

1 **CollabColor: Creative Support System for Human-Human Synchronous**
2 **Collaboration**
3

4 ANONYMOUS AUTHOR(S)
5

6 Humans are unique in working collaboratively by sharing and understanding intentions. However, digital collaboration is daunting,
7 especially in creative design life cycles, due to non-linear workflows and lack of micro-alignments coupled with the need for robust
8 network connectivity. We present a formative study with creatives to identify key themes in conflicts that arise in this space. We
9 introduce CollabColor, an intelligent system that aids in resolving conflicts for two users synchronously collaborating on a low-touch
10 creative task. More specifically, given an uncolored line-art on a canvas and a set of reference images from the users as input, we
11 provide non-obtrusive interventions during their real-time collaboration to ensure that the final colorization of the art is coherent,
12 and all the users' aligned preferences are incorporated. CollabColor is a novel reinforcement learning-based approach that processes
13 the entire sequence of canvas states and user actions to provide assistance in the form of coloring actions that builds trust between
14 users, resulting in the discovery of a shared perspective and a win-win solution for their co-creation. We conduct extensive evaluation
15 using automated metrics and human studies showing that our interventions lead to coherent and preference-aligned colorizations. We
16 believe that CollabColor is an early step towards building automated social planning systems that facilitate and foster cooperative
17 interactions of a population to solve various global challenges.
18

19
20 CCS Concepts: • **Human-centered computing** → **User interface design; Collaborative interaction;** • **Computing methodologies**
21 → **Reinforcement learning.**
22

23 Additional Key Words and Phrases: collaboration, creativity, reinforcement learning, neural networks, support system
24

25 **ACM Reference Format:**

26 Anonymous Author(s). 2022. CollabColor: Creative Support System for Human-Human Synchronous Collaboration. In *IUI '22:*
27 *ACM Conference on Intelligent User Interfaces, Mar 22–25, 2022, University of Helsinki, Finland*. ACM, New York, NY, USA, 38 pages.
28 <https://doi.org/XX.XXXX/XXXXXXX.XXXXXXX>
29

30 **1 INTRODUCTION**
31

32 Collaboration is a process of accumulating shared information, aligning on choices, and acting on it together [112].
33 Collaboration happens everywhere – ranging from music bands to business meetings, and in co-editing documents. In
34 nature, we find several examples of species of ants or bees that live in large colonies and cooperate with optimal work
35 division to collect food and survive [51]. Ever since the Covid-19 pandemic raged, companies and institutions have
36 accelerated towards a “digital-first” approach of collaboration. However, due to lack of micro-alignments, excessive
37 hidden work, bad network connectivity, and an overall lack of common view, collaboration has shifted from unifying
38 and exciting to time-consuming and daunting [3].
39

40 The perils of digital collaboration are worsened in creative design workflows [104]. Art directors provide reviews
41 across the entire creative life cycle leading to a non-linear workflow with significant back-and-forth between ideation,
42 creation, and the feedback phases. In particular, the creation phase involves lots of conflicts among co-creators when
43

44 Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not
45 made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components
46 of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to
47 redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
48

49 © 2022 Association for Computing Machinery.
50 Manuscript submitted to ACM
51



Fig. 1. Within CollabColor’s user interface, two users S1 (top screen) and S2 (bottom screen) can understand each other’s perspective and arrive at a coherent artwork. An angry emoticon intervention (1) is displayed on S2’s screen because she chose green color most recently to fill the cloth on elephant which clashes with S1’s intentions of orange/red color for the cloth. The color for elephant’s body in S2’s Peek intervention output (2) is a very light shade of green that is the closest to S1’s preferred light colors. Interventions such as peeking, emoticons, collaboration score bars facilitate communication of intentions between users, resulting in fewer conflicts.

choosing appropriate colors, fonts, shapes, etc. wherein simple merging would lead to unpleasant outputs. Carefully delegating individual tasks to creators and forgoing collaboration could be a solution, but a recent work by Parikh and Zitnick [86] found that individual creators provide value, whereas collaboration leads to more novel and surprising designs. This motivates us to focus on easing collaboration in the hard creation phase.

Prior to 1962, machines were largely seen as tools for solving heavy numerical problems. In 1962, Douglas Engelbart’s [19, 29] proposal of using machines to augment human intelligence propelled researchers to think of them as real-time interactive systems. Today, humans and machines collaborate in various disciplines to produce highly creative outcomes [33, 41, 93]. Recent studies on human-AI co-creativity [50, 71] discuss the perception and utility of AI tools in supporting creative work, especially in the discovery stage of design projects. However, the situation of two or more humans collaborating on a creative task with AI as a support tool is still in nascent stages of exploration. A genre of creative support tools called Casual Creators¹ [24, 80, 88, 106] have only been deployed recently as test-beds to understand this scenario better. Building an AI system that aids two or more users not only in casual co-creative explorations, but also in solving creative tasks poses several challenges: (i) the system must understand the creative motivations of all users in context of the current state of artwork and arrive at a *high-level* shared perspective that adds value; then, (ii) the system must promote *low-level* conflict-resolving actions during the co-creation process while not obstructing or limiting the imagination of users.

¹Casual Creators are tools that allow users to pursue their creative intentions casually without any pressure of achieving certain goals

In this paper, we consider flat line-art colorization as our creative task motivated by the following reasons: (a) Given the ambiguity in subjective definitions of creativity, the vast literature on colorization helps us define metrics to build our system; (b) Flat colorization offers greater computational feasibility as compared to other creative tasks in exploring a vast design space [13]; and (c) Research in psychology has shown that coloring aids in mindfulness and reduction of anxiety in adults [72].

We introduce CollabColor, an intelligent reinforcement learning-based system that aids in resolving conflicts for two users synchronously collaborating on flat line-art colorization. More specifically, given an uncolored line-art on a canvas and a set of reference images from the users as input, we provide non-obtrusive interventions during their real-time collaboration to ensure that the final colorization of the art is coherent, and all the users' aligned preferences are incorporated. Drawing inspiration from Mary Scannell's Conflict Resolution games [95], we design a novel transformer component in our system that sequentially processes sub-optimal canvas interaction trajectories and learns to facilitate conflict-resolving actions among users during deployment.

This paper contributes:

- Insights into key conflicts that occur during co-creative task of flat line-art colorization, supported by a formative study with creatives and a web-deployable implementation of CollabColor system to collaborate on the task
- The design of novel data simulator and reward functions that aid in preparing realistic trajectories of canvas interactions, using techniques such as Behavior Cloning and Theory of Mind to imbue human biases
- The design of a transformer-based model architecture that learns from simulated interaction trajectories to give greater importance to conflict-resolving actions
- Non-obtrusive UI interventions such as buttons for peeking, color prediction, & feedback scores derived from the trained model to support creativity and ease collaboration of users working on CollabColor interface
- Evaluations with automated adversarial users to Stress test CollabColor, showing our effectiveness in incorporating preferences of both users and reducing conflicts in difficult real-world scenarios
- insights into flexibility that our system offers, using Recurrence Quantification Analysis
- Human studies and exploratory surveys on colorizations produced by CollabColor that show significant gains in novelty of colorization, with system-suggested interventions aligning with those of human facilitators.

We envision CollabColor as an early step towards building automated facilitators that can not only aid in creativity, but also in other social and professional domains to build trust, shared intentionality [103], and bring populations together to solve pressing global challenges [27].

2 RELATED WORK

CollabColor builds on related work in five key areas: (a) studies on support systems that enhance creativity; (b) automatic colorization techniques and datasets/metrics therein; (c) studies on simulating realistic trajectories to train multi-agent reinforcement learning (RL) systems; (d) deep neural RL models that learn from suboptimal offline trajectories; and (e) human autonomy training (HAT) literature for evaluating our proposed system. We provide further discussion on related work in Appendix B.

2.1 Support Systems for Collaboration and Creativity

CoDraw [52] presents two AI-based neural models called drawer and teller that interact to generate a clipart-based scene. Each of these models can be used in isolation to assist humans in generating creative scenes. While this provides

157 us a framework to reason about, we build a system that has a model to assist *two* humans collaborating on a creative
 158 task. There are several works that consider this teller-drawer architecture to design assistive tools. For instance, Creative
 159 Sketching Partner [50] presents an AI drawer that sketches variants for ideation based on user inputs of visual and
 160 conceptual similarity. Casual Creators such as GANimals [30] foster data collection for human-human co-creativity but
 161 do not consider ways to resolve conflicts.
 162

163 There are other examples of interactive support systems such as Vocal Shortcuts for designing [53], StreamSketch
 164 for livestreaming [70], Ideawall for collaborative ideation [99], and Scones [46] for sketching via natural language
 165 commands but none of them build strong models for processing long sequences of user interactions and canvas states,
 166 or intervene non-obtrusively to reduce conflicts.
 167

168 While no previous work defines metrics for creative conflicts during colorization, Gu et al. [38] devise techniques for
 169 understanding and aiding conflicts in textual domain. Kuiter et al. present variED [57] and define conflicts and data
 170 structures for collaborative editing in the context of coding software. Similarly, Owhadi-Kareshk et al. [84] devises
 171 features out of GitHub code versions to predict merge conflicts. We note that textual conflicts are easier to define using
 172 metrics such as Edit Distance [121], as compared to creative conflicts that are ambiguous and subjective.
 173

175 2.2 Interactive Colorization

177 Several works in Computer Vision literature tackle the problem of colorization of raster images [32, 65, 123? , 124].
 178 Zhang et al. [123] devise a convolutional neural network that colors grayscale images into RGB images. Barnes et al.
 179 [13] point out the vast design space of creative tasks like colorization and posit that user interactivity is paramount to
 180 good colorizations. Motivated by this, Zhang el al. [124] further fine-tune their architecture to allow for interactivity
 181 wherein users can pick pixels in the image and enlist their color preferences. However, none of these work in the
 182 context of co-creativity that we approach in this paper. Furthermore, we approach colorization from the perspective of
 183 SVG images instead of raster images. Prior work [14, 34] does exist for SVG colorization, but again not in the context of
 184 iterative co-creation. Nevertheless, we rely on extensive work in colorization literature to define reward signals such as
 185 harmony score [83, 94], and also to conduct human studies. We provide further discussion of colorization datasets in
 186 Appendix C.
 187

189 2.3 Simulation for Multi-agent Reinforcement Learning

191 RL training can be seen as an alternating optimization between finding good trajectories and finding good policies over
 192 those trajectories [58]. Previous methods avoid optimizing for good trajectories by fixing them to episodes played by
 193 expert demonstrators, and train RL agents to mimic the experts using imitation learning [9]. However, demonstrator
 194 data collected in the formative study is neither sufficient nor diverse enough to train data-intensive deep supervised RL
 195 algorithms. Also, RL training requires rewards either at the end of a sequence (sparse reward), or at each action turn
 196 [64], which is not practical in online setups.
 197

198 Recently, there has been a surge of self-play algorithms [42, 81] where the agents play with themselves and generate
 199 good trajectories. This forgoes the expensive process of collecting large and diverse demonstrator data but requires very
 200 long hours of training. We therefore simulate a dataset consisting of a sequence of user actions and approximate rewards
 201 at each action turn without self-play. Like many previous works [58, 75, 78] we design a rule-based simulator along with
 202 reward functions that approximate users' payoff intentions. While such a simulation process has disadvantages like the
 203 sim2real gap [47, 77, 125], the lack of interaction with real world environment can be offset by powerful supervised
 204

209 algorithms. Another difficulty is the requirement of strong domain knowledge to design reward functions [64]. We use
 210 data from our formative study to impart domain knowledge and inductive biases into rewards.
 211

212 2.4 Reinforcement Learning for Human-AI Collaboration and Creativity

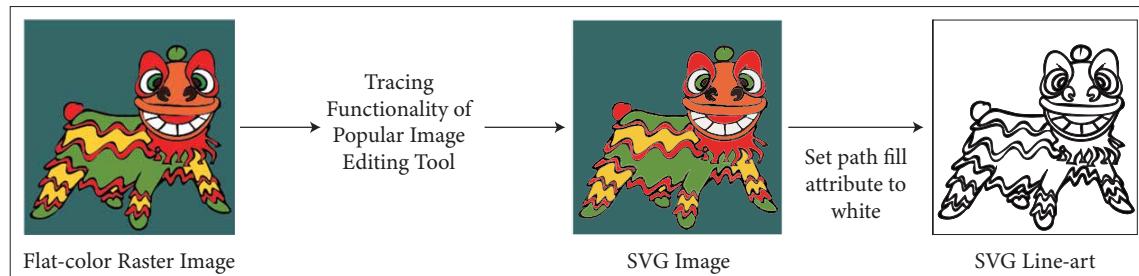
213 Knott et al. [56] & Wu et al. [118] train agents for human-AI collaboration, but their tasks are confined to the Overcooked
 214 environment [18] with well-defined sub-task division, explicit actions for cooperation, and visible spatio-temporal
 215 conflicts. For instance, Wu et al. [118] employs principles from cognitive science [54] to build a multi-agent RL with two
 216 AI agents that can coordinate to divide cooking actions into sub tasks and allocate low-level actions optimally. However,
 217 research shows that such agents fail to coordinate and generalize when one of the players is a human [18, 22, 56].
 218 Humans do not always take optimal actions that an AI agent might predict in a multi-agent setup, leading to poor
 219 rewards at inference time. We employ techniques such as Behavior Cloning [12] and Theory of Mind [37] to explicitly
 220 model humans [18] and arrive at richer simulated trajectories to alleviate issues of generalizations.
 221

222 While previous works [49, 52, 61, 87] train RL agents for tasks such as coloring, sketching, or retouching in a
 223 canvas-like environment, these tasks are not considered in the context of multiple user collaboration. We leverage a
 224 3-agent real-time interactive setup that explicitly models humans and their preferences in CollabColor. We investigate
 225 in a more challenging creative task space of flat colorization wherein no obvious creative sub task divisions exist,
 226 cooperation occurs across a wide range of turns, and conflicts are not immediately perceivable.
 227

228 3 FLAT COLORIZATION OF LINE-ARTS: INTERFACE DESIGN & CHALLENGES

229 Our goal is to obtain a sufficiently large dataset of uncolored line-arts and their corresponding colorized images that
 230 can be leveraged to simulate two users iteratively coloring an image². Since colorization is not the main contribution of
 231 our work, we reduce our task to *flat colorization* of line-arts which is easier to deal with. We also require the line-arts
 232 to have sufficient number of semantic segments so that the simulated sequences capture meaningful patterns and
 233 interactions between the users.
 234

235 3.1 Preparing Data for Colorization of Line-arts



236 Fig. 2. We use a popular image editing tool's built-in tracing functionality to extract SVG image (colored) and uncolored line-arts.
 237

238 We obtain flat colored raster images from the web and convert them to SVG images. Inspired by automatic generation
 239 of fake images using Photoshop Scripting [111], we process a large batch of raster images and convert them to SVGs
 240

241 ²Note that we did not find publicly available datasets that matched our requirements, hence the need to first create our own uncolored images.
 242

using built-in tracing functionality of a popular image editing tool. Although raster-to-vector tracing is a hard task [63], flat colored images with closed line segments are easier to trace. We thus obtain our dataset of 255797 SVG images of 4418 unique line-arts, boiling down to nearly 58 alternate colorings per line-art. This captures the inherent multimodality of the coloring process [20] and ensures that our images cater to a diverse set of users. Figure 2 shows a schematic diagram of our data preparation pipeline. Figure 3(a) shows an uncolored line-art with six alternate colorings and associated tags. From Figure 3(b), we can see that our images in our dataset contain 30 segments on average, making it feasible for two users to iteratively color and impart their preferences. For our task, each SVG image can be considered as a list of segments. Each segment contains features such as area, centroid, color, and path length which are computed with the help of the `svgpathtools` [2] library. We discuss alternate approaches of preparing data for our task in Appendix C

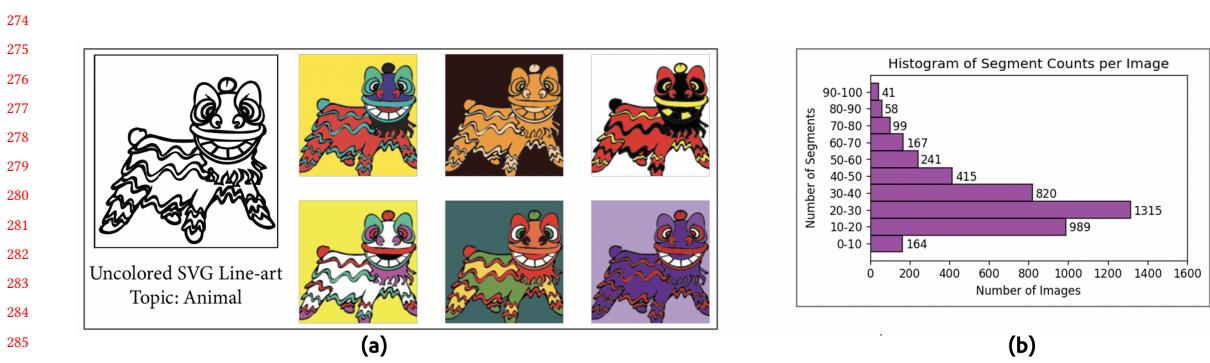


Fig. 3. Part (a) demonstrates six alternate colorings of the same line-art capturing the inherent multimodality and diversity of colorization process [20]. Part (b) demonstrates a histogram of the number of segments per SVG image in our data. Values are computed over unique line-arts.

3.2 User Interface Design

To understand the pain points of creative collaboration, we build a user interface (UI) that enables two users to color a line-art simultaneously. Figure 4 shows screens of two users (S1 and S2) coloring the line-art of an elephant. Our UI is inspired from collaborative interfaces such as Miro, Figma³, etc. to provide a realistic experience to the users. We use Twilio's[6] real-time state synchronization API to enable synchronous coloring. Collecting reference/inspiration images into a moodboard to diversely color an artwork is a standard practice in designer's toolbox [71]. We mimic this scenario by providing two inspiration images to each user that cater to certain emotions, using the Behance Artistic Media (BAM) dataset [114]. Here, user S2 gets peaceful/calm images in the form of sceneries, large green fields, etc. Colors can be picked either from the palette or from the canvas. We also display the most recent color picked by each user to aid in understanding each other's preferences. Here, users S1 and S2 have last picked violet and green colors respectively. Once both the users are done with iterative coloring the current line-art and click the “Done/Next” button, we proceed to the next line-art.

³<https://miro.com/>, <https://www.figma.com/>



Fig. 4. Screens of two users S1 (top) and S2 (bottom). Users can pick colors from the palette based on their inspiration images to color the line-art displayed in the canvas. The various components of our UI are labelled in the red boxes. User S1’s inspiration images present a happier emotion with bright hues; User S2’s images present a peaceful emotion with light green hues

3.3 Formative Study

We conduct a formative study in three modes by altering the type of reference images displayed (no images, ground-truth (GT) images, & inspiration images) with 20 user pairs consisting of novices (undergrad students) and experts (designers). We ask the users not to communicate via any other medium beyond the canvas. While several collaborative platforms provide users with options to video/voice-call or chat via text, we intend to study scenarios wherein users do not have means to communicate either due to poor network quality or lack of cognitive bandwidth [39]. At each click (action / turn) t for a given SVG image on canvas, we record a triplet (u_t, s_t, c_t) corresponding to the user u_t (u_1 or u_2) coloring a segment s_t (segment ID in the SVG image) with color c_t . We also restrict the number of turns per user to $T = 100$. Additionally, we conduct informal interviews and query users about their experience and ways in which they prefer to be assisted. We identify several key themes in **conflicts** that arise while collaborating on our creative task.

