

Exploring Crowd Co-creation Scenarios for Sketches

Devi Parikh^{1,2} and C. Lawrence Zitnick¹

¹Facebook AI Research

²Georgia Tech

parikh@gatech.edu zitnick@fb.com

Abstract

As a first step towards studying the ability of human crowds and machines to effectively co-create, we explore several human-only collaborative co-creation scenarios. The goal in each scenario is to create a digital sketch using a simple web interface. We find that settings in which multiple humans iteratively add strokes and vote on the best additions result in sketches with the highest creativity (combination of value and novelty). Lack of collaboration leads to a higher variance in quality and lower novelty or surprise. Collaboration without voting leads to high novelty but low quality.

Introduction

How can one best collaborate with humans in a creative process? Insights towards this can inform what roles machines can (or should not) play when co-creating with humans.

Specifically, we consider a scenario where agents take turns collaboratively drawing a sketch on a simple web interface (Figure 1). During each iteration, multiple agents propose strokes to add to the sketch. Agents then vote on the proposals, and the preferred set of strokes is added to the sketch. This process is repeated for a fixed number of iterations to create a final sketch.

The roles of creating stroke proposals and voting could each be fulfilled by either humans (H) or machines (M). Borrowing terminology from Generative Adversarial Networks (Goodfellow et al. 2014), we can call the former role a generator (G), and the latter a discriminator (D). This allows for $4 \{H,M\} \times \{G,D\}$ co-creation scenarios. Further, different individuals could play the role of generators/discriminators across iterations, leading to *crowd* co-creation.

In this work, as a step towards human-machine co-creation, we study various *human-human* crowd co-creation scenarios. In the first, **Individual**, a single human creates the entire sketch (no discriminator D, and no crowd). Second, in **Collaborative** the sketch is generated by multiple human agents (crowd) iteratively taking turns adding strokes. That is, all the agents act as generators G and there is no voting or discriminator D. The third, **Collaborative + voting**, is where multiple human agents (generators) propose new strokes at each iteration. Another set of human agents (discriminators) vote on which set of strokes to add to the sketch. Finally, we explore **Individual with collaborative**

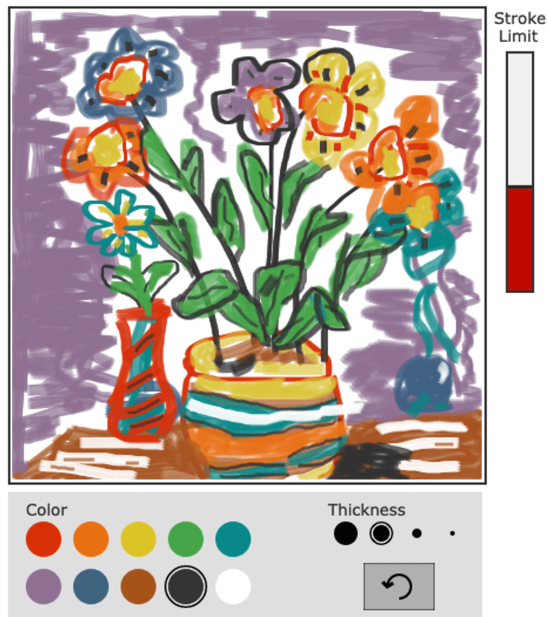


Figure 1: As a first step towards human-machine co-creation, we explore human-human collaboration for creating digital sketches on a simple web interface shown above. Video: <https://youtu.be/9fikuKPYPd0>

prompts, for which the crowd is involved indirectly. A single human creates the entire sketch, but by following text prompts that describe the evolution of a sketch that was created in the **Collaborative** scenario.

We evaluate the qualitative difference between the sketches produced via these four scenarios. We find that the collaborative setting *with* a voting mechanism (**Collaborative + voting**) leads to sketches that are rated by human subjects as most creative (and are preferred along a variety of other dimensions). The lack of either one of these components results in less creative sketches: **Individual** sketches have decent quality (value) but low novelty, while **Collaborative** sketches have high novelty but low value. **Individual with collaborative prompts** results in high novelty but even worse quality. Overall, among these four scenarios, **Collaborative + voting** best hits the sweet spot for creativity: combination of value and novelty (Boden 1992).

