CREA.blender: a GAN based casual creator for creativity assessment

Miroslav Gajdacz[1], Janet Rafner[1], Steven Langsford[1], Arthur Hjorth[1], Carsten Bergenholtz[1], Michael Mose Biskjaer[2], Lior Noy[3], Sebastian Risi[4], and Jacob Sherson[*1]

[1]Center for Hybrid Intelligence[†], Department of Management, BSS, Aarhus University, Denmark
[2]Center for Digital Creativity, dpt. Digital Design and Information Studies, Aarhus University, Denmark
[3]Business Administration, Ono Academic College, Kiryat Ono, Israel
[4]IT University of Copenhagen, Copenhagen, Denmark

Abstract

In this technical demonstration paper we document the use of a Generative Adversarial Network (GAN) based casual creator game for systematic assessment of human creativity. We discuss some of the challenges in designing GAN-based casual creators, specifically focusing on how to identify and select appropriate parts of the latent vector space - images, in our case.

Introduction

Since their invention, Generative Adversarial Networks (GANs) (Goodfellow et al. 2014) have deservedly attracted attention in the field of computational creativity (Berns and Colton 2020). GANs excel at producing realistic artifacts (Karras, Laine, and Aila 2019) and have the potential to replace humans at the most laborious parts of the manual creative process, to the extent of facilitating visual creation for people without training in the manual arts.

GANs are well-known for producing artificial images that can be nearly indistinguishable from real images (Borji 2019). More recently, artists have begun using GANs for creative purposes to create images and music (Berns and Colton 2020). However powerful these tools may be, they come with substantial requirements for the technical expertise of their users to be able to implement and execute the underlying Machine Learning models and devise the processes, which create the desired artifacts.

In a recent trend called 'casual creators', digital tools are designed to "empower autotelic and enjoyable amateur creativity" (Petrovskaya, Deterding, and Colton 2020). Casual creators afford the creation of highly elaborate artifacts with little input from users. At the core of these products often lie algorithmic generators, of which GANs are just one example. These generators map from a simple low-dimensional input domain, such as sliders, dragging gestures, numerical or multiple choice parameters, which are relatively easy to understand for the users, to a complex high-dimensional output domain, such as images or audio. In a well-designed casual creator, the mappings are intuitive enough to allow for meaningful directed search, creativity, and serendipity to emerge from people's interactions with these systems.

Thus far, casual creators have mostly been used for entertainment (Petrovskaya, Deterding, and Colton 2020). In this paper we propose a novel use of casual creators: to systematically assess human creativity. Standard psychometric tests for assessing creativity are based on simple tasks such as making as many unique drawings as possible out of circles (Torrance 1966) or coming up with alternative uses for an object (Guliford 1968). This design is intentional so that the task is widely accessible. However, in most circumstances it does not allow for complex, interesting expressions of ideas. Utilizing a casual creator has the potential to change this.

Here, we present *crea.blender*, a GAN-based image generation game, explicitly designed to assess creativity of the general public (Rafner et al. 2020). This game is part of a broader suite of games and tasks to measure creativity, called *CREA* (Rafner 2021). *crea.blender* utilizes a Big-GAN model (Brock, Donahue, and Simonyan 2019) in three distinct sub-tasks where a set of base images, either predefined or selected by the user, can be "blended" together into one output image.

The design of *crea.blender* draws heavily on existing casual creation systems, particularly Artbreeder (originally *Ganbreeder*) (Simon 2021). The *crea.blender* system is distinct in two main respects. Firstly, the *crea.blender* interface is somewhat simpler and does not expose individual 'genes' to player manipulation as Artbreeder does. More importantly, *crea.blender* is divided into three play modes, the challenge mode, divergent mode, and open play mode, in order to expose different elements of the creative process to systematic measurement.

Details of the three play modes and their theoretical foundations are given below, but in brief the challenge mode asks players to reproduce a target image, the divergent mode asks them to create as many different images of a particular theme, and the open play is unconstrained. A core goal of *crea.blender* is to explore the extent of the relationships between the different modes. Some elements of the task, like the perceptual similarity between images, are unlikely to change across modes. Others, like the search strategies employed, could potentially change drastically in response to the incrementally looser constraints. By applying different task demands to the same underlying space, *crea.blender* offers the opportunity to look for distinctive features of these strategies, and the extent to which these features recur across

---

[*]sherson@mgmt.au.dk

[†]www.hybridintelligence.eu/crea/

1

the different tasks types. We consider the extent to which these different tasks can be understood as reflecting overlapping cognitive abilities to be an important open question.