F1: Understanding intentions is hard. Users found it difficult to relay their intentions just using colors on the canvas. Given that our task involves creative decision making, communication becomes paramount to arrive at a coherent colorization. Users could not verbally communicate in our setup, leading to several conflicts. They take multiple actions with excess *recoloring* in an attempt to align their preferences. Table ?? shows the average number of actions taken to color line-arts. We see that when inspiration image were not provided, the number of actions taken were more than twice the average number of segments. Around 12% of all sequences had lengths greater than 80, indicating the long duration of the task. *Cross-recoloring* conflicts are particularly dissatisfying to the users, wherein a user alters the color of the segment colored by the other user. Formally, a recoloring action is recorded when a user alters the color

365
366
367 Table 1. Formative study of our user interface with three modes varying the reference images (no image, GT images, inspiration
368 images from BAM dataset). The value N indicates the number of SVG line-arts colored by our user pairs. Arrow mark (\uparrow) indicates
369 that higher value is typically better.

Mode	Avg no. of actions (↓)	% Recoloring (↓)	% Cross Recoloring (↓)	% Reversion (↓)	Dominance Score (↓)	Common- sense Score (↑)	Harmony Score (↑)
No Image (N = 77)	62.6	31%	17%	8.0%	0.32	0.11	0.34
GT Images (N = 20)	49.8	28%	10%	11.0%	0.37	0.26	0.41
Inspiration Images (N = 229)	33.7	30%	12%	11.7%	0.47	0.21	0.42

380
381 of an already colored segment. Let $S_1, S_2, \dots, S_{|S|}$ represent segment IDs of a given line-art where the total segment
382 count is $|S|$. Then, the count of recoloring actions is given by $\sum_{t=1}^T \sum_{i=1}^N I(s_t = S_i) - |s_1, s_2, \dots, s_T|$, where $I(\cdot)$ is the
383 indicator function ⁴, and $|\cdot|$ is the count of unique segments in each sequence.

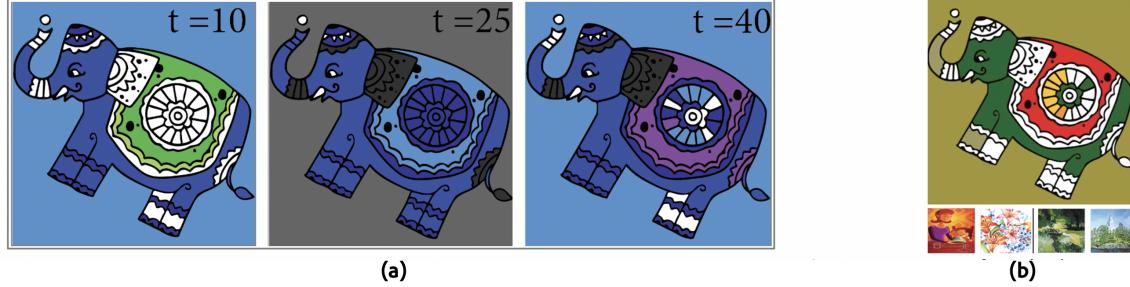
384
385 Of these actions, cross-recoloring actions (for user 1) are given by, $\sum_{t=1}^T \sum_{i=1}^N I(s_t = S_i, u_t = U_1, f_u(S_i) = U_2)$, where
386 $f_u(S_i)$ denotes the user that previously colored the segment S_i . We also compute the number of reversion actions
387 wherein users take an undo turn, that is, revert the current action by returning to the previous state. Table ?? shows the
388 percentage of all these actions. We see that the recoloring actions are lower in the mode with GT images, whereas they
389 are high with no images. This may be due to the provision of expected outputs through GT images, which helps in
390 giving a definitive goal for users to pursue at the cost of limiting their novelty and imagination.

391
392
393 **F2: Varying pace of actions leads to multiple conflicts.** Some users color extremely fast and dominate, giving
394 no opportunity to the other user to think or act. This forces the preferences of one user over the other and leads to
395 skewed turn-taking and inaction from dominated users. To measure this effect, we define a metric called *dominance*
396 score. Formally, dominance of user U_1 over user U_2 is given by $\sum_{t=1}^T \exp(-\lambda t) (Count_1(u_{1:t}) - Count_2(u_{1:t})) / Z$ where
397 $Count_k(u_{1:t}) = \sum_{i=1}^t I(u_i = U_k)$ represents the number of actions taken by user U_k so far and Z_d is a normalization
398 factor. We sum the excess number of actions taken by one user over the other after each action, to capture the cumulative
399 impact of dominance. Note that dominating actions at the beginning of the task are more impactful than those that
400 occur towards the end because early actions determine the overall mood and direction of colorization. We thus multiply
401 an additional $\exp(-\lambda t)$ factor with $\lambda = 0.05$ to capture the decreasing impact of dominating actions. Normalization
402 factor Z_d is the dominance score when only one user takes all the actions (with $T = 100$), given by is 3.84. Table ??
403 displays the average value of dominance scores over all sequences in each mode. A higher dominance score indicates
404 greater number of conflicts.

405
406 While our measure is not perfect, we found that it correlates well with users' feelings based on our informal questions.
407 Figure 5(a) shows a qualitative sequence of high dominance score leading to conflicts and overall poor output. Until
408 action turn $t = 10$, user U_1 forces his preference of blue hue while user U_2 barely imparts her greenish hue to the canvas.
409 This leads to a creative conflict and U_2 cannot visualize how to bring about the greenish hue. By action turn $t = 25$, U_1
410 recolors most of the area colored by U_2 , while U_2 is yet to find a possible middle ground. Towards action turn $t = 40$,

411
412
413
414
415 ⁴Indicator function $I(C)$ is equal to 1 if the condition C is true, else it is equal to 0

417 U_2 decides to help U_1 by altering her preferences and takes up colors with dark blueish/purple hue, recoloring and
 418 corrupting some areas that U_1 has already colored. The final output lacks coherence and is misaligned to both their
 419 preferences. Moreover, both users have an unpleasant experience collaborating, leading to a lose-lose collaboration [95].
 420



422 Fig. 5. Part (a) demonstrates a sequence showing the conflicts that arise due to varying pace of actions. One user typically takes
 423 more moves and dominates the other in terms of overall direction of the colorization. Part (b) demonstrates the line-art of an elephant
 424 colored by users in the mode with inspiration images (shown at the bottom). As evident, the quality is poor with incoherent colors for
 425 the feet and body of elephant.

426 **F3: Novice users need help in aligning preferences.** Conflicts arise when rules of symmetry are broken, such as,
 427 legs of an animal having different colors. Some users found it difficult to match each other's preferences while not
 428 violating spatial symmetry. From Figure 5(b), we can see that the body and legs of elephant are colored differently,
 429 leading to an incoherent output. We define a metric called *common-sense score* to capture this scenario, given by the
 430 average number of color matches for any two segments that have the same ground-truth color. Formally, common-sense
 431 score is

$$\sum_{t=1}^T \sum_{i=1}^{|S|} \sum_{j=1, j \neq i}^{|S|} I(C^{GT}(S_i) = C^{GT}(S_j), C_t(S_i) = C_t(S_j)) / T \quad (1)$$

432 where $C^{GT}(\cdot)$ gives the ground-truth color of a segment and $C_t(\cdot)$ gives the color of a segment at turn t . Note that our
 433 score is an approximate measure as it is impossible to take all rules of symmetry into consideration. For instance, our
 434 score penalizes the case when shirt and trousers of a boy have the same colors in ground-truth image but users in the
 435 study have colored it differently, which does not break any rule of symmetry. In our interviews, users felt that there
 436 were too many colors to choose from and it was very unclear to arrive at harmonious color combinations. We use the
 437 notion of harmony model $H(c_1, c_2)$ [15, 83, 94] to quantify this scenario, which computes the harmony between two
 438 colors by quantifying the hue, luminance, and chroma effects. Formally, *harmony score* is given by

$$\sum_{t=1}^T \sum_{i=1}^{|S|} \sum_{j=1, j \neq i}^{|S|} \frac{|(x_i, y_i) - (x_j, y_j)|_2}{Z_h} H(C_t(S_i), C_t(S_j)) / T \quad (2)$$

439 where the harmony model's outputs are scaled with the L_2 distance between centroids (x_i, y_i) of each segment i to
 440 capture the fact that the coloring is harmonious over a wide area of the line-art. Normalization factor Z_h is given
 441 by $\sum_{i=1}^{|S|} \sum_{j=1, j \neq i}^{|S|} |(x_i, y_i) - (x_j, y_j)|_2$. A low score on either of the above defined metrics indicates a high chance of
 442 conflicting experience.

443 Qualitatively, users found the mode with no inspiration images to be very tricky to deal with, since they were not sure
 444 if the overall canvas was heading in the right direction. The mode with GT images was too obtrusive and provided no
 445

means to think in a novel fashion. The mode with inspiration images was preferred greatly since it provided a reference to compare the canvas with, while also allowing users to be creative and choose novel colors. Overall, many users lacked clarity in dividing the canvas into two meaningful parts, delegate their preferences, and collaborate seamlessly. They were open to assistive systems that can ease their collaborative experience to arrive at a win-win situation [95], while being non-obtrusive. Our qualitative findings coincide with the observation in Main and Grierson [71] that creatives are open to AI-based support tools provided they do not minimize the role of the creative.

3.4 Design Goals for Interventions

Based on our formative studies, analysis, and literature review, we identify the following design goals to build interventions that aid in creative collaboration of users during our colorization task:

- **G1** *Facilitate communication of intentions and aid in resolution of conflicts.* Users generally want to align preferences and help each other out. They prefer taking fewer steps to color, avoid cross-recoloring [F1] and domination [F2], etc. that also aid in building trust. Thus, interventions that can imagine the canvas in the context of user preferences and relay intentions from one user to the other can help in converging their perspectives.
- **G2** *Provide feedback on collaborative strength and overall quality of colored line-art.* Users prefer to track their performance on the task which can help in ensuring that they are not too dominant at any stage of coloring [F2].
- **G3** *Provide assistive tools to color harmoniously.* Novice users feel overwhelmed by the multiple segments and color choices provided to them leading to spatial (symmetry) conflicts [F3]. Thus, interventions that can sensibly reduce the available choices while staying true to user preferences add significant value.
- **G4** *Interventions should be non-obtrusive.* A common prerequisite of all creative assistance tools [71] is that they should not interfere in the task and hinder the imagination of users. Previous research [86, 95] has shown that initial conflicts can lead to a more creative final product.

Simple hard-coded algorithms based on user action counts or segment color counts may help with tracking the state of the canvas, but they do not understand the semantics of the canvas or how user preferences interact with the canvas. Moreover, hard-coded algorithms do not take the sequential action history into account beyond simple count statistics. We thus approach the above goals by training a deep supervised reinforcement learning (RL) based algorithm that acts as a facilitator between the two users. We choose deep RL technique because it suits our requirement of sequentially processing the entire history of user actions and canvas states and outputting an optimal intervention that can meet the above goals. We also prefer the supervised version of RL owing to the success of supervised methods and architectures over the past decade.

4 SIMULATING USER INTERACTION TRAJECTORIES TO TRAIN THE RL SYSTEM

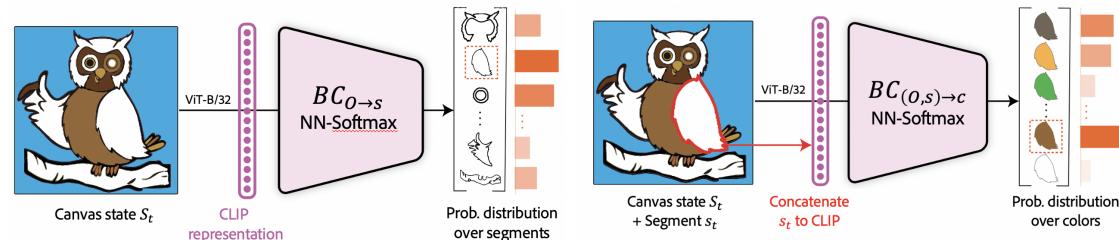
The formal setting of our problem is a multi-agent Markov Decision Process (MDP) consisting of n players that process the state (observation) O and take actions $A^{(i)}$ based on their respective policies $\pi^{(i)}$ for $i = 1 \dots n$. Given policies for all agents, a trajectory/episode τ can be sampled as follows: start with a state O_0 , repeatedly sample actions $\{a_t^{(i)}\}$ from $\{\pi^{(i)}\}$ and update the state to O_{t+1} . The shared objective is to maximise the reward which is given by the average value of $\sum_t \gamma^t R(s_t, a_t^{(1)}, \dots, a_t^{(n)})$ over all trajectories where $R(\cdot)$ is the reward function and γ is the discount factor [56]. Once we embed a human policy π^H as one of the agents in the multi-agent MDP setup, our problem reduces to a partially observed MDP (POMDP) and the optimal agent policy can be learned using an RL algorithm. Our case of line-art

521 colorization consisting of two users and an AI assistive agent to resolve conflicts boils down to a human-human-AI
 522 3-agent MDP setup, in which we embed human policies to train the AI agent.
 523

524 Having formulated a POMDP, we simulate a dataset consisting of a sequence of user interactions and approximate
 525 rewards at each action turn without self-play, as discussed in Section 2. We offset the limitation of lack of interaction
 526 with real world environment by utilizing powerful supervised algorithms for training. We use data from our formative
 527 study to impart domain knowledge and inductive biases into rewards.
 528

529 4.1 Behavior Cloning to learn Human Models

530 Building on the work of Carroll et al. [18], we use the technique of behavior cloning (BC) to train human models
 531 that can be embedded into environment to arrive at a POMDP. BC learns a human policy π^H by directly mapping
 532 human-human observations to actions using supervised learning [12]. From the $N = 129$ joint trajectories of triplets
 533 $\{(u_t, s_t, c_t)\}_{t=1}^T$ collected in mode 3 (with inspiration images) of our formative study, we prepare two single agent
 534 trajectories $\{(u_1, s^1_t, c^1_t)\}_{t=1}^T$ and $\{(u_1, s^2_t, c^2_t)\}_{t=1}^T$ where $(s^i_t, c^i_t) = (s_t, c_t)$ if user u_i has taken the action else
 535 recorded as a noop (no action). We thereby obtain $N' = 258$ trajectories of length $T = 100$. We now prepare the partially
 536 colored line-art states $\{O_t\}_{t=1}^T$ by applying colors to the segments specified by the actions. Note that states observed by
 537 both users at each turn are same. We convert each state to a vector representation using the CLIP Image (ViT-B/32)
 538 model [28, 89] and use it as input to two BC models – User Segment Selector $BC_{O \rightarrow s}$, that predicts the segment a user
 539 would select based on the line-art state; User Color Selector $BC_{(O,s) \rightarrow c}$, that predicts the color a user would select
 540 based on the state and given segment. Figure 6 depicts the input and output of both the models which together form
 541 the human policy π^H . Note that input to $BC_{(O,s) \rightarrow c}$ is a concatenation of the CLIP representation of canvas state and
 542 one-hot representation⁵ of segment.
 543



557 Fig. 6. Behavior Cloning models used: (left) User Segment Selector $BC_{O \rightarrow s}$, (right) User Color Selector $BC_{(O,s) \rightarrow c}$, which form the
 558 human policy π^H .
 559

560 Training and validation sets obtained using an 80 : 20 split on 25800 ($N' \times T$) observations to study BC models.
 561 Like all traditional multi-class classification tasks, we use a feedforward fully connected neural network (NN) with a
 562 softmax⁶ layer at output for both models and select the output with maximum probability. The models do not achieve
 563 high accuracy owing to a small non-diverse dataset. Also note that they do not take any context such as history of user
 564 actions or inspiration images into consideration.
 565

566 ⁵An array of bits where only the corresponding position of the segment is set to 1, and all other bits are set to 0. Size of the array is equal to the number
 567 of segments in the line-art.

568 ⁶Softmax layer takes a vector of K real numbers $[z_1, z_2, \dots, z_K]$ as input, and normalizes it to a probability distribution consisting of K probabilities,
 569 where the k^{th} probability is given by $\frac{\exp(z_k)}{\sum_{i=1}^K \exp(z_i)}$

573 4.2 Improving Human Models with Theory of Mind

574 To improve the quality of human policy π^H , we use the concept of Theory of Mind (ToM) [22, 37] that allows π^H to
 575 reason about other agents' decisions and mental states. ToM proposes a framework to arrive at the intentions of other
 576 players based on state observations and is widely used in human-robot interaction literature [22] to imbue models
 577 with biases that humans display. At every turn, a ToM agent first chooses a high-level subtask to pursue, known
 578 as the *strategic choice*. Once the strategic planning is complete, the agent next chooses a low-level action to tackle
 579 the subtask, known as the *motion choice*. While Knott et. al. [56] build a parameterized ToM agent by learning from
 580 human-human data and anecdotal evidence, we rely on our formative study and analysis to imbue biases and directly
 581 set the parameters. We consider biases such as tendency to recolor [F1], tendency to dominate [F2], and tendency to
 582 accept model suggestions [F3] in our formulation.
 583

584 Note that [56] works with a simpler Overcooked environment where strategic choices and motion choices are well
 585 defined. In our case, creative choices are difficult to arrive at – users may have different creative perspectives of the
 586 line-art which cannot be codified. We consider a simple definition in which segment picking and color picking are
 587 deemed as strategic and motion choices respectively. In the following paragraphs, we describe how π^H powered with
 588 ToM can be embedded into the environment and simulate valid user action sequences.
 589

590 **Prepare the Canvas.** With the goal of simulating an action sequence with conflicts that closely matches two users
 591 coloring together, we first prepare the canvas by sampling a line-art with one of its ground-truth (GT) image from our
 592 data in Subsection 3.1. Since the color palette in our UI consists of 56 fixed colors (see Figure 4), we represent the colors
 593 in the GT image as a 56-dimensional binary vector, called GT palette P^{GT} . A value of 1 at the i^{th} index in P^{GT} indicates
 594 that the GT image contains the color i . Note that we use GT image and its colors as a form of “peeking” into the output
 595 akin to [43, 124], which denotes the users’ common-sense ability to imagine the final output of an uncolored line-art.
 596

597 Recent studies have shown that CLIP model [89] contains abstract multimodal neurons that accurately represent
 598 facets such as line-arts and sketches [36]. We thus use the CLIP model to represent our canvas line-art state, akin to
 599 the vector representation in Subsection 4.1. As the line-art gets subsequently colored in the simulation, we update the
 600 canvas state and its vector representation accordingly to arrive at a sequence of canvas states $\{O_t\}_{t=1}^T$. Figure 7 shows
 601 a schematic diagram of the simulator pipeline, in which the line-art of an owl is chosen.
 602

