

ART AND MACHINE LEARNING
CMU 2019 SPRING
FINAL PROJECT

Spiral Out, Keep Going

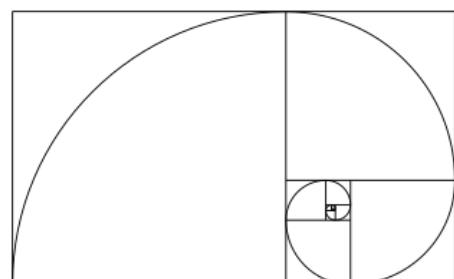


Kevin Tran

CONCEPT

The Fibonacci sequence is a simple one. It begins with 0 and 1, and each subsequent number in the sequence is the sum of the two before it. Thus the sequence is 0, 1, 1, 2, 3, 5, 8, 13... and so on. There are two renowned features of this sequence:

(1) As the sequence continues ad infinitum, the ratio of the latest number in the sequence to the second-to-last-number approaches $\frac{1+\sqrt{5}}{2} \approx 1.618$, the golden ratio; and (2) if you take each number in this sequence, k , and then construct a square of length k for each number in the sequence, it is possible to arrange these squares into a shape of a spiral (right). This is called the Fibonacci spiral.



Peculiarly, the Fibonacci sequence, the golden ratio, and the Fibonacci spiral are ubiquitous in the natural world. Spiral galaxies, nautilus shells, and various plants grow out in Fibonacci spirals. The number of possible ancestors in a human's X chromosome inheritance line follows the Fibonacci sequence as you trace

backwards in each generation, and field daisies often have petals in counts of the sequence. The Fibonacci sequence has also been used frequently in human art. The ratio between the length of the faces of the Egyptian pyramids and the width of the bases follows the Golden ratio. The canvas of Salvador Dalí's "The Sacrament of the Last Supper" has dimensions that follow the Golden ratio. Even cigarette manufacturers created rectangular cartons in proportion of the Golden ratio to improve the aesthetics. The Fibonacci is also a fundamental undertone of a song, *Lateralus*, that has great meaning in my life. I omit a detailed discussion of *Lateralus* for brevity's sake, as I can easily write an entire report on the subject.

There are other features of the Fibonacci sequence I noticed myself, too. The sequence inherently uses previous numbers to create new numbers—it uses its past to build its future. I find it serendipitous that the basis for my research lies in active optimization, where I use machines to learn from the past in order to guide the future. Separately: I noticed that even if you do not start the sequence with 0 and 1, the Golden ratio still holds. For example: If you start the sequence with 5 and 10 and then continue to 15, 25, 40, 65, etc., the ratio of the last number to the second-to-last still approaches 1.618. This means that no matter where you start, you still end at the same place. For all these reasons and more, I felt compelled to create a work of art based around the Fibonacci spiral.

This is what led me to create an algorithm—SpiralOut—that takes two inputs: a subject image and a target attribute. It then chooses new images whose styles are transferred¹ onto the subject image, and it does this iteratively. SpiralOut attempts to choose style images such that the newly styled subject image tends to characterize the target attribute better and better.

The style images symbolize new experiences for the subject, and style transfer symbolizes how those experiences change the subject. The iterative nature of SpiralOut symbolizes how the subject is constantly learning new things and how each new version of the subject is a culmination of their past experiences. SpiralOut's ability to target specific attributes symbolizes the subject's ability and desire to become who and what they want to be.

METHOD

1. Style Transfer: I used the default settings in the Tensorflow-style-transfer example from class notes.
2. Attribute classification: I trained my own model to classify an image's propensity towards certain attributes or emotions. I used the ResNet50 neural network architecture and training methods² as the base model and trained it on the WikiArt-Emotions database.³ Figure 1 illustrates this and lists the attributes.