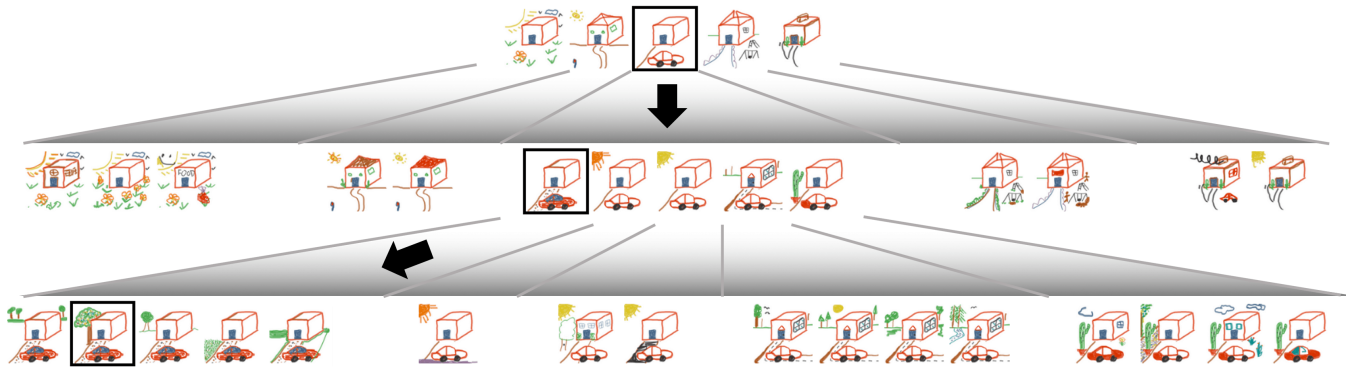


Figure 2: Few iterations of a sketch being created in the **Collaborative + voting** scenario. Once a parent sketch gets five children, it gets selected as the next iteration of the sketch (black outline), and the five children become the parents for the next iteration. Temporal visualization: <https://youtu.be/JQmGALAhhMU>.

Related Work

In the context of crowd-based sketching, (Tuite and Smith 2012) analyze user actions and large-scale behavior patterns in 50k sketches from Sketch-a-bit, a collaborative mobile drawing application. Different from our incremental contribution and voting mechanisms, (Yu and Nickerson 2011) and (Gingold et al. 2012) explore combination and averaging of sketches respectively as collaboration strategies.

Several AI systems have been trained to recognize sketches (e.g., models trained on The Quick, Draw! Dataset¹). These may form useful building blocks for the next stages of our work. However, as seen in Figure 3, our sketches tend to be complex scenes and often abstract as opposed to concrete individual objects, which has been the focus of most existing work in automatic sketch recognition. There is also work on generating images based on sketches (Chen and Hays 2018).

(Davis et al. 2016) employ a cognitive science framework called participatory sense-making to study co-creation in sketches. Central to their study is the back and forth interaction (dialog) between the human and machine as they take turns. Our work is focussed on a crowd setting where no two agents interact again in the future. (Karimi et al. 2020) study human-AI co-creativity in the context of humans sketching for a particular design goal. Our work falls in the category of “casual creators” (Compton and Mateas 2015) – systems that support exploratory as opposed to goal-driven creativity.

Sketching Interface

Human agents create sketches using the JavaScript based interface shown in Figure 1. Strokes can be varied across four thicknesses and ten colors and have a paint-like texture. The number and length of the strokes an agent may draw is limited during each iteration. Feedback on how close they are to the cutoff is provided in real-time by the stroke limit bar. Thicker strokes count more towards the limit. Strokes drawn by the agent during the current iteration may be undone. See <https://youtu.be/9fikuKPYPd0> for a video of the interface. Our interface is publicly available.