603 **Prepare the users and their preferences.** We simulate two users and their respective preferences (inspiration
 604 images) to replicate a moodboard-like scenario. We restrict our preferences to four emotions (happy, gloomy, scared,
 605 peaceful) of the BAM dataset [114] and leave further personas for future work. We sample an emotion for each user
 606 and obtain the most common colors for that emotion from the BAM dataset, denoted by a 56-dimensional binary vector
 607 P_i^E ($i = 1, 2$) [102]. We then pick a small fraction of colors from GT image to add to the preferences of both the users.
 608 These overlapping colors, denoted by P^{ol} , enable a common ground scenario when P_i^E are completely disjoint. The
 609 final representation of a user (U_i) is a 62-dimensional vector consisting of a 2-dimensional one-hot vector $v_i^{(u)}$ for the
 610 user ID, 4-dimensional one-hot vector $v_i^{(e)}$ for the emotion, and a 56-dimensional binary vector (P_i^U) conveying the
 611 color preferences.
 612

$$P_i^U = (P^{GT} \cap P_i^E) \cup P^{ol} \quad (3)$$

$$U_i = [P_i^U, v_i^{(e)}, v_i^{(u)}]; i = 1, 2 \quad (4)$$

613 We set one of the two users randomly to be more dominant to incorporate our finding [F2] from the formative study and
 614 assign a dominance probability p_{dom} . The value $p_{dom} = 0.7$ is computed from sequences of mode 3 (with inspiration
 615

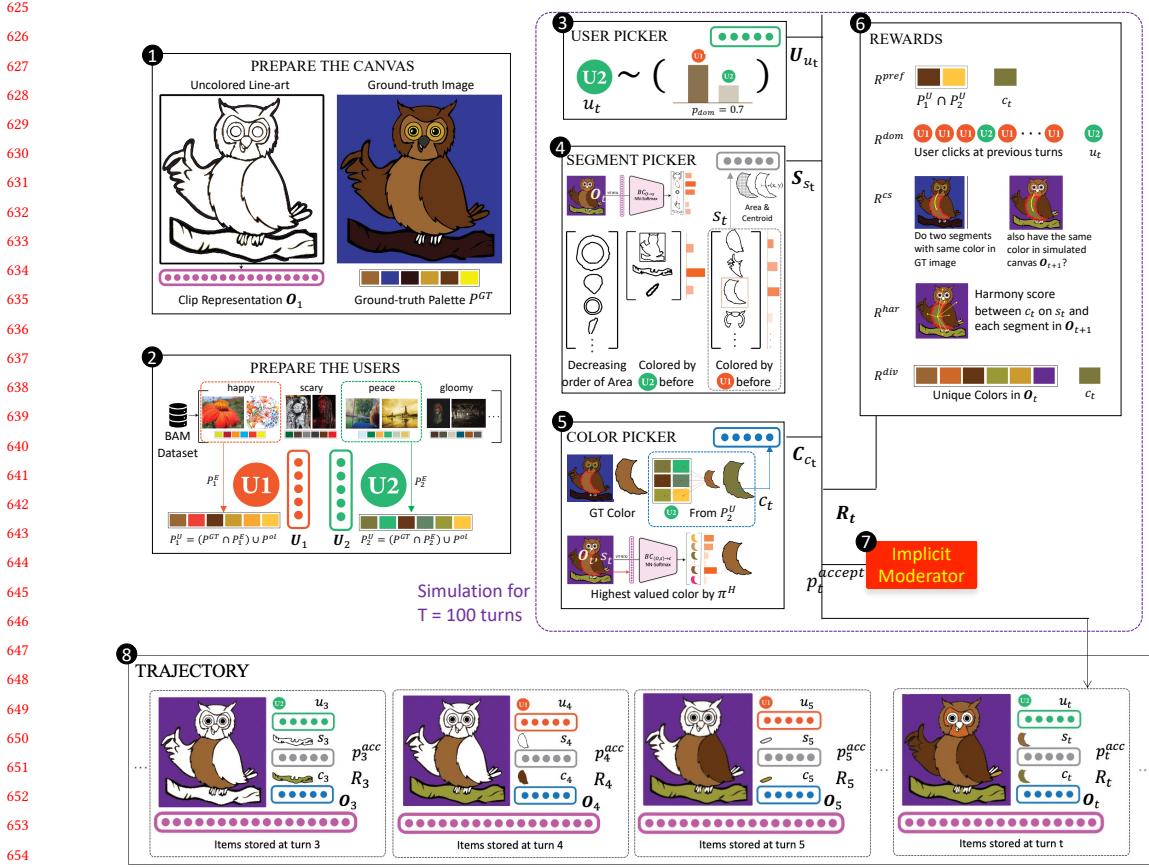


Fig. 7. Schematic diagram of our novel simulator that generates trajectories for our co-creation task. We direct readers to Appendix E for enlarged images of each step. Best viewed in color.

images) as the average ratio of the number of actions the dominating user takes to the total number of actions. In Figure 7, two users U1 and U2 are chosen with happy and peaceful preferences respectively, with user U1 being more dominant. We now describe the three picker modules that govern the simulator’s turn-taking process.

User Picker. The first module picks a user at action turn t , denoted by u_t . We sample $u_t (\in 1 \dots |S|)$ from a Bernoulli distribution given by, $u_t \sim \text{Bern}(p_{\text{dom}})$ to incorporate [F2], that is, the user’s tendency to dominate. In Figure 7, the user picker module picks U2.

Strategic Choice: Segment Picker. The second module picks a segment $s_t (\in 1 \dots |S|)$ to be colored by user u_t by following one of the three mechanisms described in Appendix E. The representation of a segment i (S_i) is a $(|S| + 3)$ -dimensional vector consisting of the concatenation of one-hot vector of the segment ($v_i^{(s)}$), area of the segment, and centroids of the segment [2], where $|S|$ is the number of segments in the chosen line-art. In Figure 7, the segment picker module chooses the third mechanism of cross-recoloring (segments colored by U1 before) and uses the human policy $\pi^H(BC_{O \rightarrow s})$ to pick the owl body segment.

Motion Choice: Color Picker. The third module picks a color c_t ($\in 1 \dots 56$) given the user u_t and segment s_t by following one of the three mechanisms described in Appendix E with equal probability. As explained earlier, each color i is represented by a 56-dimensional one-hot vector C_i . In Figure 7, the color picker module chooses the second mechanism of picking the closest color to GT color from U2's palette and recolors the body segment with an olive color. Having obtained the action triplet (u_t, s_t, c_t) , we update the line-art by coloring s_t with c_t and recompute the CLIP representation of canvas state as O_{t+1} . The turn-taking process terminates when $t = 100$.

Rewards. Our design goals [G1, G3] are to facilitate easy resolution of conflicts while obtaining a coherent colorization that respects both users' preferences. We incorporate metrics from the formative study to prepare reward signals that approximate our goals. The RL agent we subsequently train receives information about our goals via these reward signals and learns when to intervene while observing the environment. The following rewards are computed at every turn and combined using a weighted sum to form the total reward signal R_t :

- R^{pref} [G1]: *Promotes alignment of preferences.* It is set to 1 when the picked color c_t lies in the intersection of both users' color palettes, that is, $P_1^U \cap P_2^U$; else it is set to 0.
- R^{dom} [G1]: *Promotes equal turn-taking and discourages dominance of one user over the other.* It is similar to the negated value of dominance score [F2] defined in Subsection 3.3, but computed only till turn t .
- R^{cs} [G3]: *Encourages sensible colorization and falling back to ground-truth colors when unsure.* It is similar to common-sense score [F3] defined in Subsection 3.3 and is given by

$$R_t^{cs} = \sum_{j=1}^{|S|} I\left(C^{GT}(S_j) = C^{GT}(S_{s_t}), C_t(S_j) = c_t\right) / |S| \quad (5)$$

Intuitively, we reward the case when two segments that have the same color in ground-truth coloring also have the same color in simulated sequence.

- R^{har} [G3]: *Promotes harmony between color preferences of each user.* It is similar to harmony score [F3] defined in Subsection 3.3 and is given by

$$R_t^{har} = \sum_{j=1, j \neq s_t}^{|S|} \frac{\|(x_{s_t}, y_{s_t}) - (x_j, y_j)\|_2}{Z_h} H(c_t, C_t(S_j)) \quad (6)$$

- R^{div} : *Promotes diversity and discourages the colorization of line-art with a single color.* It is given by the ratio of the number of unique colors in the partially colored line-art to the total number of available colors (= 56) at turn t .

Additionally, we also sample a value $p_t^{accept} \sim Unif(0, 1)$ that denotes both users' tendencies to consider suggestions provided by our moderator agent that we will eventually train. We model this bias aspect because not all users want to consider an AI model's assistance [F3]. Note that we do not explicitly simulate an AI agent within our action sequences. We consider these sequences with an assumption that the users are already aware of the AI agent's suggestions. When $p_t^{accept} > 0.5$, it denotes that the users have taken actions by incorporating the suggestions of the agent. When $p_t^{accept} < 0.5$, it denotes that the users have ignored our agent's help. The value p_t^{accept} plays a crucial role when training the agent as it enables the loss function (defined in later sections) to penalize turns wherein the users incorporated the suggestions and yet recorded a low reward.

To summarize, we obtain the following list of items $[O_t, U_{u_t}, S_{s_t}, C_{c_t}, R_t, p_t^{accept}]$ for $t = 1 \dots 100$ as the simulated sequence for chosen line-art. We compute several such sequences for each line-art and use them as RL training data. We compute metrics from Table ?? and quantitatively verify that our sequences are close to mode 3 (with inspiration images).

729 5 SYSTEM ARCHITECTURE

730 Uncertainty in creative task divisions occur due to a lack of communication of intentions between users [118], and also
 731 due to the inherent multimodality of the coloring task [20]. Given our highly uncertain scenario with many correct
 732 plans of action, we delegate the final segment and color choices to users' creative agency and only act as a "moderator"
 733 or "facilitator" between them to ease communication and make aligned choices. We thus envision CollabColor to be a
 734 contextual memorization engine that recognizes and gives higher importance to conflict-resolving actions and conveys
 735 those actions to users via interventions on the user interface.

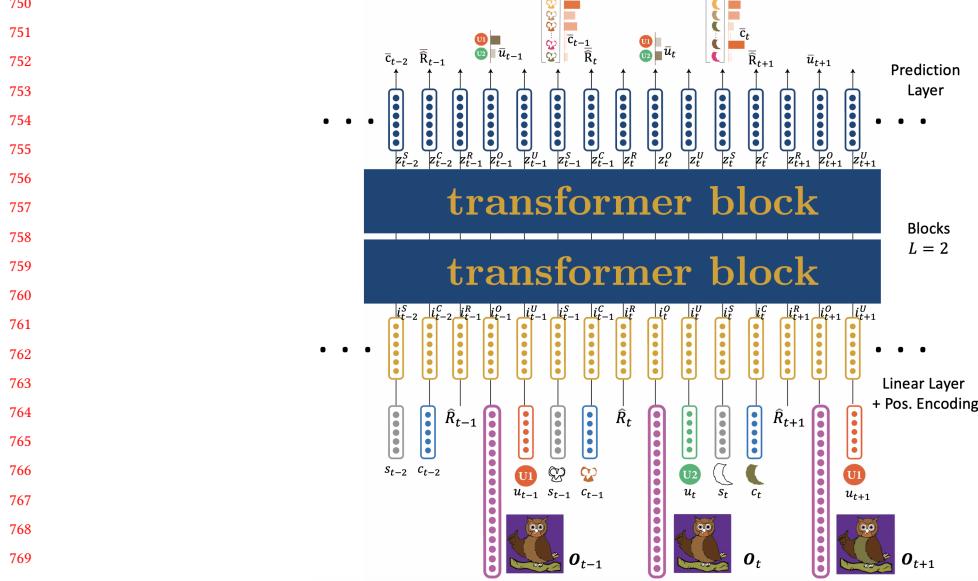
736

737 5.1 System Configuration at Train Time

738 For the CollabColor system to act as a good facilitator between two users, it is vital that the model internally develops a
 739 rich representation of all previous states and actions that the users have taken and distil key information about their
 740 preferences. The agent also has to plan ahead and imagine a future in which both their inspirations are incorporated to
 741 arrive at a collaborative win-win situation [95], while maintaining a coherent colorization. To achieve these properties,
 742 we consider a novel transformer decoder architecture building on [21] that autoregressively models the simulated
 743 trajectories. Figure 8 shows the architecture of our system at train time.

744

745



750
 751 Fig. 8. Transformer decoder (two blocks) architecture of our novel CollabColor moderator system. At train time, learning on offline
 752 simulated trajectories is modelled as a sequential learning problem, akin to language modelling in NLP. Masked self-attention in
 753 transformer block ensures that the inputs at any turn attend only to previous inputs. We employ teacher forcing to train the model
 754 using losses accumulated on predictions. Details of the transformer block are deferred to original papers [90].
 755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773 Let us consider a simulated trajectory $\tau = (\hat{R}_1, O_1, U_{u_1}, S_{s_1}, C_{c_1}, \hat{R}_2, O_2, U_{u_2}, S_{s_2}, C_{c_2}, \dots, \hat{R}_T, O_T, U_{u_T}, S_{s_T}, C_{c_T})$
 774 where $\hat{R}_t = \sum_{t'=t}^T R_{t'}$ is the returns-to-go value indicating the future desired returns. For every turn t , we use linear

layers and add a positional encoding p_t to compute the representations of inputs to the transformer block:

$$\begin{bmatrix} i_t^R ; i_t^O ; i_t^U ; i_t^S ; i_t^C \end{bmatrix} = [W^R \hat{R}_t + p_t ; W^O \mathbf{O}_t + p_t ; W^U \mathbf{U}_{u_t} + p_t ; W^S \mathbf{S}_{s_t} + p_t ; W^C \mathbf{C}_{c_t} + p_t] \quad (7)$$

where $p_t \in \mathbb{R}^{D \times 1}$, $W^R \in \mathbb{R}^{D \times 1}$, $W^O \in \mathbb{R}^{D \times 512}$, $W^U \in \mathbb{R}^{D \times 62}$, $W^S \in \mathbb{R}^{D \times (|S|_{max}+3)}$, and $W^C \in \mathbb{R}^{D \times 56}$ are learnable matrices. Note that we pad segment vectors to arrive at a common length of ($|S|_{max} + 3$) for batched computations where $|S|_{max}$ is the maximum number of segments present in any line-art of the given batch. We now prepare a sequence of context length K consisting of $5K$ inputs and use masked causal self-attention in the transformer block mirroring GPT [90] to obtain output embedding for each input:

$$h^0 = \left[i_{t-K}^R ; i_{t-K}^O ; i_{t-K}^U ; i_{t-K}^S ; i_{t-K}^C ; i_{t-K+1}^R ; i_{t-K+1}^O ; i_{t-K+1}^U ; i_{t-K+1}^S ; i_{t-K+1}^C ; \dots ; i_t^R ; i_t^O ; i_t^U ; i_t^S ; i_t^C \right] \quad (8)$$

$$h^l = \text{transformer_block}(h^{l-1}) \quad \forall l \in [1, L] \quad (9)$$

$$\Rightarrow h^L = \left[z_{t-K}^R ; z_{t-K}^O ; z_{t-K}^U ; z_{t-K}^S ; z_{t-K}^C ; z_{t-K+1}^R ; z_{t-K+1}^O ; z_{t-K+1}^U ; z_{t-K+1}^S ; z_{t-K+1}^C ; \dots ; z_t^R ; z_t^O ; z_t^U ; z_t^S ; z_t^C \right] \quad (10)$$

where L is the number of transformer blocks. We use masking to ensure that the output embeddings at any turn do not attend to future inputs. For instance, z_j^U at position j attends to the inputs at previous positions $[i_{t-K}^R ; i_{t-K}^O ; \dots ; i_{t-j+1}^S ; i_{t-j+1}^C ; i_j^R ; i_j^O]$ only. The self-attention module implicitly forms associations between state-reward, state-user, segment-color and other pairs over the entire context length. Our model thus learns long-range dependencies and meaningful patterns of color choices that achieve high reward. We train the model with teacher forcing [115] to predict the current user, current color, and future returns-to-go as follows:

$$\bar{u}_t \sim P(u_t | \hat{R}_t, \mathbf{O}_t, \tau_{<t}) = \text{softmax}(W_{pred}^U z_t^O) \quad W_{pred}^U \in \mathbb{R}^{2 \times D} \quad (11)$$

$$\bar{c}_t \sim P(c_t | \hat{R}_t, \mathbf{O}_t, \mathbf{U}_{u_t}, \mathbf{S}_{s_t}, \tau_{<t}) = \text{softmax}(W_{pred}^C z_t^S) \quad W_{pred}^C \in \mathbb{R}^{56 \times D} \quad (12)$$

$$\hat{\tilde{R}}_{t+1} = W_{pred}^R z_t^C + b_{pred}^R \quad W_{pred}^R \in \mathbb{R}^{1 \times D}, b_{pred}^R \in \mathbb{R} \quad (13)$$

where we use $\tau_{<t}$ as a shorthand for previous 5 ($K - 1$) tokens in the considered trajectory. The current user \bar{u}_t is predicted using the output embedding of current canvas state z_t^O and binary cross-entropy loss is used to train the model. Similarly, the current color \bar{c}_t is predicted using the output embedding of current segment z_t^S and multiclass cross-entropy loss is used to train the model. Finally, the future returns-to-go value $\hat{\tilde{R}}_{t+1}$ is predicted using a linear layer on the output embedding of the current color z_t^C and mean-squared loss is used to train the model. We do not predict the current segment because of its high-dimensionality and obtrusiveness [G4]. Moreover, we do not consider shape features of segments beyond its centroid and area leading to difficulties in generalization across all line-art segments. Mathematically, the loss L_t is given by,

$$L_t = p_t^{accept} \times \left(\text{BCELoss}(\bar{u}_t, u_t) + \text{CELoss}(\bar{c}_t, c_t) + \left\| \hat{\tilde{R}}_{t+1} - \hat{R}_{t+1} \right\|_2^2 \right) \quad (14)$$

where we multiply p_t^{accept} as an importance weight of the current turn t . We average L_t over the context length and compute gradient updates to learn parameters of our model that minimize the overall loss.