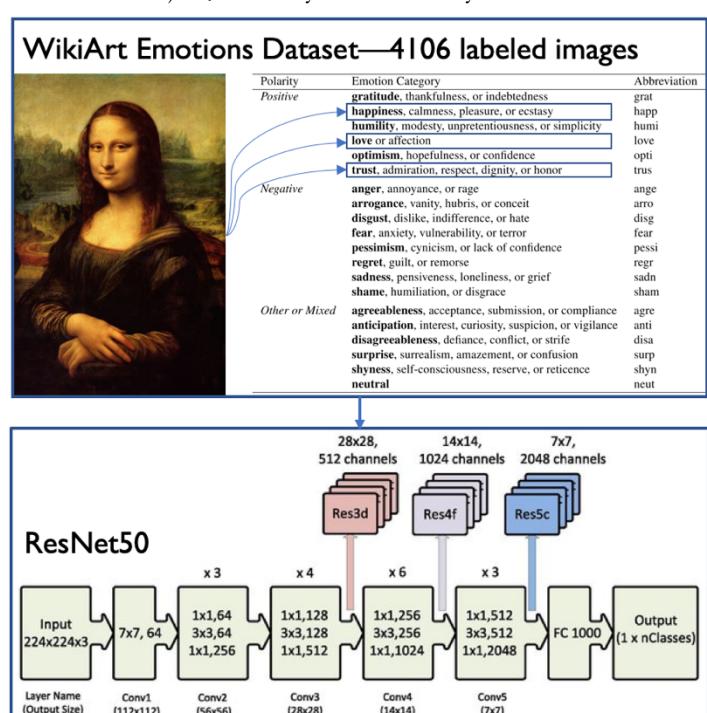


Figure 1: Representation of how I trained ResNet50 on the WikiArt Emotions dataset to create an image classifier for 20 different attributes.

3. Attribute optimization (Figure 2)

- a. I used Bayesian optimization (as implemented in Dragonfly⁴) as the style-image-selection algorithm.
 - i. I used the standard settings in Dragonfly: a pre-sampling size of 7, a batch size of 17, and a dynamic ensemble of acquisition functions.
 - ii. I varied the number of iterations—i.e., the number of style transfers—according to whichever experiment I was performing.
- b. Bayesian optimization entails the maximization of some objective function.
 - i. I defined this objective function as ResNet50’s logistic regression of the subject image’s classification as the target attribute.
 - ii. For example: Let us feed the current version of the subject image into ResNet50. Then ResNet returns a value of 0.70 for the image’s match to the “optimistic” attribute. The objective function value for the subject at this iteration is then 0.70.
 - iii. Note that the objective function depends on the target attribute.
- c. Bayesian optimization also requires us to define input variables, which it then varies to optimize the objective function.
 - i. Here I define the input variables as the output of the FC1000 layer of ResNet50, which is the final layer prior to logistic classification.
 - ii. Bayesian optimization does not perform well in high dimensional space, so I reduced the dimensionality of the FC1000 layer to 16 dimensions by using a 1-layer [auto]encoder with fully connected soft-plus nodes, the ADAM optimizer, and a mean-squared-error loss function.
 - iii. I used the Web Gallery of Art’s database of images⁵ as my sampling space for the Bayesian optimization/candidate style images.
 - iv. Dragonfly is able to account for discrete sampling spaces such as this, but I did not know that when I created SpiralOut. So instead of allowing Dragonfly to sample the images, I allowed Dragonfly to return a continuous vector of 16 input variables, and then I used a KDTree to identify whichever image within the Web Gallery of Art matched this vector most closely.