The purpose of this demo article is to outline a novel use of casual creators as a tool for systematically assessing creativity, and to discuss some of the fundamental design challenges in creating interesting creativity tasks with a GAN. We do this by providing a technical description of the GAN model and how it is utilized for cross-category image blending. We then outline and discuss challenges relating to finding suitable images which provide sufficient expressiveness in the creative process for the non-specialist users. Finally, we discuss more general implications for designing casual creators with GANs and generative ML.

The Game

CREA is a game written in Unity and runs via WebGL in a browser, making it accessible on a wide range of desktop and mobile devices. The execution of the underlying GAN model involves a large amount of parallel calculations, which can be significantly sped up (by factor 10-100) when performed on a graphics card (GPU) compared to CPU. To broaden participation and make crea.blender accessible for people without GPUs, the image generation is performed remotely on a server with a GPU. Our GAN server is running on an Azure Virtual Machine and can process around ten image-generation requests per second.

Challenges in designing with a GAN

The GAN model (Brock, Donahue, and Simonyan 2019) has been trained on ImageNet (Deng et al. 2009) and uses as input a 1,000-element class vector and 128-element noise vector, which together form a so-called latent vector. Each class represents a real-life object, i.e. a beetle, a dog, etc., and each specific latent vector, when propagated through the GAN, deterministically produces an image.

The original purpose of the GAN is to generate realistic examples of a single chosen class, which is done by inserting randomly generated instances of the noise vector, while keeping only one component of the class vector non-zero (the selected class).

Image mixing procedure

In *crea.blender* we allow for multiple class vector components to be non-zero, therefore the output images rarely resemble members of any particular class. Rather, each of the images that people use are combinations of many different real-life objects. The base images that are entering the mixing were also produced by the GAN and are each specified by their latent vectors.

The players control the mixing procedure by sliders located below each base image. Once they have chosen some desired slider settings, they push the image generation button, which sends an image generation request to the server. The latent vectors of the base images are linearly superimposed (added together) with weights proportional to the slider values. Since we do not perform normalisation of the combined latent vector, adding the images with small weights can lead to unusual outputs not resembling the base images.

Design challenges

The fundamental challenge in creating a causal creator with a GAN is selecting the images that users can blend together. The images should be interesting and have aesthetic qualities in their own right. But more importantly for a casual creator like ours, the underlying latent vector of these images should 'blend well' with the other images, meaning that if you take two images and blend them, the resulting image should meaningfully look like a combination of the two. Further, the different creative modes have slightly different requirements for their images, discussed below.

Modes of creativity assessment

crea.blender has three creative modes, each one utilizing the GAN in a different manner. We aim to measure and study player performance across tasks that vary in the specificity of their goals. This is to assess the players ability in:

- Expressing themselves, that is reaching a specific target: *Challenge mode*

- Producing many alternative solutions to a specified theme: *Divergent mode*

- Producing novelty and value in general: *Open-play mode*

Challenge mode

The design of the challenge mode is built to assess the creative process convergent thinking, which is defined as the ability to find the single best solution to a defined question (Guilford 1956). Each trial in the Challenge mode has two stages. In the first, participants are presented with three sets of three images (see Figure 1.a), on the left) and have thirty seconds to indicate which of the three sets can produce a target image (see Figure 1.a), on the right). After they have selected the correct set of images (possibly on the second or third attempt) participants progress to the blending stage, and attempt to reproduce the target image by setting contribution sliders appropriately on the three base images (see Figure 1.b). The participant-generated image is updated whenever the generate button in the center of the screen is clicked. The trial ends when a generated image is sufficiently close[1] to the target or when two minutes have elapsed in the blending stage. Feedback is given with text prompts at the end of each stage.

Divergent mode

The divergent mode is built to assess 'divergent thinking' a creative process which can be defined as the ability to come up with many different solutions to a prompt (Guilford 1956). Divergent thinking is often further broken down into the components of ideational fluency (the number of outputs: ideas, products, solutions), flexibility (how different the proposed outputs are from each other), originality