833 5.2 System Configuration at Inference Time: Updating UI with Interventions

834 As a result of above training, we can specify the desired performance via returns-to-go value and provide an uncolored
 835 line-art canvas state to initiate autoregressive generation. However, we cannot directly embed this model into our
 836 interface and color the segments following the generated sequence because it denies users the agency to express
 837 their creativity. To meet our design goal [G4] of non-obtrusiveness and ensure user agency is intact, we present the
 838 predictions of our model via five UI elements that help users in making optimal choices. For the following interventions,
 839 we assume that $(t - 1)$ turns are completed at inference time and the recorded trajectory is given by $\tau_{<t}$.
 840

841 **Peek Button** [G1, G3]. When clicked by U1, a snapshot image of the current canvas state is displayed with the
 842 uncolored segments filled by colors that U2 might prefer. Formally, for each uncolored segment i , we pick the color \bar{c}_t
 843 such that $\bar{c}_t = \text{argmax}_c P(c_t = c | \hat{R}_t, O_t, \mathbf{U}_2, \mathbf{S}_i, \tau_{<t})$ by passing \mathbf{U}_2 and S_i at turn t to our system and fill the segment
 844 i with \bar{c}_t . This intervention can be used by U1 to imagine the final output that U2 intends to have, thereby helping in
 845 forming a converged perspective.
 846

847 **Color Prediction Button** [G3]. When clicked by U1, a snapshot image of the canvas is displayed with the uncolored
 848 segments filled by colors that U1 might prefer. The functionality is similar to the previous button, expect here we pick
 849 the color \bar{c}_t such that $\bar{c}_t = \text{argmax}_c P(c_t = c | \hat{R}_t, O_t, \mathbf{U}_1, \mathbf{S}_i, \tau_{<t})$ by passing \mathbf{U}_1 and S_i for each uncolored segment i .
 850 This intervention provides great value to novice users who face difficulty in picking optimal preference-aligned colors
 851 to fill the canvas. Note that the snapshot image displayed here is *equivalent* to the image displayed when U2 clicks on
 852 the peek button at her end.
 853

854 **Floating Emoticons** [G1, G2]: A happy-face emoticon is displayed when the user who colors at turn t and the
 855 optimal user predicted to color by the model $\bar{u}_t = \text{argmax}_u P(u_t = u | \hat{R}_t, O_t, \tau_{<t})$ are the same. Contrarily, an angry-face
 856 emoticon is displayed when the user who colors is not the same as model's optimal user. This intervention facilitates
 857 a weak communication link between the users and nudges them to take actions that are in favour of each other's
 858 preferences. It also acts as a feedback mechanism to the other users; a user seeing too many angry emoticons is probably
 859 forcing her palette onto the other user. The emoticons are ephemeral and float for a few seconds before disappearing.
 860

861 **Collaboration Score Button** [G2]: When clicked by any user, a progress bar is displayed that indicates the collabora-
 862 tive strength of both users in achieving a coherent line-art colorization. It is computed using the formula $(4T - \tilde{R}_t)$
 863 where $\tilde{R}_t = W_{\text{pred}}^R z_{t-1}^C + b_{\text{pred}}^R$ is the predicted returns-to-go value at the previous turn $(t - 1)$. Note that $4T$ is the
 864 maximum attainable total reward value in our simulator at the end of any trajectory⁷. This intervention provides
 865 valuable feedback to both users; a low score implies that \tilde{R}_t is high and more rewards can be gained by taking optimal
 866 actions in future turns.
 867

868 Figure 1 shows these additions to the user interface. Users can click on “Toggle Model” button to display the various
 869 intervention elements listed above. . Note that the snapshot image of Color Prediction when chosen by user S1 is the
 870 same as the snapshot of Peek (2) when chosen by user S2; this connection is annotated with a dashed line. From the Peek
 871 output (2, 3) on both screens, we can see that our model is able to color well and capture common-sense reasoning such
 872 as legs having the same green color, tail taking a darker green shade, etc. The model also captures the user inspirations
 873 well – the color for elephant's body in S2's Peek output (2) (that is, S1's color predictions) is a very light shade of green
 874 that is the closest to S1's preferred colors. Based on interviews in the formative study, we add some more UI elements
 875 such as inspiration/emotion/joint palettes (5) to make the task easy to handle.
 876

877 ⁷Reward R_t is a weighted sum (where all weights are non-negative and less than or equal to 1) of R^{pref} , R^{dom} , R^{cs} , R^{har} , and R^{div} . Each reward
 878 component lies in the range $[0, 1]$ except R^{dom} which is in the range $[-1, 0]$. Thus, the maximum value attained by R_t at any turn is 4, and the total
 879 reward at the end of a trajectory is $T \times 4$

885 6 EXPERIMENTAL RESULTS

886 Despite the success of Deep RL systems in achieving high accuracies and rewards on various tasks, there are several
 887 works that show that these systems fail in deployment due to diversity in real-world scenarios [35, 68]. This shows
 888 that average reward or accuracy are not sufficient metrics to gauge the strength of RL systems. Since we situate our
 889 CollabColor system in human-human interactions, it accentuates the difficulty in evaluation due to subjectivity in
 890 human actions [60, 107]. In particular, Leibo et al. [60] posit that concepts such as trust, generosity, and deception are
 891 vital to evaluate AI for collaboration. User studies are also not often representative of the performance at deployment
 892 time due to a long tail of unusual and adversarial human actions [56]. Given that our task is creative, correct behaviours
 893 are *ambiguous* and arriving at reasonable behaviour is tough even for humans. Nevertheless, we design a suite of
 894 evaluation strategies to stress test our system based on real-world scenarios. We also derive inspiration from evaluation
 895 techniques used in psychological and cognitive sciences to assess team functioning [79, 116] and adapt them to our
 896 three-agent setup. Our evaluation schemes are also motivated by the increasing trend in ML communities to move
 897 beyond easily calculated metrics and consider diverse real-world scenarios to capture performance [92, 126].
 898

903 6.1 Setup

904 We prepare the canvas and users as shown in Figure 7 with uncolored line-arts and user emotion pairs, and use the
 905 trained CollabColor model to autoregressively generate action trajectories with an initial returns-to-go value $\hat{R}_1 = 4T$,
 906 context length of 10 and T set to 100. Effectively, we replace the rule-based pickers with our model that has learned to
 907 promote and generate conflict-resolving actions. Since our model does not predict segments, we consider segments
 908 following segment picker mechanisms (Appendix E) for simplicity. Additionally, we also train our system with a
 909 BiLSTM architecture [66] instead of transformer architecture as an ablation and generate trajectories as described above.
 910 The generated trajectories are equivalent to a three-player system in which the two users accept *all the suggestions* of
 911 the third player (our facilitator system). We could consider such users as novices in co-creating who are looking for
 912 assistance from our system and denote their acceptance by a quantity $p_i^{acc} = 1.0$ for each user i , to replace p_t^{accept} in
 913 Section 4.2. Note that these trajectories are only a proxy for deployment scenario if which two real users interact with
 914 our system. We show some qualitative examples of our trajectories in Appendix D.

915 **916 Stress Testing.** We consider stress testing by designing adversarial user agents to interact with our system. Note
 917 that an adversary can be very powerful in our task, leading to very low reward easily. For instance, an adversary could
 918 always color the same segment with a fixed color regardless of what our system suggests ($p_i^{acc} = 0$), leading to a very
 919 low reward. We thus consider three adversarial types for User-1 keeping User-2 at full acceptance ($p_2^{acc} = 1.0$): (i)
 920 Novice with lower acceptance $p_1^{acc} = 0.8$ who accepts system assistance for most turns, but picks random colors and
 921 segments for other turns; (ii) Stationary user with a very low acceptance of $p_1^{acc} = 0.2$ who always chooses the same
 922 segment regardless of the turn; (iii) Random user with a very low acceptance of $p_1^{acc} = 0.2$ who chooses colors and
 923 segments at random for most turns. Based on these user combinations, we generate several trajectories with $\hat{R}_1 = 4T$ as
 924 described earlier. We experimented with lower p_1^{acc} values but found the trajectories to be severely degenerate, with
 925 the same action predictions repeating across turns and terminating prematurely.

926 **927 Recurrence Quantification Analysis (RQA).** Recurrence plots [73, 113] have been previously used in measuring
 928 social and collaborative problem solving systems [25, 116]. RQA involves the construction of a 2-dimensional recurrence
 929 plot (RP) that visualizes recurrence of states in a dynamic system. When there are two or more agents interacting,
 930 joint RPs are constructed by taking a hadamard product of RP of each agent. Motivated by the literature on Human
 931

Autonomy Teaming (HAT) [74, 85], we construct a joint RP from the trajectories generated by CollabColor. Formally, let $R_1, R_2 \in \mathbb{R}^{T \times T}$ be the RPs of User-1 and User-2 respectively, then for each $(t_a^i, t_b^i) \in \{1, 2, \dots, T\} \times \{1, 2, \dots, T\}$ and $i \in 1, 2$,

$$R_i = \begin{cases} 1 & \text{if } D(t_a^i, t_b^i) < \epsilon \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

where $D(t_a^i, t_b^i) \in \mathbb{R}$ is a distance function that maps the orientation of User- u at turn t_i to a continuous value. Joint RP is therefore given by $R_1 \odot R_2$. As evident from the equations, the diagonal entries of both RPs and JRP are set to 1. Determinism is a statistic computed from JRP that measures the extent to which teams follow stable, rigid patterns of interactive behaviour. It is calculated as the proportion of recurrent points that lie on the diagonal on JRP. The inverse of determinism, termed flexibility [79], captures how fluid and receptive the team members are to state changes. Intuitively, a high value of flexibility indicates a system wherein the players are not very rigid; they respond and vary their orientations based on the actions of other players.

For RP computation, we extract triplets (u_t, s_t, c_t) from generated trajectories and prepare two single agent trajectories $\{(u_1, s^1_t, c^1_t)\}_{t=1}^T$ and $\{(u_1, s^2_t, c^2_t)\}_{t=1}^T$ where $(s^i_t, c^i_t) = (s_t, c_t)$ if user u_i has taken the action at turn t , else recorded as $(s_t, \text{argmax}_c P(c_t = c | \hat{R}_t, O_t, \mathbf{U}_{-i}, \mathbf{S}_{s_t}, \tau_{<t}))$, that is, with the color u_{-i} would pick (computed by model color predictor described in Section 5.2) to color the segment s_t picked by u_i . Note that the preparation is similar to Section 4.1, except we replace the noop action with color prediction action when the user has not taken any move because this captures the behavioural orientation of that user. Moreover, had we recorded the color as noop, then the RPs of both users would always match due to the recurrence of noops whenever the users do not take actions. We consider a simple distance function $D(\cdot)$ that takes in the orientations $(s^i_{t_a}, c^i_{t_a})$ and $(s^i_{t_b}, c^i_{t_b})$ of User- i at two turns t_a, t_b and aggregates the L_2 distance of segment centroids of $(s^i_{t_a}, s^i_{t_b})$ with LAB space distance of colors $(c^i_{t_a}, c^i_{t_b})$. We use the JRPs to analyze the variations of rewards with flexibility in our experiments.

6.2 Results

6.2.1 Stress Test Analysis. We average each of the five rewards across the entire trajectory and display values in Table 2 for all user types and model variants. Additionally, we display the average reward values obtained from the simulator. From the table, we can see that CollabColor with GPT architecture achieves higher average reward in most parts, except for diversity. This shows the strength of our GPT model in learning meaningful patterns from the simulated sequences and implicitly memorizing the useful patterns so that the model-generated trajectories achieve reward (signalled by a maximum returns-to-go value of $4T$). BiLSTM model achieves higher rewards than the simulator, but not as much as our model with GPT variant. This verifies the value of masked attention and longer context lengths.

Across the user types, we see that average rewards noticeably decrease as users take more random actions. Interestingly, our model with GPT maintains a decent reward value as compared to the model with BiLSTM even for stationary and random user types. This shows that our system is robust to adversarial agents.

Across the parts of rewards, the maximum gains are obtained in R^{cs} . This shows that our model is able to assimilate common-sense colorizations across the entire train data, showing the strength of transferability of supervised learning. R^{har} is also high for our model; showing that the final colorizations produced are harmonious and pleasing to the eye of the viewer. However, there is a noticeable drop in R^{div} ; this could be due to degenerate color actions generated in our model trajectories, especially in the beginning of the sequence. Our observation matches with that of Barnes et al.'s

Table 2. Quantitative Results of Stress Testing our CollabColor system and its BiLSTM variant with four user types. (\uparrow) indicates that higher values are better. Values shown are averaged over the entire trajectory. Note that for all the cases, User-2 type is fixed to Novice and $p_2^{acc} = 1.0$. R^{tot} is the sum of the five reward values. For reference, we also display the rewards obtained by simulated trajectories used to train CollabColor. Additionally, we also note that the maximum value of average total reward across entire sequence is 4($= 4T/T$)

User-1 Type	CollabColor Variant	R^{pref} (\uparrow)	R^{dom} (\uparrow)	R^{cs} (\uparrow)	R^{har} (\uparrow)	R^{div} (\uparrow)	R^{tot} (\uparrow)
	Simulator	0.47	-0.48	0.16	0.46	0.65	0.98
Novice $p_1^{acc} = 1.0$	With BiLSTM	0.54	-0.37	0.19	0.49	0.44	1.29
	With GPT (Ours)	0.73	-0.14	0.45	0.77	0.62	2.43
Novice $p_1^{acc} = 0.8$	With BiLSTM	0.52	-0.51	0.19	0.42	0.46	1.08
	With GPT (Ours)	0.72	-0.26	0.44	0.74	0.62	2.26
Stationary Segment $p_1^{acc} = 0.2$	With BiLSTM	0.33	-0.60	0.12	0.36	0.32	0.53
	With GPT (Ours)	0.56	-0.51	0.27	0.37	0.48	1.17
Random $p_1^{acc} = 0.2$	With BiLSTM	0.17	-0.64	0.10	0.24	0.69	0.56
	With GPT (Ours)	0.29	-0.52	0.21	0.39	0.66	1.03

Table 3. Results of clustering the trajectories generated by CollabColor for the case of $p_1^{acc} = 1, p_2^{acc} = 1$. We cluster based on the 16 sampled user-user emotion pairs, and display the reward obtained by averaging over all trajectories in each cluster. The (*) symbol next to an emotion indicates that it is more dominant for that cluster. Refer to Section 4.2 for a detailed description of how users and their emotions/dominance are prepared.

User 1 (\downarrow) & User 2 (\rightarrow) Preferences	Happy	Scary	Gloomy	Peaceful
Happy*	3.44	1.92	1.77	2.91
Scary*	1.88	3.02	2.45	1.68
Gloomy*	1.62	2.45	3.21	2.24
Peaceful*	2.95	1.72	2.18	3.42

[13] finding that without user interaction, it is difficult for models to start with sensible colorization or explore the vast color design space.

6.2.2 *Which user emotion pairs are easier to deal with?* To further analyze CollabColor (with GPT), we cluster the generated trajectories into 16 parts based on the sampled user preferences and dominance (refer Section 4.2). Table 3 shows the breakdown of average rewards across the 16 clusters. As evident from the table, our model achieves the highest rewards when sampled user emotions are the same. This is expected as there will be fewer conflicts in the users and colors chosen by our model, leading to a high R^{pref} and R^{har} . Interestingly, emotions that are closer to each other [36] such as Happy-Peaceful pair also score high. This proves that our model is able to successfully distill the preferences from the input sequences, and find aligned colors/emotions. Some emotion pairs such as Gloomy-Happy score low reward, indicating that these lead to more conflicts. However, most of the cluster rewards are higher than the total average reward of simulator, showing that our model does find and promote conflict-resolving actions even for incompatible preference/emotion pairs. We do not find any discernible trend as the dominance varies across emotions. Note that we denote the dominant user/emotion with a (*) in Table 3.

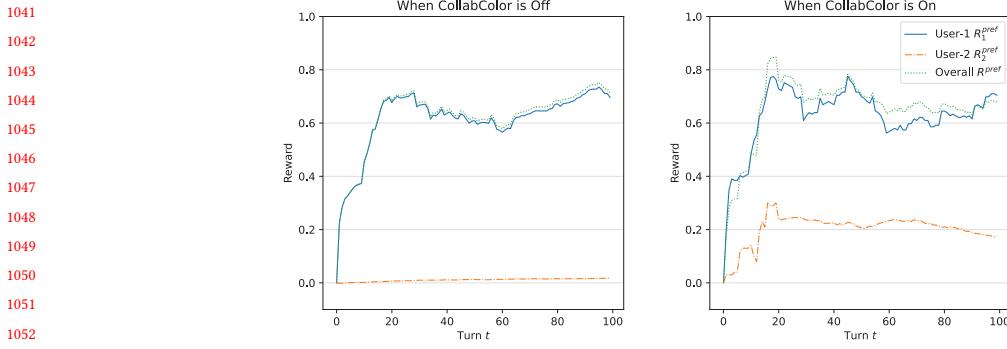


Fig. 9. Stress testing the scenario wherein User-2 does not take any actions. We consider two cases: (i) trajectories generated when CollabColor is off ($p_1^{acc} = 0$); (ii) trajectories generated when CollabColor is on ($p_1^{acc} = 1$).

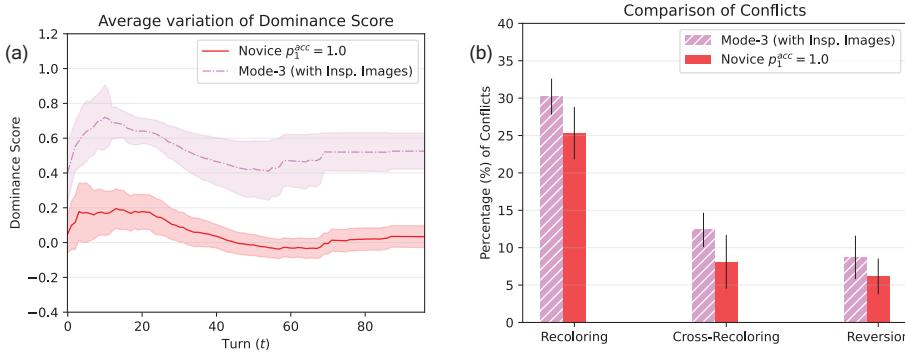


Fig. 10. Part (a) shows the average variation of dominance score with turns for two trajectories: (i) Novice user trajectory generated by CollabColor with $p_1^{acc} = 1, p_2^{acc} = 1$; (ii) trajectories extracted from mode-3 of our formative study. When trajectory lengths are not equal to 100, we pad with the average dominance score of that trajectory to reach $T = 100$. Part (b) compares the coloring conflicts such as recoloring, reversion averaged over the same trajectories in Part (a).

6.2.3 What if network connectivity is poor? We consider yet another form of stress testing in which one user does not take any actions at all. This is representative of a real-world scenario in which a user has poor network connectivity and her actions do not get relayed in the synchronous co-creation process. We first formulate two values R_1^{pref} and R_2^{pref} to represent how close an image is to the sampled emotion of each user. We do so by calculating the similarity of the canvas line-art embedding at turn t given by O_t with the average CLIP embedding of images in BAM dataset corresponding to a particular emotion. We average the similarities to obtain two values k_1 and k_2 representing the alignment to preference emotions. We then obtain R_i^{pref} as $k_i/(k_1 + k_2) \times R^{pref}$. We plot the three values $R_1^{pref}, R_2^{pref}, R^{pref}$ for two cases: when model is off, corresponding to the trajectories wherein user-2 takes no actions and user-1 takes actions $p^{acc} = 0$, and when model is on, corresponding to the trajectories wherein user-2 again takes no actions but user-1 takes actions $p^{acc} = 1$.