¹<https://github.com/googlecreativelab/quickdraw-dataset>

Co-creation Scenarios

We explore four scenarios for collaborative human-human sketch co-creation. In every scenario, the sketch starts with a blank canvas. During each iteration, a limited number of strokes may be added. The limit roughly corresponds to five medium-thickness strokes spanning the width of the canvas. 30 iterations are used to create each sketch. Unless stated otherwise, we collected 20 sketches for each scenario. All our studies were conducted on Amazon Mechanical Turk. Subjects can not submit their work till they have contributed the required amount of strokes to the canvas.

Individual. The entire sketch is created by a single individual. That is, a single human agent adds all 30 iterations of strokes to “Create a beautiful, detailed, coherent painting!”.

Collaborative. A different human agent contributes strokes for each iteration of the sketch. That is, 30 unique individuals contribute to a sketch. The first subject sees a blank canvas and adds strokes. Every subsequent subject is shown the partial sketch and asked to add to it. They cannot undo strokes from earlier contributors. The prompt is “Let’s collectively create a beautiful, detailed, coherent painting!”. Subjects are given the additional instruction to consider the kind of painting being created and the stage of the painting when deciding upon which strokes to draw.

Collaborative + voting. Each subject contributes strokes to a sketch of their choosing from a set of five starting sketch variations. We refer to the chosen starting sketch as a parent, and the sketch created by a subject as the chosen sketch’s child. During each iteration, sketches are gathered until a parent is selected five times. Its children then replace the current five parents and the process is repeated. Children of parents selected less than five times are discarded. See Figure 2 and <https://youtu.be/JQmGALAhhMU>.

This voting strategy allows for the most promising versions of a sketch to go forward. This scenario is robust to the strokes added by any one individual. Of course, it is also significantly more “expensive”. In the best case scenario where a single parent gets all 5 children and none of the other parents get a child, it takes 5 times the amount of strokes to create a sketch compared to **Collaborative**. In the worst case, all 5 parents get 4 children each before a parent



Description of [row] relative to [column]	Individual	Collaborative	Collaborative + voting	Individual with collaborative prompts
Individual	N/A	simpler, plain, lacking depth, with a focus on specific everyday objects, clear, neat, stagnant, calm, polished, refined, realistic, deliberate, playing it safe, with a central theme	more peaceful collection of artwork that is relaxing, less sharp, less refined, less colorful, more gloomy	[Did not compare]
Collaborative	more abstract, detailed, complex, wild, chaotic, sloppy, more emotion, fluid and moving, sadder, distorted and exaggerated shapes and objects, harder to understand	N/A	more abstract, drawn by younger artists, more exciting, more strokes, more scribbles, capture complex or disturbing emotions, more colorful	more psychedelic, vibrant, busy, detailed, polished, sophisticated, joyful, more mature artist, more complete, more color
Collaborative + voting	having more of a story to them, lively, complex, imaginative, stirring, happier, sharper, refined, and detailed	more clear, careful, finely drawn, of higher quality, and looked more like a finished artifact	N/A	[Did not compare]
Individual with collaborative prompts	[Did not compare]	more primitive, having simpler figures that were easier to identify, and with more harmonious strokes	[Did not compare]	N/A

Figure 3: Example sketches from four co-creation scenarios along with differences identified by human subjects between sketches from pairs of scenarios. **Collaborative + voting** involves ~ 12.5 times the individuals, and so was run for 20 instead of 30 iterations. For comparison, **Collaborative** sketches are also shown at 20 iterations.



Figure 4: Example prompts used in the **Individual with collaborative prompts** scenario.



Figure 5: Evolution of example sketches in the **Collaborative** scenario. Left: Focus of the sketch shifts from the house to the cat in the rain outside the house. Right: Faced with seemingly incoherent strokes, subjects emphasize structure they see in it so subsequent subjects can add to it.

gets a fifth child. This would result in 21 times the number of strokes. In practice we found this factor to be about 12.5 times. Given the increased cost, we reduced the number of iterations in this scenario to 20 (as opposed to 30). On average, 250 unique individuals contribute to a single sketch.