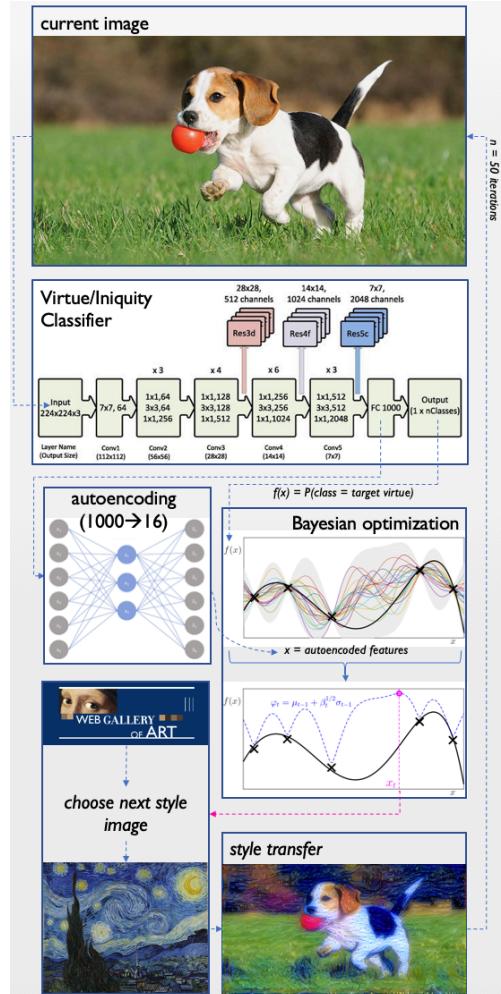


Figure 2: Outline of the “SpiralOut” algorithm.

4. Visualization
 - a. Each style transfer yielded 20 frames, and a “standard” SpiralOut experiment consistent of 50 iterations of style transfer. I concatenated this images into GIFs with 20 frames per second, which yields 50 second GIFs.
 - b. I also created “experience” GIFs that displayed the style images being chosen throughout the SpiralOut process. These experience-GIFs are timed such that they can be played alongside the subject GIFs.

TRIALS AND ERRORS

1. SentiBank
 - a. Jou & Cheng⁶ show that they were able to successfully fit ResNet50 on the Visual Sentiment Ontology (VSO) dataset and that this new model surpassed their previous classifier, DeepSentiBank.⁷
 - b. I found both the source code and compiled versions of DeepSentiBank⁷ online and tried to use them as my classifier. But after a week of failing to get their source code to run properly and multiple (ignored) email requests to the authors, I gave up.
 - c. After I realized that their “best” model was simply ResNet50 trained on their own dataset, I decided to use a Keras template for ResNet50 and then train on their publicly available dataset.
2. Visual Sentiment Ontology (VSO) dataset
 - a. The same group that created SentiBank also created the VSO dataset. This dataset contains ~40,000 different images they scraped from Flickr that were tagged with various adjective-noun-pairs (ANPs). I took these images, removed the nouns, and left the adjectives as labels. Thus I was left with ~40,000 images with ~260 labeled classes, where classes were attributes such as “happy”, “abandoned”, “lonely”, “dusty”, “loving”, or “broken”.
 - b. I successfully trained ResNet50 on the VSO dataset and used the subsequent model in the SpiralOut algorithm. After some experimentation though, I was not satisfied with the algorithm’s ability to imbue the subjects with the target qualities (Figure 3).
 - c. I hypothesized that the poor performance was due to an over-saturation in possible classifications, so I then search for a dataset that was similar to VSO, but with less classes. This is what led me to the WikiArt-Emotions dataset.³



Figure 3: Example results from the VSO-version of SpiralOut. Maynard James Keenan (left) was imbued with “trouble” (bottom-left), and Jon Snow (right) was imbued with “beauty” (bottom-right).

3. Fibonacci spirals

- a. I originally planned to create a video where each iteration was in the location of one of the Fibonacci squares within a Fibonacci spiral, and then the video zoomed out of the spiral to reveal each new iteration.
- b. The aesthetics of similar images within a Fibonacci spiral were less appealing than I anticipated (Figure 4), which is why I decided to go with videos instead of a spiraling motif.

