

‘@artbhot’ : Towards a bot with a Creative Presence on Twitter

ICCC 2022 Doctoral Consortium Submission

Amy Smith

amy.smith@qmul.ac.uk

Abstract

@artbhot is a Twitter bot that brings the generative capabilities of CLIP-guided GAN image generation to the public domain by transforming user-given text prompts into novel artistic imagery. Here, we aim to communicate several aspects of this project. Initially, we outline the progress in the project so far; discussing its implementation and some initial observations. We go on to discuss the future of the project, and what this mounting of generative technology amongst social media could mean for autonomous computationally creative agents, as well as some evaluative methods for creativity and autonomy. We also provide some illustrative examples of the bot’s usage and output.

@artbhot: The Project so far

[Note: A short paper on this work has been accepted to the conference this year under the title: ‘The @artbhot Text-To-Image Twitter Bot’].

Artbhot is a twitter bot (Veale and Cook 2018) that incorporates CLIP guided VQGAN to turn user tweets in to novel images (Esser, Rombach, and Ommer 2021; Crowson et al. 2022). Twitter users can tweet the bot using the following tweet format:

```
@artbhot #makeme tweet text
```

(e.g. @artbhot #makeme an oil painting of the Queen). Any text that follows the ‘#makeme’ hashtag will be sent to the generative model for processing.

Once an image has been generated, it is tweeted as a reply back to the user via the Twitter API, along with a reminder of the original tweet text. Please see Figure 1(a) for a diagram showing the different system interactions and flow of data. Please also see Figure 1(b) for an example of a user interaction with the bot, showing a user’s initial tweet to the bot using the tweet format mentioned, and also the reply to this tweet from the bot containing the generated image created using the text from the user’s tweet.

The initial iteration of Artbhot incorporated CLIP guided BigGAN for text-to-image generation; and at the time of building the first version of the bot this was one of the best CLIP guide GANs available. Later, VQGAN was released, and as the output from this newer model showed an improved curation co-efficient (i.e., much less cherry picking was required to find appropriate outputs) this model replaced CLIP + BigGAN for the text-to-image generation process.

The newest implementation of the bot is hosted on a remote server which runs 24 hours a day, and has currently been running for 6 months. The bot has processed over 600 tweets, taking, on average, around 2 minutes for a user to receive an image in response to their tweet. While there have been no outright failures (where images do not reflect the prompt at all) after an informal subjective evaluation (by ourselves) of the most recent 100 replies to Twitter prompts, we found only 16% of the images were not visually coherent enough to reflect the prompt satisfactorily.

The future of @artbhot

It is the overarching aim of this project to increase the creative autonomy of @artbhot. By this we mean that we would like the @artbhot to become a creative presence (Cook and Colton 2018) on Twitter, rather than an image generation service. Hence, in the future, the bot should be able to creatively interpret user tweet text and return challenging, surprising, and contextually aware creative output. Some examples of how we plan to achieve this are mentioned below.

As a first step to achieving this increased level of autonomy, we plan to introduce new functionality to the bot that will expand on the possible user interactions available with it. This also expands the scope for the ways in which the bot can respond to user input. This functionality will initially be in the form of different hashtags; that can be used as part of the current tweet format when tweeting at the bot. For example:

```
@artbhot #evolvethis prompt  
@artbhot #mergethis prompt
```

The ‘evolvethis’ hashtag will trigger a process of iteration over an image that is shown to the bot using this hashtag. The image could be iterated on in several ways, including using it as an input in to the generative model along with a text prompt. The text prompt could be derived from data on Twitter regarding trends. The ‘mergethis’ hashtag could indicate to the bot that two tweets should be merged in order to engineer a new prompt which can then be used as input to the generative process. Both these hashtags simultaneously increase the options users have and provide opportunities for more creative autonomy in @artbhot. We will discuss some more example of future bot functionality in the next section.