---

[1]Acceptance thresholds were manually specified for each image

Figure 2: *Screenshot of the game in the Divergent mode. The Open-play mode looks identical, but allows for switching the individual base images by clicking on them.*
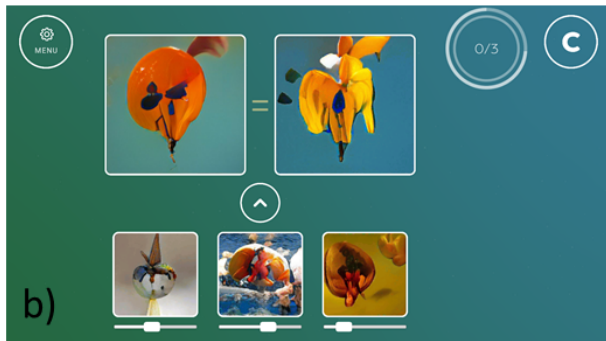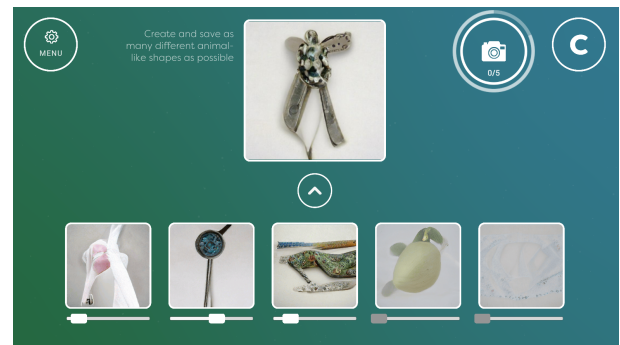
Figure 1: *Challenge mode: a task with a well defined goal. a) Choose one image set which can produce the given target image on the right. b) Chose the image mixing weights for the base images (sliders below) to produce a blended image (on the top left) such that it is close to the target image (on the top right).*

sessed through based on the commonly accepted definition of creativity: novelty and value (Runco and Jaeger 2012). Currently the images are assessed through crowdsourcing of other participants in an evaluation phase, but we are working on supplementing this with algorithmic techniques such as clustering.

Figure 3: *Examples of images that can be generated in the Open-play mode.*

(how unique an output is), and elaboration (the level of details in an output). The divergent mode asks players to create as many different animal-like images as possible in four minutes. A 'continue' option to end the task early is also available after five images have been submitted. Participants create images by setting sliders on a fixed set of five base images (see Figure 2), clicking a generate button to render the image resulting from the current slider choices in a display area at the center top of the screen.

Blended images can be submitted by clicking on a camera button in the top right. Previously submitted images can be reviewed via a gallery accessed through a menu button in the top right. An instruction reminder prompt to create as many animal-like figures as possible is visible in top left of the display throughout. No feedback is presented in this task.

Open-play mode

The open-play mode asks participants to "create as many creative images as you can" in four minutes. An option to end the task early is available after five images have been submitted. Visual presentation of this task is the same as in the Divergent mode, but the set of base images is not fixed. Clicking on a base image replaces it with another base image drawn randomly from a pool of 33 items. This new mechanic is introduced to participants with a short tutorial at the beginning of the task. The open play mode is as-

Pre-selection of the base images

The image mixing process on a pre-trained GAN is relatively simple. The difficult part is to provide users with aesthetically pleasing base image sets, which do not produce offensive or distasteful outputs when blended, e.g. weirdly disfigured creatures. The output of the GAN can sometimes be quite unpredictable, especially when the slider weights are set low. As a case in point, we have seen that a mixture of an image resembling a mango with an image resembling a shower head can give a rise to a dog, a cat or a squished human head. Weeding out such base images can be a tedious manual process, which we have partially alleviated by creating systematic line scans of the latent vector space for the different base image sets and glanced at the results by watching a fast-paced movie compiled from the output images.

Another issue in the selection of the base images is the support of creative intent. Users should be able to some extent predict the output of the GAN. When mixing images containing multiple objects or features, it is hard to judge

which features will make it into the resulting mix. Generally, we observe that difficulty of predicting the GAN output depends on the following factors:

- The number of base images provided

- The number of base images used in the target blend

- Presence of characteristic features in the base images that can be identified in the target image

- Distinctness of features across the alternative base image sets

- The relative volume of the parameter space that produces something similar to the target (provided the correct base images are blended)

Challenge mode image sets

In order to obtain suitable images for the Challenge mode, we have developed and applied a Monte Carlo-style algorithm, which generates random examples of basis image candidates and then evaluates them by simulating their use in blending. In order to reduce the count and location of distinct features in the base images, we require that a major fraction of the output images has:

- Uniform background: limit on the maximum color difference between selected pixels along the image border

- Smooth background: upper limit on the average Sobel gradient magnitude (density of sharp edges) in the border region

- Good foreground contrast: lower bound on the color difference between the background and selection of centrally located pixels

Once we have obtained a sufficiently large pool of base images, we have clustered them by the color of their background using the k-means algorithm with three clusters. We then draw randomly one image per cluster for each of the three alternative image set, which ideally produces balanced (similar) challenge item sets. In practice, manual selection was involved afterwards, to swap base images with inappropriate or very obvious features. We have also performed systematic slider space scans to ensure no majorly disturbing images arise when the users explore the space.