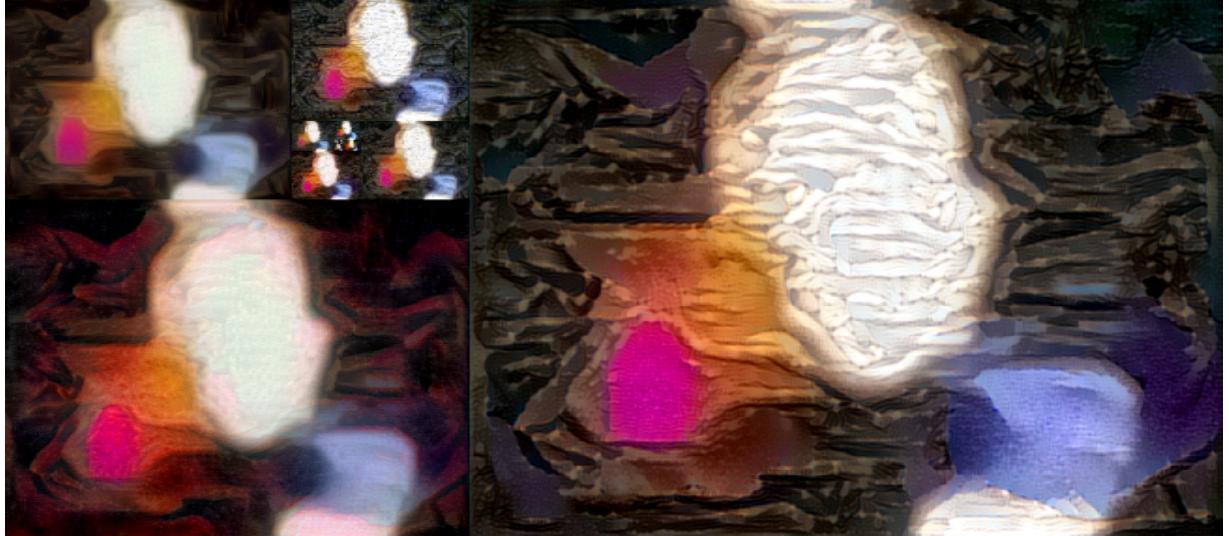


Figure 4: Original “spiraling” idea where each new image would be shown on the outer-most part of the Fibonacci spiral.

RESULTS

I experimented with various configurations of subjects and attributes. One “standard” experiment shows what happens when we try to make a painting of Benedict Cumberbatch become more optimistic (Figure 5). Here we see a color binarization similar to what is seen in Figures 3 and 4. Darker colors trend towards black, lighter colors trend towards white, and then a patch of red and a patch of blue appear. In my opinion, the final image is not necessarily “optimistic”. This suggests that transitioning from the VSO training set to the WikiArt-Emotions dataset did not help.

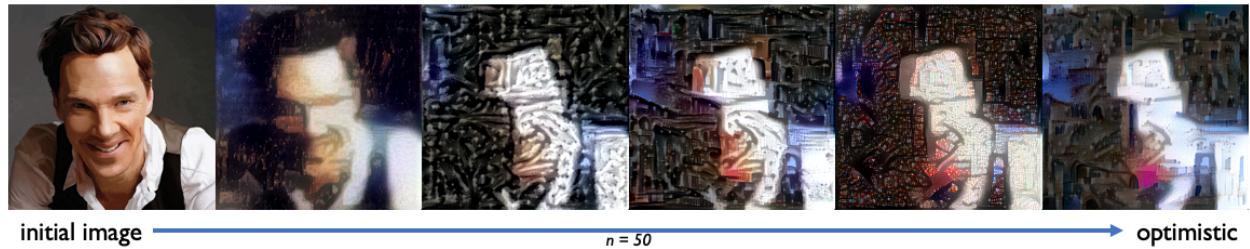


Figure 5: A painting of Benedict Cumberbatch with the SpiralOut algorithm applied using the “optimistic” attribute target.

