Generating artwork according to an emotional profile

Konrad Cybulski

March 2019

1 Introduction

In the domain of creative artificial intelligence and evolutionary art, the desire exists for processes that produce imagery that is not only visually appealing, but images that exhibit abstract and emotive characteristics. There exists extensive research on the production of realistic, and target label accurate images such as work by Nguyen et al. (2015b) using quality-diverse (QD) algorithms in combination with deep neural networks (DNN), or Bao et al. (2017) that uses a variational generative adversarial network (GAN) architecture. Recent work by Tan et al. (2017) has explored techniques for synthesizing artwork according to a target artist or genre through the use of a GAN architecture with highly accurate and creative results.

Little exploration however has been done on incorporating emotion into the process of art and image generation. Ali and Ali (2017) explored the idea of *emotion transfer*, using techniques such as image emotion assignment, and color/style transfer with the aim of altering image composition to reach a target emotion. Examples given use a target profile, with varying levels of emotions such as joy, anger, and fear, to alter the image's color composition. The classification of an image's affective emotion, the emotion with which a viewer classifies the image, has been explored in various works (Machajdik and Hanbury, 2010; Chen et al., 2015; Kim et al., 2018). Kim et al. (2018) produced a classifier for recognizing the emotion attributed to an image. This was done through the application of a DNN to decompose an image to a two-dimensional feature vector (valence and arousal) representing the image's emotion mapped to a continuous plane (see Figure 2).

Recent work by Nguyen et al. (2015a,b) in evolutionary image generation has aimed to address the desire for a system that not only produces visually appealing images, but a diverse collection. To address the need for a high quality, yet diverse solution space in related optimization domains, algorithms such as Multi-dimensional Archive of Phenotypic Elites (MAP-Elites) (Mouret and Clune, 2015) and Novelty Search (Lehman and Stanley, 2008, 2011) have been developed. The use of such QD algorithms has shown great promise in its efficiency and accuracy on a number of hard optimization problems (Pugh et al.,

2016) such as maze navigation (Lehman and Stanley, 2011). The combination of MAP-Elites and DNN classifiers used by Nguyen et al. (2015a) and Nguyen et al. (2015b) in conjunction with a deep neural network (DNN) image classifier; assigning individual fitness according to the accuracy with which a generated image is classified. Nguyen et al. (2015b) applied this architecture to the domain of image generation and has demonstrated the visual diversity that arises from the use of a fitness evaluation that "can provide informative, abstract distance functions in high-dimensional spaces" (Nguyen et al., 2015b, p. 8).

==

The work conducted by Tan et al. (2017) aimed to create a process by which artwork could be generated with more abstract characteristics. Through the use of a target artist or genre, and a respective discriminator assigned label, the generative network was able to produce images that were stylistically similar to the desired artist or genre.

There has been extensive research however into the synthesis of visually appealing and aesthetic images using evolutionary algorithms. With human-guided evolution (Machado and Cardoso, 2000) the process of fitness evaluation relies entirely on both subjective appeal, and unquantifiable metrics. Quantifiably estimating the aesthetic appeal of an image has been studied by den Heijer and Eiben (2010b,a, 2011) with varying results. A number of metrics such as image compression complexity, distribution of color gradients, fractal dimension, and contrast have shown to improve the quality of images generated through an unsupervised evolutionary process (den Heijer and Eiben, 2014).

==

Discuss the paper Johnson et al. (2019) which looks at the dissonance between aesthetic measures used in evolutionary art and the distributions of features ratings given by humans.

==

2 Aims

The aim of this research is to explore the synthesis of artwork with a target emotional profile. Primarily leveraging work by Tan et al. (2017) and Kim et al. (2018) to produce a generative architecture whose output not only has desirable abstract characteristics, but shows emotive capabilities. The proposed system would generate an image that satisfies a set of emotions provided. This will investigate both the efficacy with which a generative system can create emotive images, as well as give insight into the properties attributed to various emotions portrayed in image form.

This research will explore both the use of a GAN architecture for producing emotionally driven artwork, as well as a quality-diverse approach based on the findings of Nguyen et al. (2015b). Due to the feature-space exploration performed by the MAP-Elites algorithm, it may have a comparative advantage over the GAN model used by Tan et al. (2017). As noted by Nguyen et al. (2015b), the feature-space maintained by MAP-Elites readily allows further exploration of the distribution of images generated as a function of their emotional feature profile.

In order to test, and verify the output of such a system, generated images will be exhibited to explore their emotional effect on humans, and any dissonance between the intended, and resulting emotional profile. This will further verify the accuracy with which an emotional profile can be synthesized into affective artwork with such an architecture.

