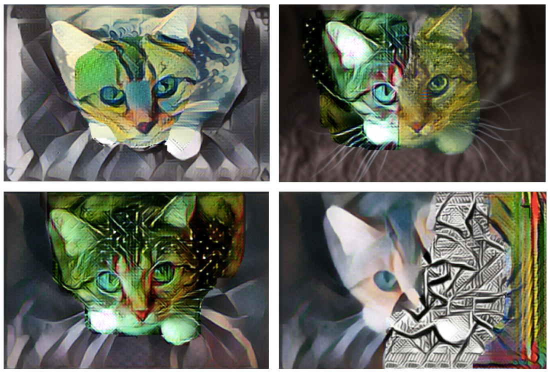

Figure 3: Results of using the Magic Marker Application loaded on default cat image. We combine multiple different styles and orderings to create creative new paintings and art pieces.

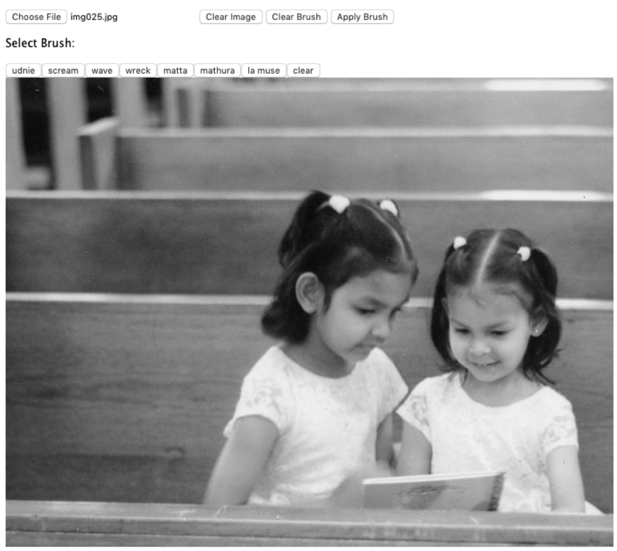

Figure 4: Compositing Stamps Application loaded with a custom image. Buttons are avilable to clear the image, clear the brush, or apply brushed changes to the image. Brushes switch between a set of pre-trained style models based on user selection.

on the selected style. Figure 4 demonstrates the initial state of the application when opened. Though similar to the Magic Marker page, image size is not as largely scaled in this implementation, and the application button process is different to reflect the unique patterning method. Due to browser limitations, it is recommended to avoid creating brush selections larger than 800x800 pixels.

Application usage simulates the physical application of stamps colored with paint. Select a brush to switch to the relevant style, and use the mouse to click and drag to select a rectangular shape. This patch will be replaced with a stylized version of the original content using the brush style selected by the user. Rectangular brush selections can be moved around the canvas, but the content will only be sent to the style transfer library on a mouse up event. ’Apply Brush’ merges the changes on the preview canvas to the main canvas, ’Clear Brush’ removes the brush box from the canvas area, and ’Clear Image’ removes any changes on the preview layer that haven’t been applied to the main canvas.

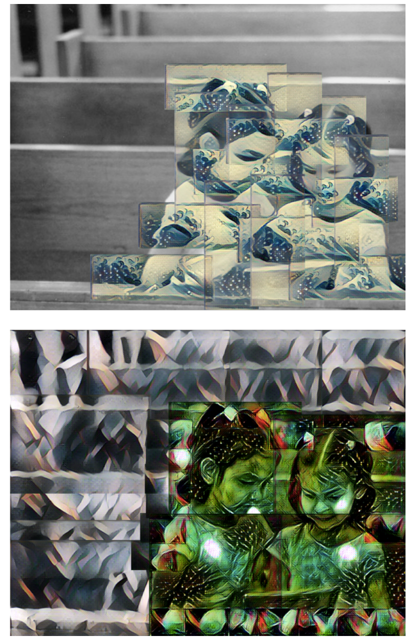

Figure 5: Results of using the Compositing Stamps Application on a custom image. We combine multiple different styles and orderings to create creative new paintings and representations of old content.