Discussion

GAN systems are particularly well-suited to producing rich and complex outputs from a relatively simple interaction. Although there are significant design choices involved in setting up such a system for casual creators, as described above, users of the system can produce a highly diverse array of possible outputs even with just three to five base images. Most importantly for assessment purposes, they do so with a constrained set of tools (in this case slider manipulation and base image selections) which, unlike paintbrushes or chisels, can be wielded in much the same way by almost anyone.

A primary advantage of the ease-of-use of such a GAN system is how the basic interaction can be used in a range of

task designs targeting specific cognitive components of creativity. Here, we have begun to exploit this by designing a series of tasks that vary in the specificity of the goal, allowing for a contrast between abilities supporting open-ended creativity and more goal-oriented creative tasks. These are relevant to studying divergent and convergent thinking (Guilford 1956), which may be supported by a common set of cognitive abilities, but are also considered to dissociate under some conditions (Chermahini and Hommel 2010; Chermahini and Hommel 2012). A deep dive into how divergent and convergent thinking are operationalized in the CREA suite can be found here (Rafner 2021)

Another advantage of a casual creators is their accessibility to a broad audience. Widening participation is important for testing subtle effects that require large participation numbers to be detected reliably. Additionally, and perhaps more importantly, it is a crucial property for creativity assessment tools as it must be broadly usable and capable of supporting creativity without relying on craft-specific competencies. *crea.blender* meets these requirements, since participants only need to choose some base images and manipulate their associated sliders in order to create a vast range of distinct artefacts and explore the complex high dimensional output space.

We hope *crea.blender* will help pave the path for the use of casual creators in studying creativity at scale and making GAN-based generators accessible to the general public.

Acknowledgments

References

[Berns and Colton 2020] Berns, S., and Colton, S. 2020. Bridging Generative Deep Learning and Computational Creativity. In *ICCC*, 406–409.

[Borji 2019] Borji, A. 2019. Pros and cons of GAN evaluation measures. *Computer Vision and Image Understanding* 179:41–65.

[Brock, Donahue, and Simonyan 2019] Brock, A.; Donahue, J.; and Simonyan, K. 2019. Large scale GAN training for high fidelity natural image synthesis. *arXiv preprint: 1809.11096*.

[Chermahini and Hommel 2010] Chermahini, S. A., and Hommel, B. 2010. The (b)link between creativity and dopamine: spontaneous eye blink rates predict and dissociate divergent and convergent thinking. *Cognition* 115(3):458–465.

[Chermahini and Hommel 2012] Chermahini, S. A., and Hommel, B. 2012. Creative mood swings: divergent and convergent thinking affect mood in opposite ways. *Psychological research* 76(5):634–640.

[Deng et al. 2009] Deng, J.; Dong, W.; Socher, R.; Li, L.-J.; Li, K.; and Fei-Fei, L. 2009. ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255.

[Goodfellow et al. 2014] Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014. Generative adversarial networks. *arXiv preprint: 1406.2661*.

[Guilford 1956] Guilford, J. P. 1956. The structure of intellect. *Psychological bulletin* 53(4):267.

[Guliford 1968] Guliford, J. P. 1968. *Intelligence, creativity, and their educational implications*. Edits Pub.

[Karras, Laine, and Aila 2019] Karras, T.; Laine, S.; and Aila, T. 2019. A style-based generator architecture for generative adversarial networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. IEEE. 4396–4405.

[Petrovskaya, Deterding, and Colton 2020] Petrovskaya, E.; Deterding, C. S.; and Colton, S. 2020. Casual creators in the wild: A typology of commercial generative creativity support tools. In *ICCC'20: Eleventh International Conference on Computational Creativity*. Association for Computational Creativity (ACC).

[Rafner et al. 2020] Rafner, J.; Hjorth, A.; Risi, S.; Phillipsen, L.; Dumas, C.; Biskjaer; Michael, M.; Noy, L.; Carsten, B.; Zana, B.; and Sherson, J. 2020. Crea.blender: A neural network-based image generation game to assess creativity. In *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, 340–344.

[Rafner 2021] Rafner, J. 2021. Creativity assessment games and crowdsourcing. In *Proceedings to the 2021 Annual Conference on Creativity and Cognition*.

[Runco and Jaeger 2012] Runco, M. A., and Jaeger, G. J. 2012. The standard definition of creativity. *Creativity research journal* 24(1):92–96.

[Simon 2021] Simon, J. 2021. Artbreeder. `https://www.artbreeder.com/browse`. Accessed: 2021-08-02.

[Torrance 1966] Torrance, E. P. 1966. *The Torrance Tests of Creative Thinking-Norms-Technical Manual Research Edition-Verbal Tests. Forms A and B-Figural Tests, Forms A and B*. NJ: Princeton, Personnel Press.