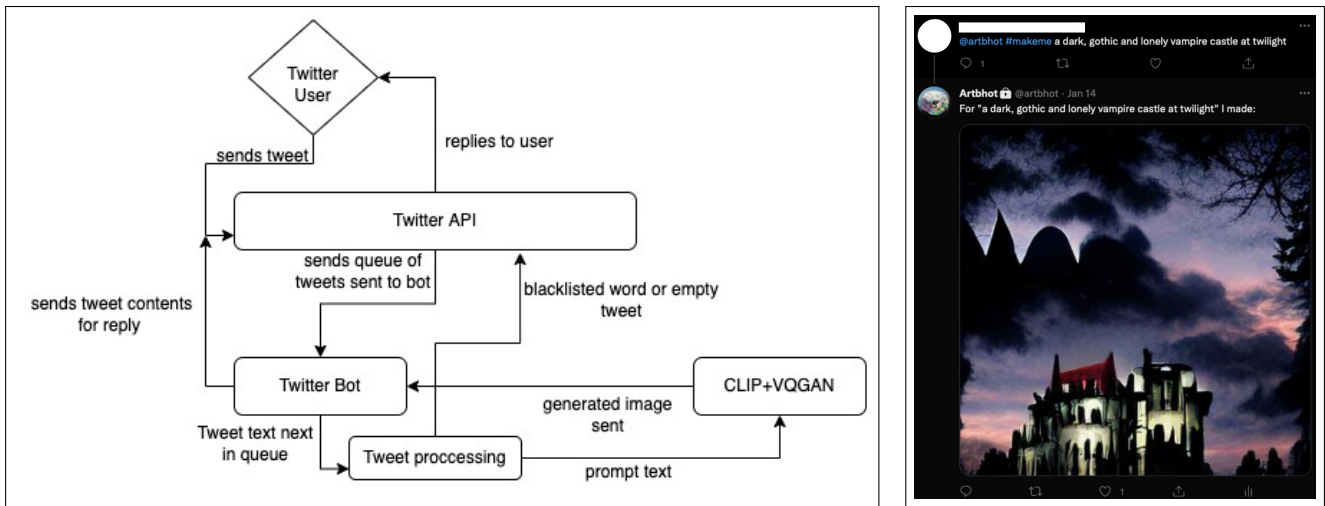


Figure 1: (a) Processing of a tweet by @artbhot (b) Example user interaction on Twitter.

Evaluating creativity and autonomy

One approach to assessing creativity is set out in Anna Jordanous' work on the '4 P's' (Jordanous 2016). This works explores the assessment of creativity from four different perspectives, considering: Person, Process, Product and Press. The following examples of future bot functionality illustrate how we aim to enhance @artbhot's creative presence and autonomy using these four areas as a basis for evaluation.

Person: The individual that is creative. Future functionality: The code can track which posts receive the most likes, and use this data to form preferences towards certain content. Data about the content of images that receive the most likes and retweets can be determined through various types of analysis (object detection, CLIP scoring, colour prevalence). Decisions about bot processes can be influenced by this preference when augmenting a user's tweet text to automatically engineer a prompt. For example, when augmenting a user prompt about the weather, Natural Language Processing (NLP) techniques such as syntactic analysis (Linzen and Baroni 2021) can be used to derive an understanding of the text that would allow the bot to engineer a text prompt stating that the sky in the generated image be depicted in the bot's favourite colour of the day. The bot could also tweet its favourite colour of the day based on this analysis.

Process: What the creative individual does to be creative Future functionality: The 'FloWr Framework' (Charnley, Colton, and Llano 2014) can be used for the automatic invention of bot processes. This framework is specifically designed for the optimisation and alteration of creative systems, and will enable the system to be innovative at the level of processes, as well as output. The framework generates novel flowcharts that map out bot processes, meaning that the system can decide for each different user how it should proceed when generating output.

Product: What is produced as a result of the creative process. Future functionality: Automatic prompt engi-

neering can produce more diverse and novel results. For example, merging a user's tweet text with information accessed through the Twitter API on what is trending on Twitter that day will produce images with visual content that contain contextually aware elements. It is also possible to:

- Introduce an implementation of the 'Disco Diffusion' text-to-image generative architecture as an alternative image generation technique.
- Implement CLIP + VQGAN in such a way that it accepts an image, as well as a combination of image and text, as the input.
- To implement CLIP + VQGAN to output animations based on the text prompt.