By sending rectangular selections of canvas directly to the style transfer algorithm, we can study the trained network on a continuous deformation of content. This interactive process reveals the artifactual texture of the network, for instance in its response to boundaries and aspect ratios besides the literal content. That is, our system allows polling of the fast style transfer network’s expressive range. Figure 5 demonstrates some of the unique patch-like results from using the Compositing Stamps method to alter images.

4 CONCLUSION ON ART CREATION PROCESS

In the development of this application we sought to understand the role of artistic vision in art creation. In the physical world, does intended layout take precedence over colors? Or do artists bring to their canvas a preconceived personal bias towards a certain style, and let the layout appear as a result of their own creativity? Thinking of style transfer inherently separates the creative process for creating an art piece by forcing the user to pick out an image to modify (layout) in its entirety, and identifying an existing style to mimic (colors) [4]. Our tool brings these two processes together by adding the flexibility of choice and application pattern within a single image, breaking down the separation and engaging the user in a new creative process. The traditional roles of artist and curator are intertwined.

5 FURTHER WORK

In experimenting with web frameworks for machine learning in an artistic context, we were successful in making an easy-to-use set of style transfer painting systems for users of various levels of skill and creativity. Further iterations of this work will include options to scale to larger images, submit new paintings for the creation of custom brushes, and the addition of a live mirror mode for video style transfer attached to users' web cams. The greatest challenge in accomplishing these tasks will be optimizing for browser constraints. As with any image manipulation task, style transfer is a computationally heavy process that scales with time as the size of the input increases.

As we progress in this field we also move towards understanding the motivations behind two very different audiences: researchers and content creators. While researchers are heavily invested in the technical implementations of such creative processes, content creators seek out a greater array of features for authoring and versioning their work. This difference guides development as we aim to cater to the needs of both while also reducing the knowledge and philosophical gap between the two roles.

REFERENCES

[1] [n. d.]. ml5js · Friendly Machine Learning For The Web. ([n. d.]). https://ml5js.org/
[2] [n. d.]. Short-coated Gray Cat · Free Stock Photo. ([n. d.]). https://www.pexels.com/photo/cat-whiskers-kitty-tabby-20787/
[3] Mike Bostock. [n. d.]. Data-Driven Documents. ([n. d.]). https://d3js.org/
[4] Alex J. Champandard. 2016. Semantic Style Transfer and Turning Two-Bit Doodles into Fine Artworks. *CoRR* abs/1603.01768 (2016). arXiv:1603.01768 http://arxiv.org/abs/1603.01768
[5] Dongdong Chen, Jing Liao, Lu Yuan, Nenghai Yu, and Gang Hua. 2017. Coherent Online Video Style Transfer. In *The IEEE International Conference on Computer Vision (ICCV)*.
[6] Kate Compton and Michael Mateas. [n. d.]. Casual Creators. http://axon.cs.byu.edu/ICCC2015proceedings/10.2Compton.pdf
[7] Leon A. Gatys, Alexander S. Ecker, and Matthias Bethge. 2015. A Neural Algorithm of Artistic Style. *CoRR* abs/1508.06576 (2015). arXiv:1508.06576 http://arxiv.org/abs/1508.06576
[8] Justin Johnson, Alexandre Alahi, and Fei-Fei Li. 2016. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. (2016). https://cs.stanford.edu/people/jcjohns/eccv16/
[9] Yifang Men, Zhouhui Lian, Yingmin Tang, and Jianguo Xiao. 2018. A Common Framework for Interactive Texture Transfer. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE, Salt Lake City, UT, 6353–6362. https://doi.org/10.1109/CVPR.2018.00665
[10] Helena Sarin. 2018. Playing a game of GANstruction. (Sept. 2018). https://thegradient.pub/playing-a-game-of-ganstruction/