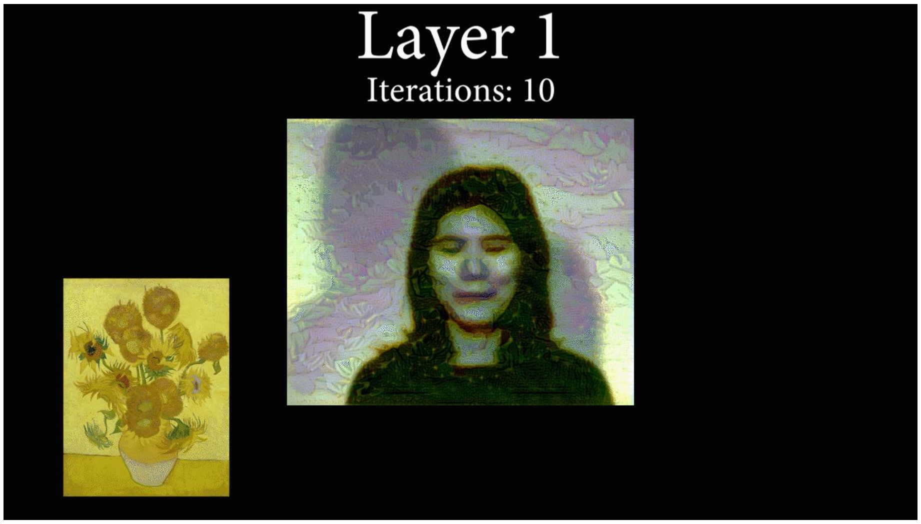
ART AND MACHINE LEARNING
CMU 2019 SPRING
PROJECT 1

Self-Identity in the Stillie & the Selfie



Iris Hwang, Frank Kovacs, Subhodeep Mitra, Marisa Lu, Gillis Bernard

DESCRIPTION

Concept:

"I only wanted to find great people and let them be themselves...and I'd film them for a certain length of time and that would be the movie." – Andy Warhol, 1980

"In the future everybody will be world famous for fifteen minutes." – Andy Warhol, 1965

Andy Warhol's famous screen tests served as inspiration for our project. The video portraits, shot and filmed for three minutes on 16mm film, are meant to be projected at the silent film speed of 18 fps, slowing down the action and extending the running time to four minutes. Screen tests, or "stillies", were the medium through which Warhol gave his subjects their spontaneous piece of fame.

Fame has been democratized in the 21st century. New media like Instagram and YouTube provide outlets for bootstrapped celebrity. Self-representation is carefully filtered and curated (a far cry from the candid screen test) for "likes" from friends and strangers. Content on these channels is even more fleeting than the foretold 15 minutes: The ephemeral Snapchat with its host of filters makes Warhol's four-minute medium look like an epic of uncensored personhood.

We use Warhol's 16mm screen test medium as a starting point to explore our respective self-identities and the perceptions that others have of us. With multi-style transfer, we layer styles that represent (1) who we think we are, and (2) who others perceive us to be, over stills from one second of our respective screen tests. In an effort to take back control of our online identities, we run more iterations on styles that resonated with us. Ultimately, we each combine these stills to create a GIF, a content type that infinitely loops. We also experimented with DeepDream and Guided Deep Dream to explore the idea of what others [in this case, ImageNet] might see in us.

Questions considered:

- *How might we repurpose the screen test as a vehicle for self-expression?*
- *In what ways is a self-initiated screen test similar to and different than a modern-day selfie?*
- *Is it possible to create authentic self-portraits using new media?*
- *How is our self-identity online influenced by others' expectations and biases of gender, ethnicity, and cultural norms?*
- *Is our online identity the sum of our self-perception and the perceptions of others?*
- *What if we could use Google's DeepDream as a tool to deepen our understanding of ourselves and others?*
- *[Assuming the ImageNet dataset is more diverse,] what might DeepDream see in us that we are unable to see in ourselves?*

Technique:

Multi-style transfer

- Inspired by Andy Warhol's screen test, the team decided to use the screen test video obtained from visiting the Andy Warhol Museum. We each selected one second of the video we liked, and broke down the video into a 16 to 18 image sequence in order to replicate the slow motion effect from the screen test video.
- Each team member selected their own styles to transfer onto their base images to represent different aspects of their identity. Therefore, we decided to modify the style transfer algorithm provided in the Jupyter notebooks instead of training new models for the fast style transfer algorithm.
- Based on the style transfer technique, we fed the image sequences to iterations of different style layers to create visuals that are unique to each person. First, the user passes a series of styles they would like to apply to their base images. The program then transfers the first style to the base images, and uses the resulting images as the new base images for the subsequent style transfer iterations.
- We also allow the user to choose how many iterations they want to run on each style to symbolize how important and relevant the style–and aspect of identity–is to them.
- Finally, we combined our resulting styled images to create a video-GIF sequence running on approximately 17 frames/second, which is the same speed at which Warhol ran his screen tests.