Individual with collaborative prompts. A single individual creates an entire sketch using instructive text prompts provided at each iteration. The individual is instructed to follow the prompts when drawing. The text prompts are generated by asking another individual to describe what changed in a sketch from one iteration to the next in the **Collaborative** scenario. All text prompts for a sketch are written by a single individual. This is an interesting hybrid of having a single creator, but being guided through prompts that describe the evolution of a sketch as created by 30 unique individuals. We collected three sets of text descriptions for each of the 20 **Collaborative** sketches. This resulted in a total of 60 **Individual with collaborative prompts** sketches. In our evaluation, we consider 20 sketches (randomly picking 1 out

of the set of 3). See Figure 4 for example prompts.

Evaluation

Example sketches from these scenarios are shown in Figure 3. Before we discuss properties of the final sketch, it is worth considering the evolution of a sketch as it is being created. **Collaborative** sketches evolve in several interesting ways: what seems like the main subject of a sketch changes in a few iterations (Figure 5, left), given seemingly incoherent strokes, subsequent subjects try and emphasize regions that could lead to meaningful structures in the sketch for future subjects to build on (Figure 5, right), and subjects use the color white or other strategies to try and cover parts of the sketch they think are contributing negatively to it.

To assess the qualitative differences between sketches produced from the 4 scenarios, we created a collage of 20 sketches from each scenario (at 20 iterations for **Collaborative** and **Collaborative + voting**, 30 for the rest). We showed pairs of collages to subjects on Amazon Mechanical Turk and asked them to describe differences that stood out. Snippets from subjects' responses are shown in Figure 3.

For a quantitative evaluation, we showed subjects pairs of sketches from two different scenarios (i, j). Each subject picked which sketch they prefer along 12 axes. Every pair was evaluated by 5 subjects resulting in 144,000 assessments: 20 (sketches from scenario i) \times 20 (sketches from scenario j) \times 6 (pairs of scenarios) \times 12 (axes) \times 5 (subjects per sketch-pair). The 12 axes were: which painting (1) seems more strange / unusual / different than typical paintings? (2) is more creative? (3) do you like looking at more? (4) is more creative? (5) is more interesting? (6) is more original? (7) took more skill? (8) is made by an artist more likely to be an adult? (9) would you pay more for? (10) are you more likely to put up in your home? (11) is more likely to be in an art museum? (12) would you be more proud to have made yourself? Some of these axes (e.g., originality, novelty, skill) are from (van der Velde et al. 2015).

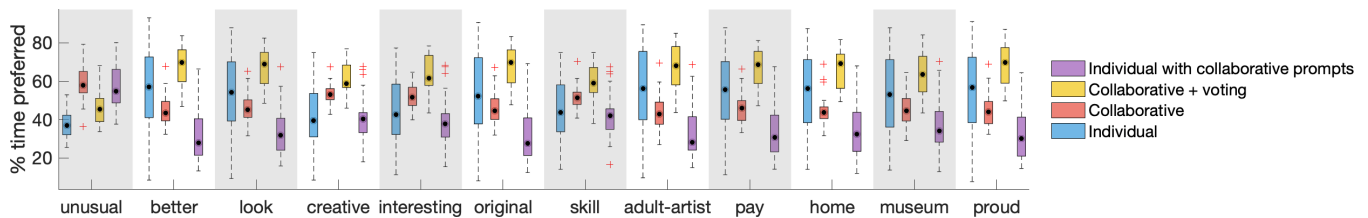


Figure 6: **Collaborative + voting** sketches are consistently preferred by human subjects over sketches from other scenarios across a variety of dimensions, and notably are rated as most creative. Notice the high variance in **Individual** sketches.

The percentage of times each scenario was picked over a competing scenario is shown in Figure 6. For 11 out of the 12 axes, including creativity, **Collaborative + voting** is preferred. **Collaborative + voting** scores well for both novelty (unusual) and quality (better, look), which we hypothesize increases its perceived creativity. **Individual** is rated well for quality but scores poorly on novelty. Across 11 out of the 12 axes, **Individual** has high variance due to differences in skill and motivation of individuals creating the sketches. **Collaborative** scores well on novelty, but worse on quality. **Individual with collaborative prompts** does poorly across all axes except for unusual. Of all scenarios, **Collaborative + voting** falls in the sweet spot for maximizing creativity (combination of value and novelty).