When CollabColor is off, R^{pref} is entirely governed by user-1's preferences, with R_1^{pref} curve almost coinciding with R^{pref} . User-2's preferences are seldom incorporated, leading to complete lack of co-creativity. When CollabColor is on, R_2^{pref} increases, indicating that the actions suggested by the model try to find a balance between the two preferences,

1093 leading user-1 to find a shared perspective. This is possible because CollabColor is trained to arrive at a coherent and
1094 harmonious colorization that incorporates both users' preferences. Thus, even when the second user loses connectivity,
1095 our CollabColor system ensures that her presence is not *entirely* lost in the final line-art colorization. However, we note
1096 the limitation that the preferences are unevenly incorporated.
1097

1098 **6.2.4 Reduction in Coloring Conflicts.** We now compare trajectories generated by CollabColor for Novice user type
1099 with $p_i^{acc} = 1.0$ with the user-user data obtained in mode-3 (with inspiration images) of formative study on the metrics
1100 of dominance scores and coloring action conflicts defined in Section 3.3. This comparison allows us to see how well our
1101 sequences improve upon a real world user study. Note that when our model is deployed, our sequences are not applied
1102 on the line-art, but only suggested to the user if model is toggled (non-obtrusive). Thus, the data we present here with
1103 $p_1^{acc} = 1$ is an upper bound on the expected performance in deployment.
1104

1105 We first plot the variation of dominance score of the dominant user (can be user-1 or user-2) across turns. From
1106 Figure 10(a), we can see that CollabColor trajectories have lesser dominance than the trajectories collected in user study.
1107 This is also reflected in a higher dominance score in Table 2. Furthermore, our trajectories also allow for a very feeble
1108 negative dominance score towards the middle of the entire turn-taking process. This shows that our model also gives
1109 the less-dominant user a chance to take extra turns. Note that we do not explicitly train the model to produce such
1110 a scenario. Our model is able to learn the connections between the R^{dom} signals and the simulated user sequence in
1111 input to stitch useful user sequences together for inference-time generation. Note that the mode-3 curves becoming
1112 constant towards the end of turns could be an artifact of the padding scheme (appending average dominance values)
1113 we employed to make the length of sequences $T = 100$. Next, we plot the coloring action conflicts in Figure 10(b). As
1114 evident from the figure, there are fewer recoloring and reversion conflicts in generated trajectories. This again shows
1115 the strength of our model in learning to reduce conflicts and promoting conflict resolutions with only implicit signals
1116 from the rewards.
1117

1118 **6.2.5 Does Flexibility lead to increase in rewards?** We compute the flexibility using Joint Recurrence Plots (JRP)s for
1119 three cases as shown in Figure 11, and display a scatter plot of flexibilities of trajectories with the total reward obtained.
1120 The three cases we consider are (i) user simulator trajectory; (ii) Novice User trajectory generated by CollabColor with
1121 $p_1^{acc} = 1, p_2^{acc} = 1$; (iii) Stationary User trajectory $p_1^{acc} = 0, p_2^{acc} = 1$ with generated by CollabColor, wherein both the
1122 user chooses the same segment and same color for all turns.
1123

1124 From Figure 11, we see that the simulator has low rewards (also in Table 2) and comparatively low flexibility. Novice
1125 user trajectories have higher reward and high flexibility, showing that the model facilitates and promotes actions that
1126 lead to less rigidity and more cooperation. Despite the overall flexibility of the simulator being low, the fact that the
1127 model has achieved higher flexibility indicates that it is able to distill parts of the simulated trajectories with higher
1128 flexibility and stitch them together implicitly during learning and inference. Although there is no evidence of exact
1129 trend between flexibility and total reward, we find that the trend for novice case amounts to an inverted 'U' curve, that
1130 is, the reward first increases with more flexibility, and then decreases for very high flexibility. The case of stationary
1131 user is interesting. Stationary user takes the same orientation throughout the trajectory, leading to recurrence at every
1132 turn. Thus, entries of stationary user's RP R_1 are set to 1 as seen in Figure 11(c). Flexibility of JRP is then equal to the
1133 flexibility of the RP R_2 ⁸. As evident from the scatter plot in Figure 11, stationary user enjoys a wide range of flexibility
1134 albeit with lower rewards. This is expected due to adversarial actions of User-1, but the high flexibility it offers is
1135

1136⁸JRP = $R_1 \odot R_2$ where all entires of R_1 are 1; so JRP = R_2

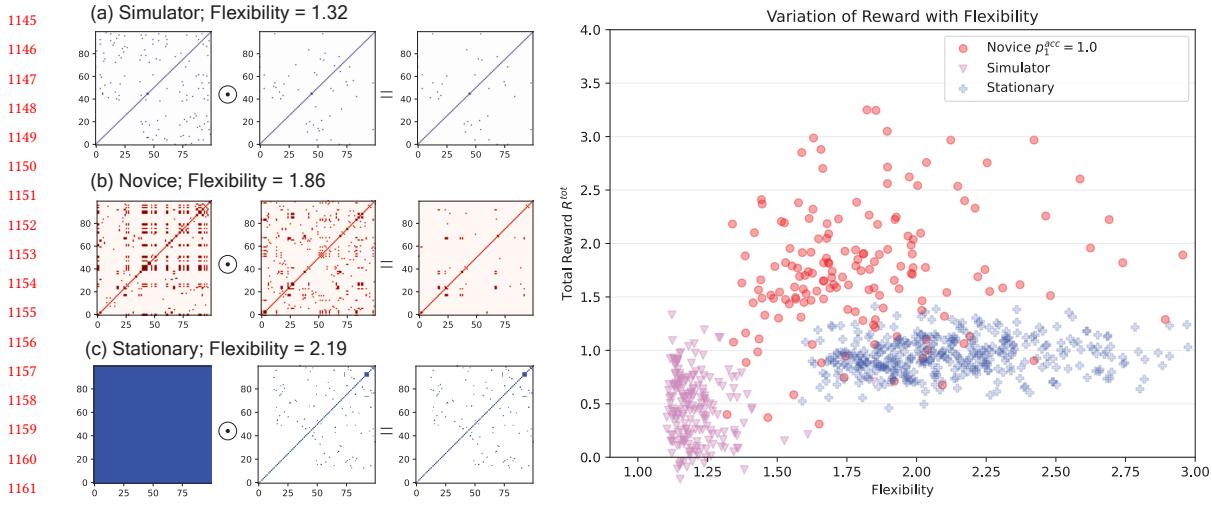


Fig. 11. Recurrence Quantification Analysis of CollabColor system. In the left part of the figure, we display the RPs of User-1 and User-2, namely R_1 (left) and R_2 (middle) respectively, and their joint RP given by $R_1 \odot R_2$ (right). In the right part of the figure, we display a scatter plot of flexibility versus average total reward. We consider trajectories from three cases: (i) Trajectories from Simulator (a); (ii) Trajectories from CollabColor for Novice user with $p_1^{\text{acc}} = 1, p_2^{\text{acc}} = 1$ (b); (iii) Trajectories from CollabColor for Stationary user who picks the same color and same segment always with $p_1^{\text{acc}} = 0, p_2^{\text{acc}} = 1$.

counter-intuitive. A user taking the same actions always actually is the most rigid with no flexibility or consideration of other user's actions. This is a limitation of the RQA method when used in our system's context. Challenges of such coordination measures are discussed at length in Wiltshire et al.'s work [117]. We note the limitation of RQA in our context and delegate further analysis and more robust measures of coordination to future work.

7 HUMAN EVALUATION

Human perception of the output artwork of a creative task is vital in creative design cycle [71]. Several previous works [11, 124, 126] have pointed out the limitations of automated metrics in measuring subjective opinions of creative outputs such as that of image generation, artificial chatbots, etc. Easily calculable metrics such as PSNR, SSIM do not reflect the true quality of colorization, as discussed in detail in Zhang et al.'s work [124]. We conduct an extensive human evaluation to measure the performance of final line-art colorizations created by CollabColor system, assuming $p_i^{\text{acc}} = 1$ for each User- i . As discussed earlier, this serves only as an upper bound of the accuracy obtained when our system is deployed in real-world.

7.1 Setup

We use Amazon Mechanical Turk for crowdsourcing annotations for human evaluation through two surveys that capture subjective opinions such as value, novelty, surprise of colorizations generated by CollabColor. We set the annotator prerequisites as "MTurk Masters" located in the United States having an approval rate more than 95% and at least 30 annotations approved in the past. To arrive at optimal budgeting of our surveys, we conduct several trial runs and estimate the mean time taken to complete them. For our second experiment which displays videos to annotators,

1197 we ensure that the video length is less than one minute by sampling for appropriate number of turns. This is to ensure
1198 that our annotators are not put under severe cognitive load. We pay our annotators at 12 dollars per hour.
1199

1200 *7.1.1 Experiment-1: Estimate the creativity of final colorization.* We build on Parikh et al.'s work [86] to choose two axes
1201 (novelty, value) to measure the subjective notion of creativity. In our trial runs, we displayed the uncolored line-art and
1202 model's colorizations. However, we received feedback from respondents that they needed a reference to ascertain the
1203 novelty or value. We thus displayed the ground-truth colorization too in the main survey as a reference. A sample of
1204 survey questions are displayed in Appendix F. Since novelty and value are abstract concepts, we break them down
1205 into two or more components following See et al. [97]. We consider 50 (uncolored lineart, GT colorization, model
1206 colorization) triplets for this survey, with 10 annotations per triplets.
1207

1208

1209
1210 *7.1.2 Experiment-2: How would you assist in the coloring process?* We create videos at 0.8 frames per second for
1211 turn-taking generated by CollabColor by sampling the number of turns from $\text{Unif}(30, 50)$. We do not show all the 100
1212 turns since it would lead to very long videos. Note that each frame corresponds to the canvas state after each turn, as
1213 shown in the Experiment-2 questionnaire box in Appendix G. Respondents are asked to make certain corrections and
1214 assistance at the last frame of the video. We compare their responses with what the model decides at the last frame
1215 turn. Furthermore, we ask a set of exploratory questions to understand whether a human facilitator can contribute well
1216 for our co-creation task.
1217

1218

1219 7.2 Results

1220

1221

7.2.1 CollabColor achieves significant gains in novelty while adding comparable value. Average across all responses
1222 in Experiment-1 is reported in Table 4. A value of 50% in "More Purchasable?" column for Image-A refers to 50%
1223 respondents feel ground-truth colorization scores more than or equal to CollabColor colorization in purchasability. To
1224 obtain the gain in value of a colorization, we average the gains in sensibility and purchasability. The overall gain in
1225 novelty is reported as the average of gains in surprising-ness and likeliness. Table 4 corroborates the evidence in Parikh
1226 and Zitnick [86] that co-creativity leads to more novelty. The gains in novelty for our system are high, but the value
1227 offered is slightly lower than the ground-truth coloring. Despite a lower value, the average of gains in value and novelty
1228 (corresponding to a vague notion of creativity) is higher for our system. Furthermore, we ask the respondents to choose
1229 an emotion for Image-B and compare with the dominant emotion in the trajectory for Image-B. The emotions match
1230 74% of the time, indicating that CollabColor perceptually meets the requirements of emotion preferences of users. Thus,
1231 CollabColor not only leads to more creativity, but is also a Game with a Purpose (GWAP) [110] unlike Casual Creators.
1232

1233

1234

1235

1236

1237

1238

1239

1240

1241

1242

1243

1244

1245

1246

1247

1248

Qualitatively from the feedback, respondents largely found our survey interesting and fun to complete. One respondent
1233 stated, "*It's like A is traditional and B is modern!*"; another stated, "*Though B is not colored as I feel it should be, if gives a*
1234 *more happy feeling*". We received some negative feedback too; one respondent found a line-art colored by CollabColor
1235 to contain the same color for all segments and did not find it appealing. This could be due to degeneracy in generated
1236 trajectories. Although such colorizations are rare, we intend to tackle this limitation in future work.

Table 4. Results of Experiment-1 in Human Evaluation.

Colorization Type	More Sensible? (↑)	More Purchasable? (↑)	More Value? (↑)	More Surprising? (↑)	Less Likely to color? (↑)	More Novelty (↑)
Ground-truth (Image A)	53%	50%	51.5%	42%	35%	38.5%
CollabColor (Image B)	45%	47%	46%	53%	65%	59%

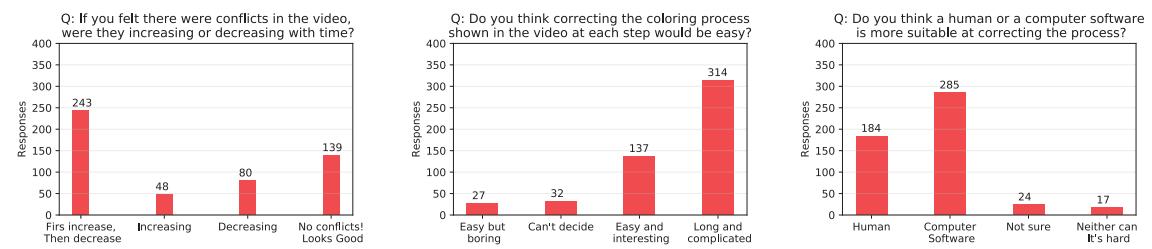


Fig. 12. Results of Experiment-2 in Human Evaluation

can somewhat replace a human facilitator. Although the alignment is not too high, exploratory queries suggest that respondents (61%) did not find the task easy to handle, as shown in Figure 12. 55% of the respondents felt that a software is more suitable to facilitate than a human. These responses encourage the idea of building creative support tools [50, 71] for co-creation, wherein observing and catering to conflicts of two or more creators can get significantly hard for human facilitators.

8 DISCUSSION, LIMITATIONS, AND FUTURE WORK

Our work can be extended by sampling more than two users (discussed in Section 4.2) with a myriad preference traits beyond emotions. Mladenov et al. [76] provide a framework of simulation with evolving preference traits in the context of recommender systems, which can be used to extend our work. We believe that CollabColor is an important step forward in designing facilitator systems that promote collaboration. We envision systems that can not only aid in creativity, but also in other social and professional domains to build trust, shared intentionally [103], and bring populations together to solve pressing global challenges [27]. We provide further discussion of using CollabColor to foster data collection and study co-creativity akin to Casual Creators [24] or GWAP systems [110] in Appendix H.

We do acknowledge some limitations of this work. Our simulation assumes complete observability of user preferences, which is often not desirable by users in real-world scenarios. We also do not consider the consequences of poor synchronization and undo-ing of actions, both of which are vital in building real-time collaborative data structures such as CRDTs [55, 98]. Conflicts in initial stages of collaboration can often lead to significant improvements in final product [95]. We do not explicitly encode this in our reward signals, however our human evaluation results in Figure 12 suggest that conflicts indeed increase and then decrease in CollabColor trajectories. We intend to investigate such evolution of conflicts in future work. In terms of modelling, we do not subdivide and allocate creative tasks to users [118] due to the ambiguity in creative choices. Moreover, our model predicts optimal choices for next user and next color, but does

1301 not predict for next segment. We found segment prediction hard owing to the diversity in shapes, sizes, and counts of
 1302 segments in various line-art SVGs. One way to improve segment prediction is by obtaining rich vector representations
 1303 of shapes of segments using recent innovations such as DeepSVG [17, 91], which can aid in generalization across
 1304 segments in our data.
 1305

1306

9 SOCIETAL IMPACT

1308

1309 Cooperative Intelligence is a desirable feature of all societies, however it can quickly reduce to coercive competition
 1310 if not handled well by social planners [27]. Our work can be considered as computing used for formalism, wherein
 1311 we explicitly define the subjective notions of conflicts and resolving actions/rewards. As discussed in Computing for
 1312 Social Change [10], formalism can easily be gamed to favor one set of population over the other. We take special care in
 1313 CollabColor to ensure that both users' preferences are incorporated and there is no selective favoring of any one user.
 1314 This can be corroborated by our experimental results in Figure 10, wherein the dominance of one user reduces over time
 1315 and a window is provided for other user to incorporate her preferences. Figure 9 shows that even if one user has poor
 1316 network connectivity, our system incorporates her preferences in the final colorization. Furthermore, we consider user
 1317 agency paramount to our interface design and carefully investigate the consequences of obtrusiveness. Our interface
 1318 thus encodes rarely considered values in ML research [16] such as user agency, and focuses on UI interventions that
 1319 respect user's creativity [71]. We point the reader to survey on Cooperative AI [27] for extensive discussion on societal
 1320 impact of facilitators in cooperative tasks.
 1321

1322

1323 Evans et al. [31] show that rewards in RL systems can be tampered to polarize and alter the innate preferences of
 1324 users. While we do not analyze our rewards in context of tampering, we provide a detailed description of each reward in
 1325 Section 4.2 and request readers to carefully consider the consequences when adapting our reward functions. We further
 1326 note that our rewards for conflicts are meant only for the co-creation task of line-art colorization and will probably not
 1327 generalize to any other societal conflicts. We delegate further investigation of our rewards in other societal contexts
 1328 to future work. In particular, we intend to investigate the explainability of action predictions by CollabColor using
 1329 techniques such as attribution [82] for RL Vision as discussed in Hilton et al.'s work [44]. This would help us interpret
 1330 the impact of canvas states and actions in context of final reward obtained, thereby enabling principled fine-tuning of
 1331 reward function choices.
 1332

1333

10 CONCLUSION

1334

1335 We introduce CollabColor, a reinforcement learning-based creative support system that eases the collaboration of two
 1336 users working on a co-creation task of line-art colorization. We present a transformer-based architecture that processes
 1337 the sequence of user actions and canvas states to contextually memorize and suggest conflict-resolving actions via
 1338 non-obtrusive UI interventions. We identify and define key conflicts that occur in the considered co-creation task
 1339 using a formative study. We further use insights from the study to design data simulator and reward signals that aid
 1340 in preparing offline trajectories for training the transformer. Our novel simulator is imbued with techniques such as
 1341 Behavior Cloning and Theory of Mind to simulate realistic trajectories that capture human biases.
 1342

1343

1344 Our evaluation consisting of stress testing with adversarial users and recurrence quantification analysis (RQA)
 1345 demonstrate the effectiveness of CollabColor in incorporating the preferences of both users and reducing conflicts
 1346 between them. Our human studies and exploratory evaluation show that the colorizations produced by CollabColor
 1347 provide significant gains in novelty while maintaining a good value as compared to ground-truth colorizations. Survey
 1348 responses indicate that CollabColor's interventions during co-creation are aligned to those of human facilitators.
 1349