Dreaming different layers

- In the regular DeepDream, there is a *Laplacian Pyramid Gradient Normalization* step to boost lower frequencies of the gradient. We modified the algorithm to instead combine multiple layers, but at different frequencies: we selected two layers A and B, and combined low frequencies of layer A with high frequencies of layer B.
- Specifically, we first passed layer A through a Gaussian low-pass filter to filter out the higher frequencies. For layer B, we experimented with some different filters and corresponding thresholds, including the Laplacian of Gaussian. We settled on subtracting the layer B from the Gaussian low-pass filtered version of B (which thus contained only the higher frequencies). Finally, we combined together the two layers, with each layer weighted by some linear combination that we experimented with to find the most visually appealing results.
- We tried this with just the layers themselves (the visualization with image-space gradient ascent), as well as combining the actual DeepDream results of these layers. Especially on the DeepDream results, we see good results, using a dog-like-feature layer as the higher frequencies and a building-like-feature layer as the lower frequencies. When the combined "dream image" is seen from up close, the higher frequencies dominate and we see dog-like features in the image. However from afar/when we squint our eyes, these dog-like-features seem to blend into the background image, and the building-like features become more prominent.

Dreaming vividly

- In order to create even more vivid dreams than the dream images created by regular DeepDream, we decided to implement a "dream sharpening" function. We experimented with different smoothing filters such as Laplacian smoothing, low-pass filtering, and Savitsky-Golay smoothing. All of these methods blur the edges and corners of an object; we realized that subtracting this edges-blurred version of the dream from the original dream creates an image X consisting of only the sharp edges and corners. We then added this image X to the original image, resulting in the edges and corners being much sharper than before. The result of this is that the dream now relies less on the original input image and more on the applied dream layer features.

Process:

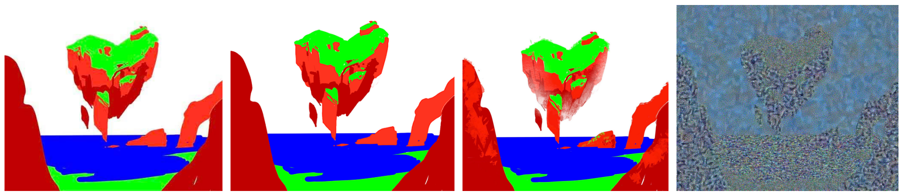
Multistyle Transfer

Trip to the Warhol Museum: We went to the Warhol Museum to learn more about Andy Warhol's screen test process and discover this style. We then each created a three-minute screen test at an interactive exhibit within the museum.

Challenges with fast style transfer: Originally, we thought we would be able to transfer styles on the entirety of the screen test videos we created. We quickly realized that it would take an incredibly long time to complete this process.

Decision to create GIF: In order to execute style transfer in a realistic time frame, we used regular style transfer on stills from our screen test. We selected one second of our screen tests (around 16-18 stills) to use as our source images. This limitation forced us to pivot and inspired us to change the final deliverable to a looping GIF made of 16 stills–essentially a mini screen test. Using a GIF maker and Photoshop, we clipped the screen test and adjusted the grayscale of our stills to prepare them for style transfer. To explore different processes quickly we also had to do image manipulation efficiently on multiple images at a time (16 per group member so 80 total). We used Batch automated script actions in Photoshop to apply grayscale adjustment, pixel clustering for posterization, and later RGB color mapping for target masks.

Exploration of NeuralDoodle: We tried several variations of a Valentine's Day-themed target mask. The first image was too large for the 2x.large GPU spot instance to handle (error thrown was resource overloaded) and so the next was a 400px one; the iteration after that, we tried to give it a much more complex mask with shading between white and red to see how much of the detail would get picked up. As it turns out, there wasn't really any visibly noticeable difference. Despite tweaking parameters, the perlin noise-flavored look to the sky stayed persistent.



Changing model to handle multiple styles and layering: Rather than doing one-off style transfers, we were interested in layering different styles on top of one another to create a composite of style "identities". Using the original style transfer code from the class Jupyter Notebook, we looped through folders of multiple style images and source images. Each team member set their desired number of iterations for each style, and each style is applied to a single image.