Discussion

In what way may a machine best contribute to the collaborative creation of a sketch? It is often the case that humans may not be good at generating strokes, but can tell if a sketch looks good or not. This may suggest using machines to generate candidate strokes and having humans vote on which versions should proceed next. The machine may also contribute in a manner similar to the humans in our fourth scenario, i.e., the machine could generate textual prompts as a human draws a sketch. The prompter can have different “personalities” based on whether it is trained on sketches generated from **Individual** (coherent), **Collaborative** (rich but chaotic) or **Collaborative + voting** (rich with subtle details and coherent). Humans and machine can generate strokes as a team, either in co-painting scenarios as in (Cabannes et al. 2019), or where the machine provides some visual guidance as in (Lee, Zitnick, and Cohen 2011) or via suggestions for where to draw, what colors to use, etc. as explored in (Oh et al. 2018). We can also train a machine to be a discriminator: given a few different strokes from a human, select which stroke should be added to the sketch next.

All our sketches started with a blank canvas. We could instead start sketches with a prompt (subject of the sketch, adjective describing a desired property of the sketch, a picture to be used as inspiration for the sketch, etc.), and have this prompt persist across iterations (or not).

Our collaboration scenarios can be enriched by allowing for repeated interactions between the individuals contributing to the sketch, creating opportunities for a dialog.

It is interesting to consider ideas of ownership in the context of crowd co-creation. While no one individual may feel a complete sense of ownership of the final piece, crowd col-

laboration may lead to a sense of community and the satisfaction of contributing to a common cause. Finally, while our motivation was human-machine co-creation, studying human-human collaboration in general is, obviously, important and interesting in and of itself. Collaborative creative endeavors may be a fertile ground for such explorations.

References

- Boden, M. 1992. *The Creative Mind*. London: Abacus.
- Cabannes, V.; Kerdreux, T.; Thiry, L.; Campana, T.; and Ferrandes, C. 2019. Dialog on a Canvas with a Machine. In *Creativity Workshop at NeurIPS*.
- Chen, W., and Hays, J. 2018. SketchyGAN: Towards Diverse and Realistic Sketch to Image Synthesis. *arXiv:1801.02753*.
- Compton, K., and Mateas, M. 2015. Casual Creators. In *ICCC*.
- Davis, N.; Hsiao, C.-P.; Singh, K. Y.; Li, L.; and Magerko, B. 2016. Empirically Studying Participatory Sense-Making in Abstract Drawing with a Co-Creative Cognitive Agent. In *IUI*.
- Gingold, Y.; Vouga, E.; Grinspun, E.; and Hirsh, H. 2012. Diamonds From the Rough: Improving Drawing, Painting, and Singing via Crowdsourcing. In *HCOMP*.
- Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A. C.; and Bengio, Y. 2014. Generative Adversarial Networks. *NIPS*.
- Karimi, P.; Rezwana, J.; Siddiqui, S.; Maher, M. L.; and Dehbozorgi, N. 2020. Creative Sketching Partner: An Analysis of Human-AI Co-creativity. In *IUI*.
- Lee, Y. J.; Zitnick, C. L.; and Cohen, M. F. 2011. ShadowDraw: Real-time User Guidance for Freehand Drawing. *SIGGRAPH*.
- Oh, C.; Song, J.; Choi, J.; Kim, S.; Lee, S.; and Suh, B. 2018. I Lead, You Help But Only with Enough Details: Understanding the User Experience of Co-Creation with Artificial Intelligence. In *CHI*.
- Tuite, K., and Smith, A. 2012. Emergent Remix Culture in an Anonymous Collaborative Art System. *AIIDE*.
- van der Velde, F.; Wolf, R. A.; Schmorrow, M.; and Nazareth, D. S. 2015. A Semantic Map for Evaluating Creativity. *ICCC*.
- Yu, L., and Nickerson, J. V. 2011. Cooks or Cobblers? Crowd Creativity through Combination. In *CHI*.