1350

1351

1352

1353 REFERENCES

- [1] 2016. The Quick, Draw! dataset. <https://github.com/googlecreativelab/quickdraw-dataset>
- [2] 2016. Svgpathtools. <https://github.com/mathandy/svgpathtools>
- [3] 2020. *Lucid Research*. <https://lucidspark.com/blog/report-collaboration-and-creativity-during-covid>
- [4] 2021. Icons8. <https://icons8.com/illustrations/>
- [5] 2021. LAB Color Space. https://en.wikipedia.org/wiki/CIELAB_color_space
- [6] 2021. Twilio. <https://www.twilio.com/sync>
- [7] 2021. Undraw.co. <https://undraw.co/illustrations>
- [8] 2021. *The W3C SVG Working Group*. <https://www.w3.org/Graphics/SVG>
- [9] Pieter Abbeel and Andrew Y. Ng. 2004. Apprenticeship learning via inverse reinforcement learning. In *Machine Learning, Proceedings of the Twenty-first International Conference (ICML 2004), Banff, Alberta, Canada, July 4-8, 2004 (ACM International Conference Proceeding Series, Vol. 69)*, Carla E. Brodley (Ed.). ACM. <https://doi.org/10.1145/1015330.1015430>
- [10] Rediet Abebe, Solon Barocas, Jon M. Kleinberg, Karen Levy, Manish Raghavan, and David G. Robinson. 2020. Roles for computing in social change. In *FAT* '20: Conference on Fairness, Accountability, and Transparency, Barcelona, Spain, January 27-30, 2020*, Mireille Hildebrandt, Carlos Castillo, L. Elisa Celis, Salvatore Ruggieri, Linnet Taylor, and Gabriela Zanfir-Fortuna (Eds.). ACM, 252–260. <https://doi.org/10.1145/3351095.3372871>
- [11] Daniel Adiwardana, Minh-Thang Luong, David R. So, Jamie Hall, Noah Fiedel, Romal Thoppilan, Zi Yang, Apoorv Kulshreshtha, Gaurav Nemade, Yifeng Lu, and Quoc V. Le. 2020. Towards a Human-like Open-Domain Chatbot. *CoRR* abs/2001.09977 (2020). arXiv:2001.09977 <https://arxiv.org/abs/2001.09977>
- [12] Michael Bain and Claude Sammut. 1995. A Framework for Behavioural Cloning. In *Machine Intelligence 15, Intelligent Agents [St. Catherine's College, Oxford, UK, July 1995]*, Koichi Furukawa, Donald Michie, and Stephen Muggleton (Eds.). Oxford University Press, 103–129.
- [13] Connelly Barnes, Eli Shechtman, Adam Finkelstein, and Dan B. Goldman. 2009. PatchMatch: a randomized correspondence algorithm for structural image editing. *ACM Trans. Graph.* 28, 3 (2009), 24. <https://doi.org/10.1145/1531326.1531330>
- [14] Vineet Batra, Ankit Phogat, and Mridul Kavidiyal. 2018. General primitives for smooth coloring of vector graphics. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference, SIGGRAPH 2018, Vancouver, BC, Canada, August 12-16, 2018, Posters Proceedings*. ACM, 35:1–35:2. <https://doi.org/10.1145/3230744.3230786>
- [15] Yoann Baveye, Fabrice Urban, Christel Chamaret, Vincent Demoulin, and Pierre Hellier. 2013. Saliency-Guided Consistent Color Harmonization. In *Computational Color Imaging - 4th International Workshop, CCIW 2013, Chiba, Japan, March 3-5, 2013. Proceedings (Lecture Notes in Computer Science, Vol. 7786)*, Shoji Tominaga, Raimondo Schettini, and Alain Trémeau (Eds.). Springer, 105–118. https://doi.org/10.1007/978-3-642-36700-7_9
- [16] Abeba Birhane, Pratyusha Kalluri, Dallas Card, William Agnew, Ravit Dotan, and Michelle Bao. 2021. The Values Encoded in Machine Learning Research. *CoRR* abs/2106.15590 (2021). arXiv:2106.15590 <https://arxiv.org/abs/2106.15590>
- [17] Alexandre Carlier, Martin Danelljan, Alexandre Alahi, and Radu Timofte. 2020. DeepSVG: A Hierarchical Generative Network for Vector Graphics Animation. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (Eds.). <https://proceedings.neurips.cc/paper/2020/hash/bcf9d6bd14a2095866ce8c950b702341-Abstract.html>
- [18] Micah Carroll, Rohin Shah, Mark K. Ho, Thomas L. Griffiths, Sanjit A. Seshia, Pieter Abbeel, and Anca D. Dragan. 2019. On the Utility of Learning about Humans for Human-AI Coordination. *CoRR* abs/1910.05789 (2019). arXiv:1910.05789 <http://arxiv.org/abs/1910.05789>
- [19] Shan Carter and Michael Nielsen. 2017. Using artificial intelligence to augment human intelligence. *Distill* 2, 12 (2017), e9.
- [20] Guillaume Charpiat, Matthias Hofmann, and Bernhard Schölkopf. 2008. Automatic Image Colorization Via Multimodal Predictions. In *Computer Vision - ECCV 2008, 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III (Lecture Notes in Computer Science, Vol. 5304)*, David A. Forsyth, Philip H. S. Torr, and Andrew Zisserman (Eds.). Springer, 126–139. https://doi.org/10.1007/978-3-540-88690-7_10
- [21] Lili Chen, Kevin Lu, Aravind Rajeswaran, Kimin Lee, Aditya Grover, Michael Laskin, Pieter Abbeel, Aravind Srinivas, and Igor Mordatch. 2021. Decision Transformer: Reinforcement Learning via Sequence Modeling. *CoRR* abs/2106.01345 (2021). arXiv:2106.01345 <https://arxiv.org/abs/2106.01345>
- [22] Rohan Choudhury, Gokul Swamy, Dylan Hadfield-Menell, and Anca D. Dragan. 2019. On the Utility of Model Learning in HRI. *CoRR* abs/1901.01291 (2019). arXiv:1901.01291 <http://arxiv.org/abs/1901.01291>
- [23] Yuanzheng Ci, Xinzhu Ma, Zhihui Wang, Haojie Li, and Zhongxuan Luo. 2018. User-Guided Deep Anime Line Art Colorization with Conditional Adversarial Networks. *CoRR* abs/1808.03240 (2018). arXiv:1808.03240 <http://arxiv.org/abs/1808.03240>
- [24] Katherine Compton. 2019. *Casual creators: Defining a genre of autotelic creativity support systems*. University of California, Santa Cruz.
- [25] Nancy J. Cooke, Jamie C. Gorman, Christopher W. Myers, and Jasmine L. Duran. 2013. Interactive Team Cognition. *Cogn. Sci.* 37, 2 (2013), 255–285. <https://doi.org/10.1111/cogs.12009>
- [26] Marius Cordts, Mohamed Omran, Sebastian Ramos, Timo Rehfeld, Markus Enzweiler, Rodrigo Benenson, Uwe Franke, Stefan Roth, and Bernt Schiele. 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. *CoRR* abs/1604.01685 (2016). arXiv:1604.01685 <http://arxiv.org/abs/1604.01685>
- [27] Allan Dafoe, Edward Hughes, Yoram Bachrach, Tantum Collins, Kevin R. McKee, Joel Z. Leibo, Kate Larson, and Thore Graepel. 2020. Open Problems in Cooperative AI. *CoRR* abs/2012.08630 (2020). arXiv:2012.08630 <https://arxiv.org/abs/2012.08630>

- [1405] [28] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, and Neil Houlsby. 2021. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net. <https://openreview.net/forum?id=YicbFdNTTy>
- [1406] [29] Douglas C Engelbart. 1962. Augmenting human intellect: A conceptual framework. *Menlo Park, CA* (1962).
- [1407] [30] Ziv Epstein, Oc  ane Boulais, Skylar Gordon, and Matt Groh. 2020. Interpolating GANs to Scaffold Autotelic Creativity. arXiv:2007.11119 [cs.HC]
- [1408] [31] Charles Evans and Atoosa Kasirzadeh. 2021. User Tampering in Reinforcement Learning Recommender Systems. *arXiv preprint arXiv:2109.04083* (2021).
- [1409] [32] S  bastien Fourey, David Tschumperl  , and David Revoy. 2018. A Fast and Efficient Semi-guided Algorithm for Flat Coloring Line-arts. In *23rd International Symposium on Vision, Modeling, and Visualization, VMV 2018, Stuttgart, Germany, October 10-12, 2018*, Fabian Beck, Carsten Dachsbaier, and Filip Sadlo (Eds.). Eurographics Association, 1–9. <https://doi.org/10.2312/vmv.20181247>
- [1410] [33] Giorgia Franceschelli and Mirco Musolesi. 2021. Creativity and Machine Learning: A Survey. *CoRR abs/2104.02726* (2021). arXiv:2104.02726 <https://arxiv.org/abs/2104.02726>
- [1411] [34] Sayan Ghosh, Jose Echevarria, Vineet Batra, and Ankit Phogat. 2019. Exploring color variations for vector graphics. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference, SIGGRAPH 2019, Los Angeles, CA, USA, July 28 - August 1, 2019, Posters*. ACM, 10:1–10:2. <https://doi.org/10.1145/3306214.3338552>
- [1412] [35] Adam Gleave, Michael Dennis, Cody Wild, Neel Kant, Sergey Levine, and Stuart Russell. 2020. Adversarial Policies: Attacking Deep Reinforcement Learning. In *8th International Conference on Learning Representations, ICLR 2020, Addis Ababa, Ethiopia, April 26-30, 2020*. OpenReview.net. <https://openreview.net/forum?id=HJgEMpVFwB>
- [1413] [36] Gabriel Goh, Nick Cammarata †, Chelsea Voss †, Shan Carter, Michael Petrov, Ludwig Schubert, Alec Radford, and Chris Olah. 2021. Multimodal Neurons in Artificial Neural Networks. *Distill* (2021). <https://doi.org/10.23915/distill.00030> <https://distill.pub/2021/multimodal-neurons>
- [1414] [37] Alison Gopnik and Henry M Wellman. 1994. The theory theory.. In *An earlier version of this chapter was presented at the Society for Research in Child Development Meeting, 1991*. Cambridge University Press.
- [1415] [38] Ning Gu, Jun Xu, Xiaoyuan Wu, Jiangming Yang, and Wei Ye. 2005. Ontology based semantic conflicts resolution in collaborative editing of design documents. *Adv. Eng. Informatics* 19, 2 (2005), 103–111. <https://doi.org/10.1016/j.aei.2005.05.005>
- [1416] [39] Raja Gumienny, Lutz Gericke, Matthias Wenzel, and Christoph Meinel. 2013. Supporting creative collaboration in globally distributed companies. In *Computer Supported Cooperative Work, CSCW 2013, San Antonio, TX, USA, February 23-27, 2013*, Amy S. Bruckman, Scott Counts, Cliff Lampe, and Loren G. Terveen (Eds.). ACM, 995–1007. <https://doi.org/10.1145/2441776.2441890>
- [1417] [40] David Ha and Douglas Eck. 2018. A Neural Representation of Sketch Drawings. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*. OpenReview.net. <https://openreview.net/forum?id=Hy6GHpkCW>
- [1418] [41] David Hanson, Frankie Storm, Wenwei Huang, Vytaas Krisciunas, Tiger Darrow, Audrey Brown, Mengna Lei, Matthew Aylett, Adam Pickrell, and Sophia the Robot. 2020. SophiaPop: Experiments in Human-AI Collaboration on Popular Music. *CoRR abs/2011.10363* (2020). arXiv:2011.10363 <https://arxiv.org/abs/2011.10363>
- [1419] [42] Daniel Hern  ndez, Kevin Denamgana  , Sam Devlin, Spyridon Samothrakis, and James Alfred Walker. 2020. A Comparison of Self-Play Algorithms Under a Generalized Framework. *CoRR abs/2006.04471* (2020). arXiv:2006.04471 <https://arxiv.org/abs/2006.04471>
- [1420] [43] Aaron Hertzmann, Charles E. Jacobs, Nuria Oliver, Brian Curless, and David Salesin. 2001. Image analogies. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2001, Los Angeles, California, USA, August 12-17, 2001*, Lynn Pocock (Ed.). ACM, 327–340. <https://doi.org/10.1145/383259.383295>
- [1421] [44] Jacob Hilton, Nick Cammarata, Shan Carter, Gabriel Goh, and Chris Olah. 2020. Understanding RL Vision. *Distill* (2020). <https://doi.org/10.23915/distill.00029> <https://distill.pub/2020/understanding-rl-vision>.
- [1422] [45] Sepp Hochreiter and J  rgen Schmidhuber. 1997. Long Short-Term Memory. *Neural Comput.* 9, 8 (1997), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [1423] [46] Forrest Huang, Eldon Schoop, David Ha, and John F. Canny. 2020. Scones: towards conversational authoring of sketches. In *IUI '20: 25th International Conference on Intelligent User Interfaces, Cagliari, Italy, March 17-20, 2020*, Fabio Patern  , Nuria Oliver, Cristina Conati, Lucio Davide Spano, and Nava Tintarev (Eds.). ACM, 313–323. <https://doi.org/10.1145/3377325.3377485>
- [1424] [47] Michael Janner, Justin Fu, Marvin Zhang, and Sergey Levine. 2019. When to Trust Your Model: Model-Based Policy Optimization. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alch  -Buc, Emily B. Fox, and Roman Garnett (Eds.). 12498–12509. <https://proceedings.neurips.cc/paper/2019/hash/5faf461eff3099671ad63c6f3f094f7f-Abstract.html>
- [1425] [48] Michael Janner, Qiyang Li, and Sergey Levine. 2021. Reinforcement Learning as One Big Sequence Modeling Problem. *CoRR abs/2106.02039* (2021). arXiv:2106.02039 <https://arxiv.org/abs/2106.02039>
- [1426] [49] Miguel Alonso Jr. 2019. Learning User Preferences via Reinforcement Learning with Spatial Interface Valuing. In *Universal Access in Human-Computer Interaction. Multimodality and Assistive Environments - 13th International Conference, UAHCI 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part II (Lecture Notes in Computer Science, Vol. 11573)*, Margherita Antona and Constantine Stephanidis (Eds.). Springer, 403–418. https://doi.org/10.1007/978-3-030-23563-5_32

- [50] Pegah Karimi, Jeba Rezwana, Safat Siddiqui, Mary Lou Maher, and Nasrin Dehbozorgi. 2020. Creative sketching partner: an analysis of human-AI co-creativity. In *IUI '20: 25th International Conference on Intelligent User Interfaces, Cagliari, Italy, March 17-20, 2020*, Fabio Paternò, Nuria Oliver, Cristina Conati, Lucio Davide Spano, and Nava Tintarev (Eds.). ACM, 221–230. <https://doi.org/10.1145/3377325.3377522>
- [51] Kevin Kelly. 2009. *Out of control: The new biology of machines, social systems, and the economic world*. Hachette UK.
- [52] Jin-Hwa Kim, Nikita Kitaev, Xinlei Chen, Marcus Rohrbach, Byoung-Tak Zhang, Yuandong Tian, Dhruv Batra, and Devi Parikh. 2019. CoDraw: Collaborative Drawing as a Testbed for Grounded Goal-driven Communication. In *Proceedings of the 57th Conference of the Association for Computational Linguistics, ACL 2019, Florence, Italy, July 28- August 2, 2019, Volume 1: Long Papers*, Anna Korhonen, David R. Traum, and Lluís Márquez (Eds.). Association for Computational Linguistics, 6495–6513. <https://doi.org/10.18653/v1/p19-1651>
- [53] Yea-Seul Kim, Mira Dontcheva, Eytan Adar, and Jessica Hullman. 2019. Vocal Shortcuts for Creative Experts. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, CHI 2019, Glasgow, Scotland, UK, May 04-09, 2019*, Stephen A. Brewster, Geraldine Fitzpatrick, Anna L. Cox, and Vassilis Kostakos (Eds.). ACM, 332. <https://doi.org/10.1145/3290605.3300562>
- [54] Max Kleiman-Weiner, Mark K. Ho, Joseph L. Austerweil, Michael L. Littman, and Josh Tenenbaum. 2016. Coordinate to cooperate or compete: Abstract goals and joint intentions in social interaction. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society, Recognizing and Representing Events, CogSci 2016, Philadelphia, PA, USA, August 10-13, 2016*, Anna Papafragou, Daniel Grodner, Daniel Mirman, and John C. Trueswell (Eds.). cognitivesciencesociety.org. <https://mindmodeling.org/cogsci2016/papers/0295/index.html>
- [55] Martin Kleppmann, Adam Wiggins, Peter van Hardenberg, and Mark McGranaghan. 2019. Local-first software: you own your data, in spite of the cloud. In *Proceedings of the 2019 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software, Onward! 2019, Athens, Greece, October 23-24, 2019*, Hidehiko Masuhara and Tomas Petricek (Eds.). ACM, 154–178. <https://doi.org/10.1145/3359591.3359737>
- [56] Paul Knott, Micah Carroll, Sam Devlin, Kamil Ciosek, Katja Hofmann, Anca D. Dragan, and Rohin Shah. 2021. Evaluating the Robustness of Collaborative Agents. In *AAMAS '21: 20th International Conference on Autonomous Agents and Multiagent Systems, Virtual Event, United Kingdom, May 3-7, 2021*, Frank Dignum, Alessio Lomuscio, Ulle Endriss, and Ann Nowé (Eds.). ACM, 1560–1562. <https://dl.acm.org/doi/10.5555/3463952.3464159>
- [57] Elias Kuiter, Sebastian Krieter, Jacob Krüger, Gunter Saake, and Thomas Leich. 2021. variED: an editor for collaborative, real-time feature modeling. *Empir. Softw. Eng.* 26, 2 (2021), 24. <https://doi.org/10.1007/s10664-020-09892-x>
- [58] Aviral Kumar, Xue Bin Peng, and Sergey Levine. 2019. Reward-Conditioned Policies. *CoRR* abs/1912.13465 (2019). arXiv:1912.13465 <http://arxiv.org/abs/1912.13465>
- [59] Aviral Kumar, Aurick Zhou, George Tucker, and Sergey Levine. 2020. Conservative Q-Learning for Offline Reinforcement Learning. In *Advances in Neural Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, December 6-12, 2020, virtual*, Hugo Larochelle, Marc'Aurelio Ranzato, Raia Hadsell, Maria-Florina Balcan, and Hsuan-Tien Lin (Eds.). <https://proceedings.neurips.cc/paper/2020/hash/0d2b2061826a5df3221116a5085a6052-Abstract.html>
- [60] Joel Z Leibo, Edgar A Dueñez-Guzman, Alexander Vezhnevets, John P Agapiou, Peter Sunehag, Raphael Koster, Jayd Matyas, Charlie Beattie, Igor Mordatch, and Thore Graepel. 2021. Scalable Evaluation of Multi-Agent Reinforcement Learning with Melting Pot. In *International Conference on Machine Learning*. PMLR, 6187–6199.
- [61] Debang Li, Huikai Wu, Junge Zhang, and Kaiqi Huang. 2019. Fast A3RL: Aesthetics-Aware Adversarial Reinforcement Learning for Image Cropping. *IEEE Trans. Image Process.* 28, 10 (2019), 5105–5120. <https://doi.org/10.1109/TIP.2019.2914360>
- [62] Mengtian Li, Zhe Lin, Radomir Mech, Ersin Yumer, and Deva Ramanan. 2019. Photo-sketching: Inferring contour drawings from images. In *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*. IEEE, 1403–1412.
- [63] Tzu-Mao Li, Michal Lukáć, Michaël Gharbi, and Jonathan Ragan-Kelley. 2020. Differentiable vector graphics rasterization for editing and learning. *ACM Trans. Graph.* 39, 6 (2020), 193:1–193:15. <https://doi.org/10.1145/3414685.3417871>
- [64] Ziming Li, Julia Kiseleva, and Maarten de Rijke. 2020. Rethinking Supervised Learning and Reinforcement Learning in Task-Oriented Dialogue Systems. In *Findings of the Association for Computational Linguistics: EMNLP 2020, Online Event, 16-20 November 2020 (Findings of ACL, Vol. EMNLP 2020)*, Trevor Cohn, Yulan He, and Yang Liu (Eds.). Association for Computational Linguistics, 3537–3546. <https://doi.org/10.18653/v1/2020.findings-emnlp.316>
- [65] Sharon Lin, Daniel Ritchie, Matthew Fisher, and Pat Hanrahan. 2013. Probabilistic color-by-numbers: suggesting pattern colorizations using factor graphs. *ACM Trans. Graph.* 32, 4 (2013), 37:1–37:12. <https://doi.org/10.1145/2461912.2461988>
- [66] Gang Liu and Jiabao Guo. 2019. Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing* 337 (2019), 325–338.
- [67] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang. 2015. Deep Learning Face Attributes in the Wild. In *Proceedings of International Conference on Computer Vision (ICCV)*.
- [68] Andrew J. Lohn. 2020. Estimating the Brittleness of AI: Safety Integrity Levels and the Need for Testing Out-Of-Distribution Performance. *CoRR* abs/2009.00802 (2020). arXiv:2009.00802 <https://arxiv.org/abs/2009.00802>
- [69] Raphael Gontijo Lopes, David Ha, Douglas Eck, and Jonathon Shlens. 2019. A Learned Representation for Scalable Vector Graphics. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 7929–7938. <https://doi.org/10.1109/ICCV.2019.00802>
- [70] Zhicong Lu, Rubaiat Habib Kazi, Li-Yi Wei, Mira Dontcheva, and Karrie Karahalios. 2021. StreamSketch: Exploring Multi-Modal Interactions in Creative Live Streams. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–26.