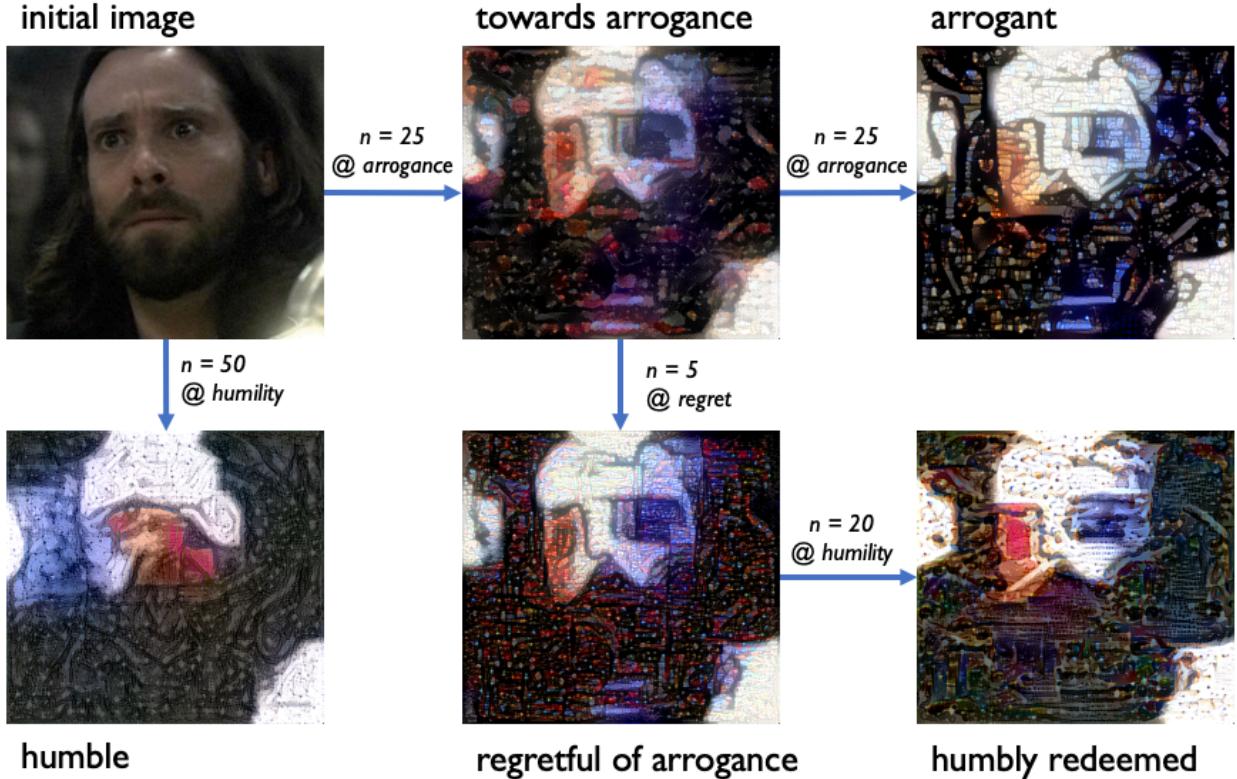


Figure 6: A thought experiment where I used the SpiralOut algorithm with various targets and numbers of iterations to yield different versions of Gaius Baltar.

Figure 6 shows another thought experiment I performed with Gaius Baltar as a test subject. First I created an “arrogant” version of him (top-right of Figure 6), then I created a humble version of him (bottom-left of Figure 6). I had hoped to see a drastic difference between these two diametrically opposed attributes, but the images did not appear to be different. I then tried to take the “arrogant” version of Gaius and then “redeem” him by imparting regret and then humility on him. I had hoped that the “redeemed” Gaius would look more similar to the humble Gaius, but he appeared to look similar to the arrogant version. I know that this phenomena is an artifact of the SpiralOut algorithm, but I was still disappointed by the artistic symbolism of this: Gaius could not be redeemed.

Note that my “final submission” is the video I used for the art show, which contains various examples of the SpiralOut algorithm.