3 Background

3.1 Unsupervised image synthesis

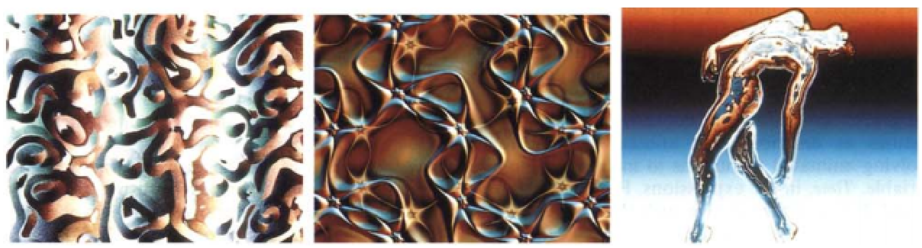

Figure 1: Images generated through the process of interactive evolution introduced by Sims (1993)

The area of evolutionary art and image generation has been explored for many years. Some of the first *human-in-the-loop* systems such as *NEvAr* (Machado and Cardoso, 2000) produced greatly impressive images leveraging methods introduced and exemplified by Sims (1993) such as those shown in Figure 1. Sims (1993) proposed using *Lisp* expressions for genotype definitions, which accepted a coordinate (x, y) which could be evaluated into a grayscale or RGB value, thus producing images. This genotype expression has been used in numerous further research into the process of both supervised and unsupervised image synthesis (Machado and Cardoso, 2000; Sims, 1993; den Heijer and Eiben, 2011, 2013; Ross et al., 2006).

Sims (1993) and Machado and Cardoso (2000) were able to produce images with visually striking characteristics, despite the slow nature of the interactive process. Ross et al. (2006) investigated the use of measures of aesthetics for fitness evaluation in artificially evolving images. This research primarily used the observation of Ralph (2006), that the distribution of colour gradients in fine art tend towards normal. While the images produced through this method did not meet the level of intricacy and detail as the results of Sims (1993) or Machado and Cardoso (2000), it represented a self-contained system able to generate appealing art without human interaction.

The introduction of the generative adversarial network architecture (GAN) by Goodfellow et al. (2014) allowed the process of image generation to be completely unsupervised. Common GAN application has involved the generation of realistic images, such as has been done by Bao et al. (2017), where images have been synthesized to fine-detailed target labels such as bird species' and actors. Zhang et al. (2017) and Reed et al. (2016) have recently explored text to image synthesis, in which detailed descriptions of birds and flowers have been converted into photo-realistic images using the GAN model. Tan et al. (2017) has explored the generation of art according to target genre and artist. Where throughout the training process,

3.2 Quality-diversity algorithms for generative and its effect on solution quality

In the area of evolutionary art, numerous methods for maintaining genotype and phenotype diversity have been explored den Heijer and Eiben (2013, 2012), and recent research using state-of-the-art quality-diversity (QD) algorithms such as MAP-Elites have shown great promise (Nguyen et al., 2015b).

As mentioned by Nguyen et al. (2015b) in the context of an *Innovation Engine*, high fitness individuals for a given target classification label (e.g. water tower, volcano, dome, etc.) often arise from varying fitness individuals in a different domain. The example shown by Nguyen et al. (2015b) shows the path taken by numerous images descending from the classification of *abaya* with 47% confidence. A *volcano* (99%) image descended from a *castle* image (4%), a *planetarium* (95%) from a *boathouse* (10%), a *beacon* (96%) from a *cocker spaniel* (2%). The evolutionary trajectory of individuals does not resemble the path taken in common genetic algorithms, where high fitness individuals arise from the exploitation of a local optima. The evolutionary path in this context conflicts with the findings of Mouret and Clune (2015) in which the MAP-Elites algorithm applied to neural network optimization resulted in the fittest individuals arising due to mutations of their direct parents in the feature-space.

3.3 Deep neural networks for image classification and fitness evaluation

Deep neural networks (DNN) have grown tremendously in popularity in the domain of image generation, and classification, and the accuracy with which they perform.

Research by Burton and Vladimirova (1998) showed that a genetic algorithm composing music benefited from fitness evaluation that relied on phenotype clustering, favoring those that showed diversity from existing clusters.

Recent work by Nguyen et al. (2015a) and Nguyen et al. (2015b) has used pre-trained DNN image classifiers for fitness evaluation throughout the evolutionary process.

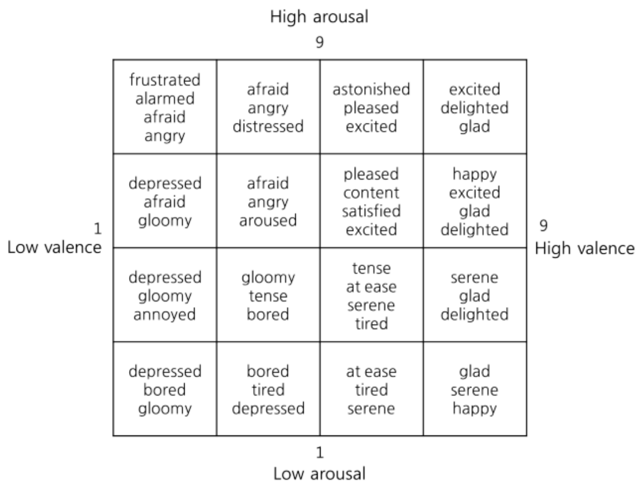

Figure 2: Distribution of emotions associated with levels of valence and arousal determined by DNN classifier produced by Kim et al. (2018)

3.4 Image emotion recognition

4 Methodology

5 Expected Outcomes & Contributions

References

Ali, A. R. and Ali, M. (2017). Emotional filters: Automatic image transformation for inducing affect. *arXiv preprint arXiv:1707.08148*.