- [71] Angus Main and Mick Grierson. 2020. Guru, Partner, or Pencil Sharpener? Understanding Designers' Attitudes Towards Intelligent Creativity Support Tools. *CoRR* abs/2007.04848 (2020). arXiv:2007.04848 <https://arxiv.org/abs/2007.04848>
- [72] Michail Mantzios and Kyriaki Giannou. 2018. When did coloring books become mindful? Exploring the effectiveness of a novel method of mindfulness-guided instructions for coloring books to increase mindfulness and decrease anxiety. *Frontiers in psychology* 9 (2018), 56.
- [73] Norbert Marwan, M Carmen Romano, Marco Thiel, and Jürgen Kurths. 2007. Recurrence plots for the analysis of complex systems. *Physics reports* 438, 5-6 (2007), 237–329.
- [74] Nathan J McNeese, Mustafa Demir, Nancy J Cooke, and Christopher Myers. 2018. Teaming with a synthetic teammate: Insights into human-autonomy teaming. *Human factors* 60, 2 (2018), 262–273.
- [75] Azalia Mirhoseini, Anna Goldie, Mustafa Yazgan, Joe Jiang, Ebrahim M. Songhori, Shen Wang, Young-Joon Lee, Eric Johnson, Omkar Pathak, Sungmin Bae, Azade Nazi, Jiwoo Pak, Andy Tong, Kavya Srinivasa, William Hang, Emre Tuncer, Anand Babu, Quoc V. Le, James Laudon, Richard C. Ho, Roger Carpenter, and Jeff Dean. 2020. Chip Placement with Deep Reinforcement Learning. *CoRR* abs/2004.10746 (2020). arXiv:2004.10746 <https://arxiv.org/abs/2004.10746>
- [76] Martin Mladenov, Chih-Wei Hsu, Vihan Jain, Eugene Ie, Christopher Colby, Nicolas Mayoraz, Hubert Pham, Dustin Tran, Ivan Vendrov, and Craig Boutilier. 2020. Demonstrating Principled Uncertainty Modeling for Recommender Ecosystems with RecSim NG. In *RecSys 2020: Fourteenth ACM Conference on Recommender Systems, Virtual Event, Brazil, September 22–26, 2020*, Rodrygo L. T. Santos, Leandro Balby Marinho, Elizabeth M. Daly, Li Chen, Kim Falk, Noam Koenigstein, and Edleno Silva de Moura (Eds.). ACM, 591–593. <https://doi.org/10.1145/3383313.3411527>
- [77] Fabio Muratore, Michael Gienger, and Jan Peters. 2021. Assessing Transferability From Simulation to Reality for Reinforcement Learning. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 4 (2021), 1172–1183. <https://doi.org/10.1109/TPAMI.2019.2952353>
- [78] Md Sultan Al Nahian, Spencer Frazier, Brent Harrison, and Mark Riedl. 2021. Training Value-Aligned Reinforcement Learning Agents Using a Normative Prior. *CoRR* abs/2104.09469 (2021). arXiv:2104.09469 <https://arxiv.org/abs/2104.09469>
- [79] Patrick Nalepka, Jordan P Gregory-Dunsmore, James Simpson, Gaurav Patil, and Michael J Richardson. 2021. Interaction flexibility in artificial agents teaming with humans. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, Vol. 43.
- [80] Mark J. Nelson, Swen E. Gaudl, Simon Colton, and Sebastian Deterding. 2018. Curious users of casual creators. In *Proceedings of the 13th International Conference on the Foundations of Digital Games, FDG 2018, Malmö, Sweden, August 07–10, 2018*, Steve Dahlskog, Sebastian Deterding, José M. Font, Mitu Khandaker, Carl Magnus Olsson, Sebastian Risi, and Christoph Salge (Eds.). ACM, 61:1–61:6. <https://doi.org/10.1145/3235765.3235826>
- [81] Junhyuk Oh, Yijie Guo, Satinder Singh, and Honglak Lee. 2018. Self-Imitation Learning. *CoRR* abs/1806.05635 (2018). arXiv:1806.05635 <https://arxiv.org/abs/1806.05635>
- [82] Chris Olah, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye, and Alexander Mordvintsev. 2018. The Building Blocks of Interpretability. *Distill* (2018). <https://doi.org/10.23915/distill.00010> <https://distill.pub/2018/building-blocks>.
- [83] Li-Chen Ou and M Ronnier Luo. 2006. A colour harmony model for two-colour combinations. *Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur* 31, 3 (2006), 191–204.
- [84] Moein Owhadi-Kareshk, Sarah Nadi, and Julia Rubin. 2019. Predicting Merge Conflicts in Collaborative Software Development. In *2019 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement, ESEM 2019, Porto de Galinhas, Recife, Brazil, September 19–20, 2019*. IEEE, 1–11. <https://doi.org/10.1109/ESEM.2019.8870173>
- [85] Thomas O'Neill, Nathan McNeese, Amy Barron, and Beau Schelble. 2020. Human–autonomy teaming: A review and analysis of the empirical literature. *Human Factors* (2020), 0018720820960865.
- [86] Devi Parikh and C. Lawrence Zitnick. 2020. Exploring Crowd Co-creation Scenarios for Sketches. In *Proceedings of the Eleventh International Conference on Computational Creativity, ICCC 2020, Coimbra, Portugal, September 7–11, 2020*, F. Amílcar Cardoso, Penousal Machado, Tony Veale, and João Miguel Cunha (Eds.). Association for Computational Creativity (ACC), 73–76. <http://computationalcreativity.net/iccc20/papers/092-iccc20.pdf>
- [87] Jongchan Park, Joon-Young Lee, Donggeun Yoo, and In So Kweon. 2018. Distort-and-recover: Color enhancement using deep reinforcement learning. In *Proceedings of the IEEE Conference on computer vision and pattern recognition*, 5928–5936.
- [88] Elena Petrovskaya, Sebastian Deterding, and Simon Colton. 2020. Casual Creators in the Wild: A Typology of Commercial Generative Creativity Support Tools. In *Proceedings of the Eleventh International Conference on Computational Creativity, ICCC 2020, Coimbra, Portugal, September 7–11, 2020*, F. Amílcar Cardoso, Penousal Machado, Tony Veale, and João Miguel Cunha (Eds.). Association for Computational Creativity (ACC), 65–72. <http://computationalcreativity.net/iccc20/papers/068-iccc20.pdf>
- [89] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. 2021. Learning Transferable Visual Models From Natural Language Supervision. *CoRR* abs/2103.00020 (2021). arXiv:2103.00020 <https://arxiv.org/abs/2103.00020>
- [90] Alec Radford and Karthik Narasimhan. 2018. Improving Language Understanding by Generative Pre-Training.
- [91] Pradyumna Reddy, Michaël Gharbi, Michal Lukáć, and Niloy J. Mitra. 2021. Im2Vec: Synthesizing Vector Graphics Without Vector Supervision. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19–25, 2021*. Computer Vision Foundation / IEEE, 7342–7351. https://openaccess.thecvf.com/content/CVPR2021/html/Reddy_Im2Vec_Synthesizing_Vector_Graphics_Without_Vector_Supervision_CVPR_2021_paper.html
- [92] Marco Tulio Ribeiro, Tongshuang Wu, Carlos Guestrin, and Sameer Singh. 2020. Beyond Accuracy: Behavioral Testing of NLP Models with CheckList. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics,

- 1561 Online, 4902–4912. <https://doi.org/10.18653/v1/2020.acl-main.442>
- 1562 [93] Adam Roberts, Curtis Hawthorne, and Ian Simon. 2018. Magenta.js: a JavaScript API for augmenting creativity with deep learning. (2018).
- 1563 [94] Catherine Sauvaget, Stéphane Manuel, Jean-Noël Vittaut, Jordane Suarez, and Vincent Boyer. 2010. Segmented Images Colorization Using Harmony. In *Sixth International Conference on Signal-Image Technology and Internet-Based Systems, SITIS 2010, Kuala Lumpur, Malaysia, December 15–18, 2010*, Kokou Yétongonon, Albert Dipanda, and Richard Chbeir (Eds.). IEEE Computer Society, 153–160. <https://doi.org/10.1109/SITIS.2010.35>
- 1564 [95] Mary Scannell. 2010. *The big book of conflict resolution games: Quick, effective activities to improve communication, trust and collaboration*. McGraw Hill Professional.
- 1565 [96] J. Schmidhuber. 2019. Reinforcement Learning Upside Down: Don't Predict Rewards - Just Map Them to Actions. *ArXiv* abs/1912.02875 (2019).
- 1566 [97] Abigail See, Stephen Roller, Douwe Kiela, and Jason Weston. 2019. What makes a good conversation? How controllable attributes affect human judgments. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2–7, 2019, Volume 1 (Long and Short Papers)*, Jill Burstein, Christy Doran, and Thamar Solorio (Eds.). Association for Computational Linguistics, 1702–1723. <https://doi.org/10.18653/v1/n19-1170>
- 1567 [98] Marc Shapiro, Nuno Preguiça, Carlos Baquero, and Marek Zawirski. 2011. Conflict-free replicated data types. In *Symposium on Self-Stabilizing Systems*. Springer, 386–400.
- 1568 [99] Yang Shi, Yang Wang, Ye Qi, John Chen, Xiaoyao Xu, and Kwan-Liu Ma. 2017. IdeaWall: Improving Creative Collaboration through Combinatorial Visual Stimuli. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW 2017, Portland, OR, USA, February 25 – March 1, 2017*, Charlotte P. Lee, Steven E. Poltrock, Louise Barkhuus, Marcos Borges, and Wendy A. Kellogg (Eds.). ACM, 594–603. <https://doi.org/10.1145/2998181.2998208>
- 1569 [100] Rupesh Kumar Srivastava, Pranav Shyam, Filipe Mutz, Wojciech Jaskowski, and Jürgen Schmidhuber. 2019. Training Agents using Upside-Down Reinforcement Learning. *CoRR* abs/1912.02877 (2019). arXiv:1912.02877 <http://arxiv.org/abs/1912.02877>
- 1570 [101] Tsai-Ho Sun, Chien-Hsun Lai, Sai-Keung Wong, and Yu-Shuen Wang. 2019. Adversarial Colorization of Icons Based on Contour and Color Conditions. In *Proceedings of the 27th ACM International Conference on Multimedia, MM 2019, Nice, France, October 21–25, 2019*, Laurent Amsaleg, Benoit Huet, Martha A. Larson, Guillaume Gravier, Hayley Hung, Chong-Wah Ngo, and Wei Tsang Ooi (Eds.). ACM, 683–691. <https://doi.org/10.1145/3343031.3351041>
- 1571 [102] Jianchao Tan, Jose I. Echevarria, and Yotam I. Gingold. 2018. Palette-based image decomposition, harmonization, and color transfer. *CoRR* abs/1804.01225 (2018). arXiv:1804.01225 <http://arxiv.org/abs/1804.01225>
- 1572 [103] Michael Tomasello, Malinda Carpenter, Josep Call, Tanya Behne, and Henrike Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and brain sciences* 28, 5 (2005), 675–691.
- 1573 [104] Hung-Yu Tseng, Matthew Fisher, Jingwan Lu, Yijun Li, Vladimir G. Kim, and Ming-Hsuan Yang. 2020. Modeling Artistic Workflows for Image Generation and Editing. In *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XVIII (Lecture Notes in Computer Science, Vol. 12363)*, Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm (Eds.). Springer, 158–174. https://doi.org/10.1007/978-3-030-58523-5_10
- 1574 [105] Tao Tu, Qing Ping, Govindarajan Thatтай, Gökhan Tür, and Prem Natarajan. 2021. Learning Better Visual Dialog Agents With Pretrained Visual-Linguistic Representation. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2021, virtual, June 19–25, 2021*. Computer Vision Foundation / IEEE, 5622–5631. https://openaccess.thecvf.com/content/CVPR2021/html/Tu_Learning_Better_Visual_Dialog_Agents_With_Pretrained_Visual-Linguistic_Representation_CVPR_2021_paper.html
- 1575 [106] Kathleen Tuite and Adam M Smith. 2012. Emergent remix culture in an anonymous collaborative art system. In *Eighth artificial intelligence and interactive digital entertainment conference*.
- 1576 [107] Jonathan Uesato, Ananya Kumar, Csaba Szepesvári, Tom Erez, Avraham Ruderman, Keith Anderson, Krishnamurthy (Dj) Dvijotham, Nicolas Heess, and Pushmeet Kohli. 2019. Rigorous Agent Evaluation: An Adversarial Approach to Uncover Catastrophic Failures. In *7th International Conference on Learning Representations, ICLR 2019, New Orleans, LA, USA, May 6–9, 2019*. OpenReview.net. <https://openreview.net/forum?id=B1xhQhRcK7>
- 1577 [108] Hado van Hasselt, Yotam Doron, Florian Strub, Matteo Hessel, Nicolas Sonnerat, and Joseph Modayil. 2018. Deep Reinforcement Learning and the Deadly Triad. *CoRR* abs/1812.02648 (2018). arXiv:1812.02648 <http://arxiv.org/abs/1812.02648>
- 1578 [109] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention Is All You Need. *CoRR* abs/1706.03762 (2017). arXiv:1706.03762 <http://arxiv.org/abs/1706.03762>
- 1579 [110] Luis Von Ahn and Laura Dabbish. 2008. Designing games with a purpose. *Commun. ACM* 51, 8 (2008), 58–67.
- 1580 [111] Sheng-Yi Wang, Oliver Wang, Andrew Owens, Richard Zhang, and Alexei A. Efros. 2019. Detecting Photoshopped Faces by Scripting Photoshop. *CoRR* abs/1906.05856 (2019). arXiv:1906.05856 <http://arxiv.org/abs/1906.05856>
- 1581 [112] Mary Beth Watson-Manheim. 2019. *Discontinuities, Continuities, and Hidden Work in Virtual Collaboration*. Springer International Publishing, Cham, 121–132. https://doi.org/10.1007/978-3-319-94487-6_6
- 1582 [113] Charles L Webber Jr and Joseph P Zbilut. 1994. Dynamical assessment of physiological systems and states using recurrence plot strategies. *Journal of applied physiology* 76, 2 (1994), 965–973.
- 1583 [114] Michael J. Wilber, Chen Fang, Hailin Jin, Aaron Hertzmann, John P. Collomosse, and Serge J. Belongie. 2017. BAM! The Behance Artistic Media Dataset for Recognition Beyond Photography. *CoRR* abs/1704.08614 (2017). arXiv:1704.08614 <http://arxiv.org/abs/1704.08614>
- 1584 [115] Ronald J. Williams and David Zipser. 1989. A Learning Algorithm for Continually Running Fully Recurrent Neural Networks. *Neural Comput.* 1, 2 (1989), 270–280. <https://doi.org/10.1162/neco.1989.1.2.270>

- 1613 [116] Travis J Wiltshire, Jonathan E Butner, and Stephen M Fiore. 2018. Problem-solving phase transitions during team collaboration. *Cognitive science*
1614 42, 1 (2018), 129–167.
- 1615 [117] Travis J Wiltshire, Sune Vork Steffensen, and Aaron D Likens. 2020. Challenges for using coordination-based measures to augment collaborative
1616 social interactions. In *Selbstorganisation—ein Paradigma für die Humanwissenschaften*. Springer, 215–230.
- 1617 [118] Sarah A. Wu, Rose E. Wang, James A. Evans, Joshua B. Tenenbaum, David C. Parkes, and Max Kleiman-Weiner. 2021. Too Many Cooks: Bayesian
1618 Inference for Coordinating Multi-Agent Collaboration. *Top. Cogn. Sci.* 13, 2 (2021), 414–432. <https://doi.org/10.1111/tops.12525>
- 1619 [119] Peng Xu, Timothy M Hospedales, Qiyue Yin, Yi-Zhe Song, Tao Xiang, and Liang Wang. 2020. Deep learning for free-hand sketch: A survey and a
1620 toolbox. *arXiv preprint arXiv:2001.02600* (2020).
- 1621 [120] Jimei Yang, Brian L. Price, Scott Cohen, Honglak Lee, and Ming-Hsuan Yang. 2016. Object Contour Detection with a Fully Convolutional
1622 Encoder-Decoder Network. In *2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2016, Las Vegas, NV, USA, June 27–30, 2016*.
1623 IEEE Computer Society, 193–202. <https://doi.org/10.1109/CVPR.2016.28>
- 1624 [121] Li Yujian and Liu Bo. 2007. A normalized Levenshtein distance metric. *IEEE transactions on pattern analysis and machine intelligence* 29, 6 (2007),
1625 1091–1095.
- 1626 [122] Lvmin Zhang, Chengze Li, Tien-Tsin Wong, Yi Ji, and Chunping Liu. 2018. Two-stage sketch colorization. *ACM Trans. Graph.* 37, 6 (2018),
1627 261:1–261:14. <https://doi.org/10.1145/3272127.3275090>
- 1628 [123] Richard Zhang, Phillip Isola, and Alexei A. Efros. 2016. Colorful Image Colorization. In *Computer Vision - ECCV 2016 - 14th European Conference,
1629 Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part III (Lecture Notes in Computer Science, Vol. 9907)*, Bastian Leibe, Jiri Matas, Nicu
Sebe, and Max Welling (Eds.). Springer, 649–666. https://doi.org/10.1007/978-3-319-46487-9_40
- 1630 [124] Richard Zhang, Jun-Yan Zhu, Phillip Isola, Xinyang Geng, Angela S. Lin, Tianhe Yu, and Alexei A. Efros. 2017. Real-Time User-Guided Image
1631 Colorization with Learned Deep Priors. *CoRR* abs/1705.02999 (2017). arXiv:1705.02999 <http://arxiv.org/abs/1705.02999>
- 1632 [125] Wenshuai Zhao, Jorge Peña Queralta, and Tomi Westerlund. 2020. Sim-to-Real Transfer in Deep Reinforcement Learning for Robotics: a
1633 Survey. In *2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020, Canberra, Australia, December 1–4, 2020*. IEEE, 737–744. <https://doi.org/10.1109/SSCI47803.2020.9308468>
- 1634 [126] Sharon Zhou, Mitchell L. Gordon, Ranjay Krishna, Austin Narcomey, Li Fei-Fei, and Michael Bernstein. 2019. HYPE: A Benchmark for Human eYe Per-
1635 ceptual Evaluation of Generative Models. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing
1636 Systems 2019, NeurIPS 2019, December 8–14, 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d’Alché-
1637 Buc, Emily B. Fox, and Roman Garnett (Eds.). 3444–3456. <https://proceedings.neurips.cc/paper/2019/hash/65699726a3c601b9f31bf04019c8593c-Abstract.html>
- 1638
- 1639
- 1640
- 1641
- 1642
- 1643
- 1644
- 1645
- 1646
- 1647
- 1648
- 1649
- 1650
- 1651
- 1652
- 1653
- 1654
- 1655
- 1656
- 1657
- 1658
- 1659
- 1660
- 1661
- 1662
- 1663
- 1664

1665 A ANECDOTAL EVIDENCE SUGGESTS OUR INTERFACE IS USEFUL

1666 At the time of writing, we did not conduct a large-scale exploratory study of our interface with CollabColor toggled on.
 1667 From anecdotal evidence of 10 user-user pairs (novices, design researchers), we find that our system is fairly easy to
 1668 understand and use. They also felt our system eases collaboration, leading to coherent colorizations. We have included
 1669 a supplementary video in the attached zip file⁹, and invite readers to view this video for better understanding of our
 1670 interface in action. While our quantitative and qualitative results including stress tests and adversarial users show that
 1671 our system fares well in facilitation, we intend to pursue a large exploration of our interface in future work.
 1672

1673 B FURTHER RELATED WORK

1674 B.1 Upside-down Reinforcement Learning for training on suboptimal trajectories

1675 Our simulated trajectories are suboptimal since we neither use self-play nor collect expensive expert data. Recent results
 1676 by Kumar et al. [58] show that learning from suboptimal policies is very much possible using offline reinforcement
 1677 learning (RL) and reward-conditioned policies. Offline RL is a “data-driven” paradigm of RL that forgoes active data
 1678 collection and instead utilizes previously collected large experience data to train high-capacity scalable models that
 1679 are introduced in supervised learning (SL) literature. This enables RL practitioners to use stable and reliable super-
 1680 vised algorithms that have resulted in substantial progress in AI over the last decade and impart a “common-sense”
 1681 generalization in their systems [59]. However, offline RL methods have been shown to overfit [59] due to excessive
 1682 training on given datasets and lack of active feedback from the environment. We partially bypass these concerns since
 1683 our simulated data is created using several diverse line-arts that decreases the chances of overfitting.