REFLECTION

I am proud of being able to conceptualize and construct this entire framework on my own, but I am still disappointed in the algorithm’s inability to imbue the subject images with the target attributes. If I decide to spend more time on this project, I would like to try using an Emotional GAN⁸ to create style images in some way. The aesthetics of the images created by Alvarez-Melis & Amores⁸ (Figure 7) look closer to the images I was intending to create. What I hope to create is a final image whose styles look like the ones in Figure 7, but whose contents are “echoes” of the original subject image. I also hope to perform experiments such as those in Figure 6, but actually find that we are redeemable.

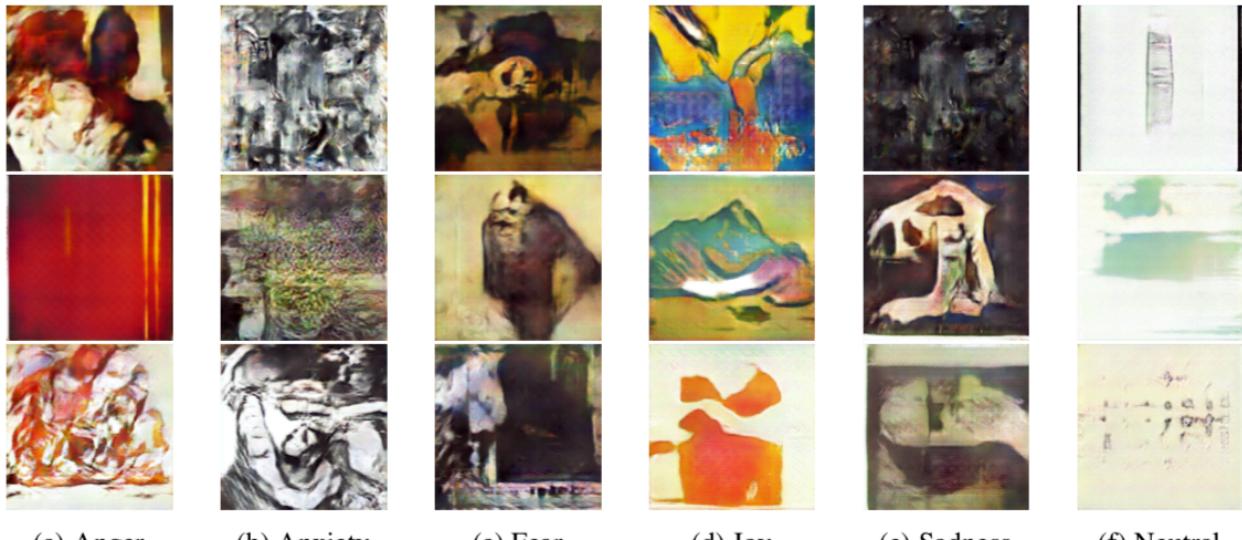


Figure 7: Images created by the Emotional GAN, created by Alvarez-Melis & Amores.

CODE: https://github.com/ktran9891/spiral_out

REFERENCES

1. Gatys, L. A., Ecker, A. S. & Bethge, M. A Neural Algorithm of Artistic Style. arXiv: 1508.06576 (2015).
2. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. arXiv: 1512.03385 (2015).
3. Mohammad, S. M. & Kiritchenko, S. WikiArt Emotions: An Annotated Dataset of Emotions Evoked by Art. Proceedings of the 11th Edition of the Language Resources and Evaluation Conference (LREC-2018), 1225–1238 (2018).
4. Kandasamy, K. et al. Tuning Hyperparameters without Grad Students: Scalable and Robust Bayesian Optimisation with Dragonfly. arXiv: 1903.06694 (2019).
5. Web Gallery of Art, accessed April 2019. <https://www.wga.hu/>
6. Jou, B. & Chang, S.F. Deep Cross Residual Learning for Multitask Visual Recognition. arXiv: 1604.01335 (2016).
7. Chen, T. et al. DeepSentiBank: Visual Sentiment Concept Classification with Deep Convolutional Neural Networks. arXiv: 1410.8586 (2014).
8. Alvarez-Melis CSAIL, D. & Amores, J. The Emotional GAN: Priming Adversarial Generation of Art with Emotion. Nips (2017).