Press: The environment in which the creativity is situated Future functionality: As @artbhot is a Twitter bot, it already has some visibility and presence in its environment; in the sense that it is already interacting with twitter users and has published tweets that are accessible to other users. It can also derive information relevant to its processes from this environment, as discussed earlier regarding Twitter trends. Increased autonomous interactions in the future with twitter users will increase its presence on social media. These interactions can be influenced by information that can be stored other than trends information, for example: the bot may use information from tweets previously published by a user when interacting with that user. The bot can use the social media environment to not only publicise itself, but to derive a history of data that can influence its processes and responses. Its also possible that @artbhot interacts with other autonomous systems, such as PUCK (Johansen and Cook 2021). For example: PUCK could request art assets from @artbhot to be used in its automatic game invention processes.

There is also an existing methodology mapped out in the work by Colton et al. (Colton et al. 2014) on assessing



Figure 2: Generated images for prompts. **First row:** “A portrait painting of Frankenstein’s monster in the style of Monet”; “☀️☁️🌈”; “Aliens invading Newcastle Upon Tyne”; “A teddy bear in the art style of Salvador Dali”. **Second row:** “A positive lateral flow test”; “A velociraptor in an Edwardian bakery”; “A painting of Boris Johnson as the lion king in the style of Jean-Michel Basquiat”; “A cathedral made from rainbow gems”.

progress when building creative AI systems. This system can be used to evaluate future iterations of @artbhot in order to track the system in terms of how it is developing creatively. This system allows for the bot’s output and processes to be simultaneously assessed with a diagrammatic formalism that highlights where creative acts may be occurring in the construction and execution of the @artbhot’s system throughout the engineering process.

Illustrative Examples

Figure 2 presents multiple examples of output created using CLIP + VQGAN (Esser, Rombach, and Ommer 2021; Crowson et al. 2022) through @artbhot (the figure’s caption contains the corresponding tweet texts used to create the images). We found that @artbhot was able to handle unexpected prompts, for instance ones containing emojis. As per the second image in the first row of figure 2, CLIP-guided VQGAN interpreted the weather emojis correctly and produced an image with sun and clouds.

The Doctoral Consortium

Attending the Doctoral Consortium will benefit my project and PhD experience in several ways. Being assigned a senior CC-researcher as a personal mentor during the event is an extremely useful opportunity to ask questions and really get the most from my time at the conference. Having an expert panel offer advice and criticism is an invaluable chance to get a fresh perspective on my work, my career and other skills. The DC is also a perfect opportunity to network with other PhD Students and experts from closely related fields, as its not often that we get a chance to meet peers from other universities from around the world and learn about their research.

References

- Charnley, J. W.; Colton, S.; and Llano, M. T. 2014. The flow framework: Automated flowchart construction, optimisation and alteration for creative systems. In *Proceedings of ICCG*, 315–323.
- Colton, S.; Pease, A.; Corneli, J.; Cook, M.; and Llano, T. 2014. Assessing progress in building autonomously creative systems. In *Proceedings of ICCG*, 137–145.
- Cook, M., and Colton, S. 2018. Redesigning computationally creative systems for continuous creation.
- Crowson, K.; Biderman, S.; Kornis, D.; Stander, D.; Hallahan, E.; Castricato, L.; and Raff, E. 2022. Vqgan-clip: Open domain image generation and editing with natural language guidance. *arXiv preprint arXiv:2204.08583*.
- Esser, P.; Rombach, R.; and Ommer, B. 2021. Taming Transformers for High-Resolution Image Synthesis. *arXiv:2012.09841 [cs]*. arXiv: 2012.09841.
- Johansen, M., and Cook, M. 2021. Challenges in generating juice effects for automatically designed games. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 17, 42–49.
- Jordanous, A. 2016. Four perspectives on computational creativity in theory and in practice. *Connection Science* 28(2):194–216.
- Linzen, T., and Baroni, M. 2021. Syntactic structure from deep learning. *Annual Review of Linguistics* 7:195–212.
- Veale, T., and Cook, M. 2018. *Twitterbots: Making machines that make meaning*. MIT Press.