Experimenting with iterations: We experimented with the number of style iterations to simulate how important each style (layer) is to each person. Normally, styles with higher iterations will have more influence on the final result. However, this also depends on the style itself. We ran the source image on different styles with the same number of iterations (15). The style with more distinguished features created a more emphasized image.

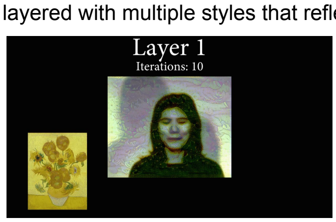
<u>Dreaming Different Layers and Vivid DeepDream</u>

Background research: We read multiple papers (see references section) to establish foundational understanding of how changing frequency of image affects visual perception and how these ideas relate to the algorithms we saw in class.

Filter experimentation: We implemented various filters from the literature based on our thoughts on how these could improve and supplement the algorithms discussed in class like Style Transfer, DeepDream, etc (details in Technique section). We also noted how some background images like the scenery image in the results section works better than faces of people to superimpose dreams on.

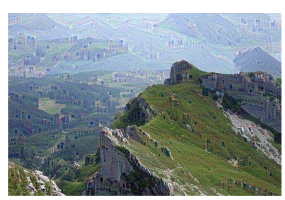
Results:

<u>Multistyle Transfer</u>: We each created a modern, machine-learning take on Warhol's stillie: a one-second GIF of screen test stills layered with multiple styles that reflect aspects of our respective identities and tastes.



Multistyle Transfer

Dreaming Different Layers



Original DeepDream

Vivid DeepDream

<u>Dreaming Different Layers</u>: While combining layers whose frequencies we set with Gaussian filters, note how we see certain layers (dog layer: low-pass filtered) better from up close, but other layers (building layer: high-pass filtered) better from afar.

<u>Vivid DeepDream</u>: Note more emphasis on the dream-features compared to the base image in Vivid DeepDream.

<u>APPENDIX</u>

Reflection & Self-Evaluation:

- We expected the styles with more iterations would have a heavier influence on the final image. However, after examining the result, it seems that the stroke of the style and coloring of the style strongly affect the final result. Styles with more distinguished strokes and features have a more visible effect on the final image.

- Ideally, we wanted to transfer styles onto a full-length screentest and show the incrementing style layers (created manually via the GIF in our supporting materials). Due to processing time, we compromised with a one-second length GIF and used fewer styles and iterations on our screen tests.
- The more styles we added, the harder it became to discern who the person was in the picture. After seven style transfers (7 iterations each), the original person was unrecognizable, thus sparking further questioning on identity erasure.
- Combining different frequencies of different images for DeepDream produced some good results, but not as good as some of the artistic illusions available online created by the same methods. The results were more clearly visible on the base image, than when visualized by space image gradient ascent (which makes sense because the latter method didn't specifically filter out frequencies but rather used the naturally low frequencies produced).
- Vivid DeepDream produced a result exactly like we expected, which is something we were quite excited about - a random idea to make the dream features more vivid ended up doing exactly that!
- We also tried to get a version of Guided DeepDream running, but ran into issues with the setup. Overall, we learnt a lot about different models, especially from running our own experiments and modifying the algorithms to see how results changed. Learning about the idea of playing with frequencies of images (that has a lot of work done in recent years including the SIGGRAPH reference below which was a great read) was fascinating to us!

Roles & Responsibilities:
Our team met frequently to complete this project and took an apprentice approach to learning. Each person contributed specific skills and shared their knowledge with the rest of the team. This method empowered each of us to complete the project deliverables and explore our concept space with confidence.
- Iris Hwang: *Multistyle Transfer* code
- Subhodeep Mitra: *Vivid DeepDream* and *DeepDream frequency modifications* to the algorithm
- Frank Kovacs: GitHub organization, report proofreading, exploration of Neural Doodle
- Marisa Lu: Photoshop batch scripting, image sequence extraction, and breadth exploration
- Gillis Bernard: Report composition, concept ideation, and project management

Detailed Results with GIFs, Photos, and more:
https://docs.google.com/document/d/17Ux0XR1OpNx9kHat2MNAUBQZ39y3aOC_03giy68DuHU/edit#

Code:
https://github.com/frank113/art-ml-project1 [Code in Multistyle_transfer/style_transfer_keras.ipynb and deepdream/notebooks/deepdream.ipynb]

References:
[1] Aude Oliva, Antonio Torralba, Philippe. G. Schyns, 2006. http://cvcl.mit.edu/publications/OlivaTorralb_Hybrid_Siggraph06.pdf
[2] Leon A. Gatys, Alexander S. Ecker, Matthias Bethge https://arxiv.org/abs/1508.06576