1684 Within the paradigm of offline RL, upside-down RL [96, 100] that predicts actions instead of rewards has gained
 1685 traction for its conceptually simple perspective and easy inclusion of powerful SL architectures. Contrary to policy
 1686 based RL methods that only look for actions that maximize the reward, upside-down RL is more general and learns
 1687 actions that achieve any required reward making it suitable to train on suboptimal sequences. Decision transformer [21]
 1688 and Trajectory transformer [48] are two instantiations of upside-down RL that offer a sequence modelling objective on
 1689 collected experience trajectories. Transformers [109] enable larger context length than LSTMs [45] and thereby help in
 1690 bypassing long-term credit assignment problem of RL [108].

1691 C DATA PREPARATION FOR LINE-ART FLAT COLORIZATION

1692 Existing colorization datasets used in works such as [23, 101, 122–124] store images pixel-wise. This storage is suitable
 1693 for tasks such as coloring a black-and-white image [123], coloring a line-art anime character based on user strokes [23]
 1694 or color style transfer [101]. However, it is not suitable for our flat colorization task which constrains pixels within a
 1695 closed segment to have the same color. One way is to store a segmentation mask for each closed segment region to bind
 1696 the pixels together, akin to various semantic segmentation datasets [26, 67]. But the images in these datasets are natural
 1697 with high frequency attributes and gradient colors, making them unsuitable for flat coloring. In addition, segmentation
 1698 mask images themselves take up extra storage. Another way to create a dataset for our task is to extract line-art and
 1699 segment flat-colored images by clustering pixels using a depth-first search approach [62, 120]. However, this method
 1700 suffers from leaky boundary issues leading to poor quality line-arts after clustering. Finally, we can generate our own
 1701 images with randomized strokes and closed paths while storing the pixel-binding information. However, these images
 1702 lack semantic meaning making it difficult for users to color.

1703 ⁹The demo video is titled *Demo_CollabColor.mp4*

In contrast to raster images, Scalable Vector Graphics (SVG) images encode different shapes in the form of hierarchical open/closed paths in an XML-based format [8]. Given such mathematical encoding, they can be scaled to any resolution without aliasing. Each closed path (segment) can be provided a fill attribute that encodes the color of the segment. Setting the fill attribute as “white” for all segments thereby gives us an uncolored line-art image. This makes SVGs suitable for our task of flat colorization of line-arts. Current SVG datasets are small and contain open sketches [1], fonts [69], or simple icons [17], all of which are uncolored with very few or no segments. Online SVG image sources are either not free [4] or contain very few and large complex illustrations for UI design [7].

D QUALITATIVE EXAMPLES

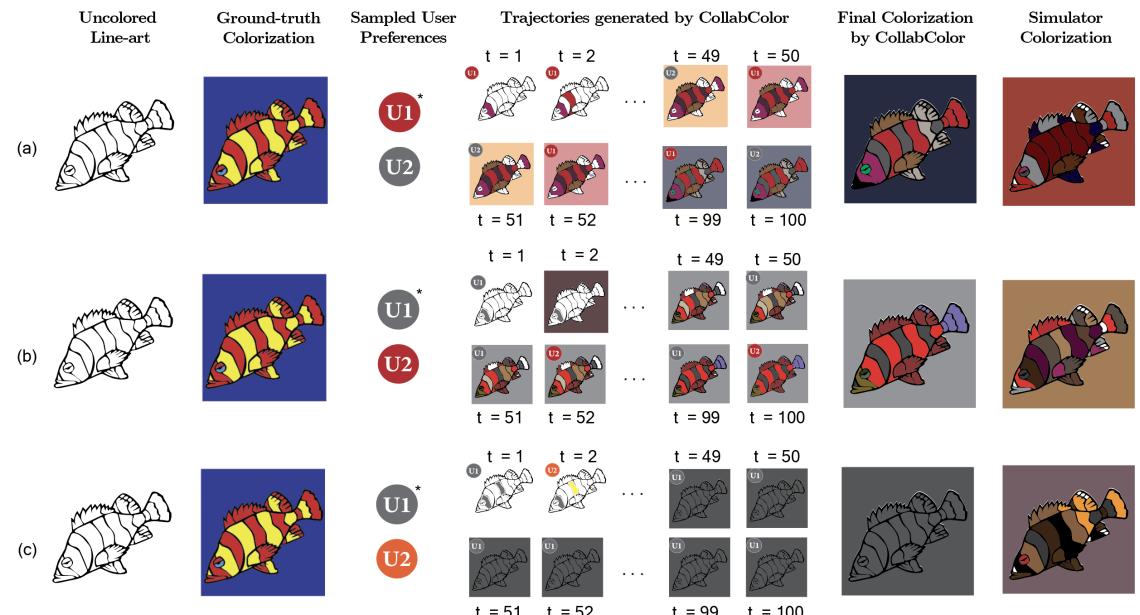


Fig. 13. Qualitative examples of our system. Users with greater dominance are labelled with a (*) symbol.

Figure 13 shows qualitative examples of trajectories generated by CollabColor for novice user type (that is, $p_{acc}^i = 1$). Additionally, we also show the ground-truth and simulator colorization (used in training) for comparison. In Part (a), the sampled emotions are “Scary” for User-1 and “Gloomy” for User-2 with User-1 having greater dominance. We can see that our trajectories are able to achieve sensible patterns in the colorization unlike the simulated output. In part (b), the dominances are same but sampled emotions are reversed. We can see that the final colorization looks more gloomy while maintaining good patterns and incorporating both users’ preferences. Some colors that lie in both palettes of users such as “purple” are also picked. In the trajectories, we can see that the model predicts the users and their respective preferences well. In part (b) at $t = 100$, U2 colors a part of the line-art in gray, showing that U2 is able to understand U1’s perspective. These color predictions are shown as interventions to promote conflict resolutions. The patterns in both colorizations show that some level of work subdivision happens within CollabColor’s trajectories, although we have not studied its consequences in detail. We delegate it to future work.

Part (c) shows an example of degenerate trajectories generated by CollabColor. Sometimes the first color or first segment picked derails the entire generated trajectory, especially if the emotions are incompatible (shown quantitatively in 3). Here, all the segments are colored gray towards $t = 50$, and such degeneracy persists towards the end with the same user U1 being predicted throughout. In fact, the simulated colorization looks better because rule-based simulators rarely result in degenerate outputs. We intend to correct this error in future work. One way is to use better sampling during Beam Search at the inference stage with temperature softmax to promote diversity of sampled actions.

E PICKER MECHANISMS IN SIMULATION

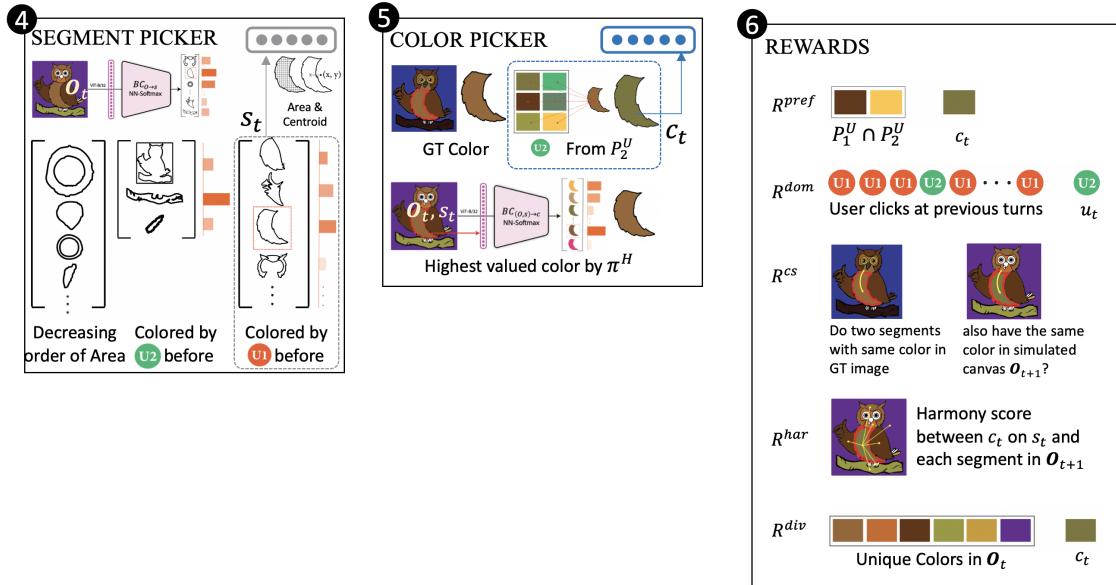


Fig. 14. Enlarged images of Segment Picker, Color Picker, and Reward Signals.

We display enlarged schematic diagrams of both segment and color pickers in Figure 15.

Strategic Choice: Segment Picker. The second module picks a segment s_t ($\in 1 \dots |S|$) to be colored by user u_t by following one of the three mechanisms described here:

- Pick deterministically from uncolored segments in decreasing order of area. Figure 14 plots the variation of areas picked by users over time across all modes in our formative study. We observe that users pick larger segments initially owing to their greater visibility.
- Pick a segment that the selected user u_t has colored before. This incorporates our finding [F1] which states that users tend to alter the colors of segments they have already colored. Of these colored segments, we pick the segment that has the highest value as computed by $BC_{S \rightarrow s}$ (from human policy π^H).
- Pick a segment that the other user has colored before. This again caters to our finding of cross-recoloring [F1]. We use $BC_{O \rightarrow s}$ to pick the highest valued segment among those colored by the other user.

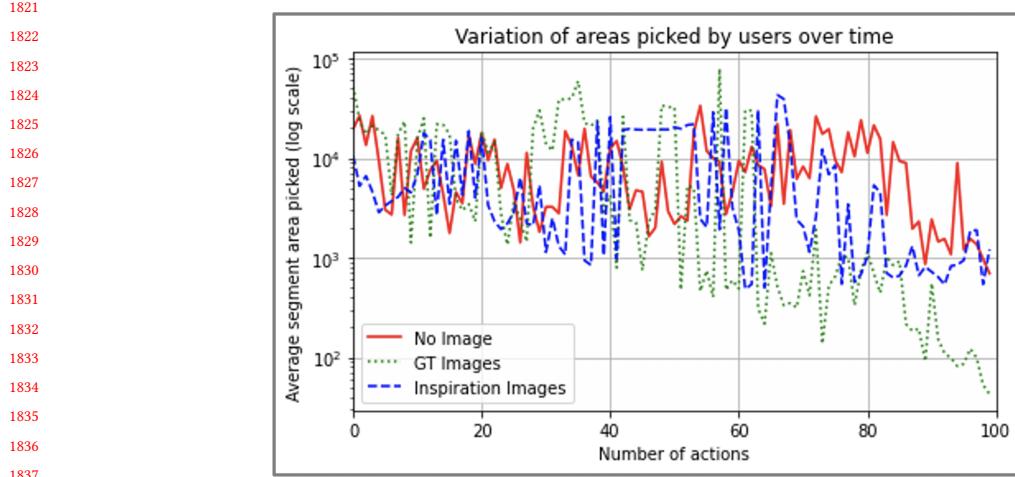


Fig. 15. Variation of Areas picked by users over time in all three modes of our formative study. This helps us in designing a mechanism for Segment Picker.

We compute the probabilities of these mechanisms from the frequencies of their occurrence in mode 3 (with inspiration images) and sample a mechanism from Multinoulli distribution. The representation of a segment i (S_i) is a $(|S| + 3)$ -dimensional vector consisting of the concatenation of one-hot vector of the segment ($v_i^{(s)}$), area of the segment, and centroids of the segment [2], where $|S|$ is the number of segments in the chosen line-art.

Motion Choice: Color Picker. The third module picks a color c_t ($\in 1 \dots 56$) given the user u_t and segment s_t by following one of these three mechanisms with equal probability. As explained earlier, each color i is represented by a 56-dimensional one-hot vector C_i .

- *Pick deterministically the color of segment s_t from GT image.* This reflects the common-sense ability of users to visualize the final output of a partially colored line-art. For instance, the background segment color in an uncolored line-art of fish is likely imagined to be blue by most users. This blue color can be readily picked by “peeking” into the GT image [43, 124] which is also likely to contain a blue background.
- *Pick a color from $P_{u_t}^U$ with the closest LAB space distance to GT color.* This is like the previous mechanism; except we stress more on user color preferences based on emotions (mood-board). We compute the widely used scaled Euclidean distance in LAB color space [5] to measure closeness of colors in $P_{u_t}^U$ to GT color.
- *Pick the highest-valued color by human policy π^H .* This caters to the model $BC(O, s) \rightarrow c$ that has learned to color from the past behavior of humans who have colored similar line-arts in the formative study.

Having obtained the action triplet (u_t, s_t, c_t) , we update the line-art by coloring s_t with c_t and recompute the CLIP representation of canvas state as O_{t+1} . The turn-taking process terminates when $t = 100$.

F HUMAN EVALUATION EXPERIMENT-1 QUESTIONNAIRE

1873
1874 Experiment-1: We display three images to each respondent: an uncolored line-art, its ground-truth
1875 colorization (Image A), and colorization by CollabColor (Image B).
1876



1881 [Sanity Check] Which category does the line-art belong to?
1882

- 1883 • Animal • Human • Scenery • None of the Above

1884 [Sensible] Which of the two colorizations makes more sense?
1885

- 1886 • Image A • Image B • Both make perfect sense • Neither makes any sense
1887

1888 [Value] Suppose you are asked to purchase one of the two colorizations at an art exhibi-
1889 tion. Which one would you prefer? • Image A • Image B • I'd take both! • Neither!
1890

1891 [Surprising] Which of the two colorizations is more surprising to you?
1892

- 1893 • Image A • Image B • Both are surprising • Neither
1894

1895 [Preference Alignment] Images can make us feel various emotions. What do you think is
1896 the emotion of coloring B? • Happy • Scary • Gloomy • Peaceful
1897

1898 [Feedback] Please provide a feedback on this survey. An empty textbox is provided.
1899

1900

1901

1902

1903

1904

1905

1906

1907

1908

1909

1910

1911

1912

1913

1914

1915

1916

1917

1918

1919

1920

1921

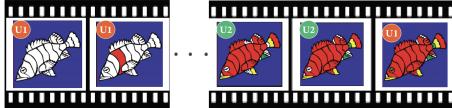
1922

1923

1924

1925 G HUMAN EVALUATION EXPERIMENT-2 QUESTIONNAIRE

1926 Experiment-2: We display a video of colorization of line-art by CollabColor, along with the users and their
 1927 emotions at each turn.



1933 [Sanity Check] Which category does the line-art belong to?

- 1934 • Animal • Human • Scenery • None of the Above

1936 [Collaboration Score] Do you think the coloring was conflicting? If you think there was
 1937 no planning in the coloring video shown, and it seemed random taking many directions of
 1938 coloring, it is conflicting • Not at all • A little • Somewhat • Very Conflicting!

1941 [Preference Alignment] Images can make us feel various emotions. What do you think is
 1942 the emotion of coloring process? • Happy • Scary • Gloomy • Peaceful

1945 [Color Prediction] If you could choose colors and apply on the frame at the end of the
 1946 coloring process in video, which color would you add? • Orange • Red • Gray • Light Green

1949 [Feedback] Please provide a feedback on this survey. An empty textbox is provided.

1951 H FURTHER DISCUSSION

1953 CollabColor demonstrates an overall framework of how support systems can be built in the context of creative
 1954 collaboration while not being obtrusive: (i) formative study to understand pain points of the considered creative task; (ii)
 1955 simulation to alleviate issues of lack of data; (iii) RL training to teach and help the system internalize what constitutes
 1956 high rewards; (iv) UI intervention and assistance design to ensure non-obtrusiveness. While we consider the co-creative
 1957 task of line-art colorization, this framework can be readily extended to other creative tasks such as co-sketching
 1958 [40, 46, 50, 119], image enhancement [87], conversational creativity [52, 105], music [41], and various other modalities.
 1959 As a starting point, CollabColor interface can be used as a test-bed for data collection to study co-creativity akin
 1960 to GWAP systems [110]. Flat colorization is an easier task for users to align preferences as compared to free-form
 1961 co-sketching that has a very vast design space. Also, research has shown that coloring aids in mindfulness and reduction
 1962 of anxiety in adults [72]. These features further encourage the use of our interface as a data-collection system for
 1963 studying co-creativity.