Bao, J., Chen, D., Wen, F., Li, H., and Hua, G. (2017). Cvae-gan: fine-grained image generation through asymmetric training. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2745–2754.

Burton, A. R. and Vladimirova, T. (1998). Genetic algorithm utilising neural network fitness evaluation for musical composition. In *Artificial Neural Nets and Genetic Algorithms*, pages 219–223. Springer.

Chen, M., Zhang, L., and Allebach, J. P. (2015). Learning deep features for image emotion classification. In *2015 IEEE International Conference on Image Processing (ICIP)*, pages 4491–4495. IEEE.

den Heijer, E. and Eiben, A. (2010a). Using aesthetic measures to evolve art. In *IEEE Congress on Evolutionary Computation*, pages 1–8. IEEE.

den Heijer, E. and Eiben, A. (2011). Evolving art using multiple aesthetic measures. In *European Conference on the Applications of Evolutionary Computation*, pages 234–243. Springer.

den Heijer, E. and Eiben, A. (2012). Maintaining population diversity in evolutionary art. In *International Conference on Evolutionary and Biologically Inspired Music and Art*, pages 60–71. Springer.

den Heijer, E. and Eiben, A. (2013). Maintaining population diversity in evolutionary art using structured populations. In *2013 IEEE Congress on Evolutionary Computation*, pages 529–536. IEEE.

den Heijer, E. and Eiben, A. (2014). Investigating aesthetic measures for unsupervised evolutionary art. *Swarm and Evolutionary Computation*, 16:52–68.

den Heijer, E. and Eiben, A. E. (2010b). Comparing aesthetic measures for evolutionary art. In *European Conference on the Applications of Evolutionary Computation*, pages 311–320. Springer.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2014). Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680.

Johnson, C. G., McCormack, J., Santos, I., and Romero, J. (2019). Understanding aesthetics and fitness measures in evolutionary art systems. *Complexity*, 2019.

Kim, H.-R., Kim, Y.-S., Kim, S. J., and Lee, I.-K. (2018). Building emotional machines: Recognizing image emotions through deep neural networks. *IEEE Transactions on Multimedia*, 20(11):2980–2992.

Lehman, J. and Stanley, K. O. (2008). Exploiting open-endedness to solve problems through the search for novelty. In *ALIFE*, pages 329–336.

Lehman, J. and Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. *Evolutionary computation*, 19(2):189–223.

Machado, P. and Cardoso, A. (2000). Nevar–the assessment of an evolutionary art tool. In *Proceedings of the AISB00 Symposium on Creative & Cultural Aspects and Applications of AI & Cognitive Science, Birmingham, UK*, volume 456.

Machajdik, J. and Hanbury, A. (2010). Affective image classification using features inspired by psychology and art theory. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 83–92. ACM.

Mouret, J.-B. and Clune, J. (2015). Illuminating search spaces by mapping elites. *arXiv preprint arXiv:1504.04909*.

Nguyen, A., Yosinski, J., and Clune, J. (2015a). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 427–436.

Nguyen, A. M., Yosinski, J., and Clune, J. (2015b). Innovation engines: Automated creativity and improved stochastic optimization via deep learning. In *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, pages 959–966. ACM.

Pugh, J. K., Soros, L. B., and Stanley, K. O. (2016). Quality diversity: A new frontier for evolutionary computation. *Frontiers in Robotics and AI*, 3:40.

Ralph, W. (2006). Painting the bell curve: The occurrence of the normal distribution in fine art.

Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., and Lee, H. (2016). Generative adversarial text to image synthesis. *arXiv preprint arXiv:1605.05396*.

Ross, B. J., Ralph, W., and Zong, H. (2006). Evolutionary image synthesis using a model of aesthetics. In *2006 IEEE International Conference on Evolutionary Computation*, pages 1087–1094. IEEE.

Sims, K. (1993). Interactive evolution of equations for procedural models. *The Visual Computer*, 9(8):466–476.

Tan, W. R., Chan, C. S., Aguirre, H. E., and Tanaka, K. (2017). Artgan: Artwork synthesis with conditional categorical gans. In *2017 IEEE International Conference on Image Processing (ICIP)*, pages 3760–3764. IEEE.

Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., and Metaxas, D. N. (2017). Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 5907–5915.