

We-toon: A Communication Support System between Writers and Artists in Collaborative Webtoon Sketch Revision

Hyung-Kwon Ko*
NAVER WEBTOON Corp.
Republic of Korea
hyungkwonko@gmail.com

Subin An*
Seoul National University
Republic of Korea
sban@hcil.snu.ac.kr

Gwanmo Park
Seoul National University
Republic of Korea
gmpark@hcil.snu.ac.kr

Seung Kwon Kim
NAVER WEBTOON Corp.
Republic of Korea
mark.kim@webtooncorp.com

Daesik Kim
NAVER WEBTOON Corp.
Republic of Korea
daesik.kim@webtooncorp.com

Bohyoung Kim
Hankuk University of Foreign Studies
Republic of Korea
bkim@hufs.ac.kr

Jaemin Jo
Sungkyunkwan University
Republic of Korea
jmjo@skku.edu

Jinwook Seo
Seoul National University
Republic of Korea
jseo@snu.ac.kr

ABSTRACT

We present a communication support system, namely *We-toon*, that can bridge the webtoon writers and artists during sketch revision (i.e., character design and draft revision). In the highly iterative design process between the webtoon writers and artists, writers often have difficulties in precisely articulating their feedback on sketches owing to their lack of drawing proficiency. This drawback makes the writers rely on textual descriptions and reference images found using search engines, leading to indirect and inefficient communications. Inspired by a formative study, we designed *We-toon* to help writers revise webtoon sketches and effectively communicate with artists. Through a GAN-based image synthesis and manipulation, *We-toon* can interactively generate diverse reference images and synthesize them locally on any user-provided image. Our user study with 24 professional webtoon authors demonstrated that *We-toon* outperforms the traditional methods in terms of communication effectiveness and the writers' satisfaction level related to the revised image.

CCS CONCEPTS

• **Human-centered computing** → **Collaborative content creation; User interface design; Interface design prototyping; Systems and tools for interaction design.**

*Both authors contributed equally to this research.

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 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 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 redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
UIST '22, October 29–November 2, 2022, Bend, OR, USA

© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-9320-1/22/10...\$15.00
<https://doi.org/10.1145/3526113.3545612>

KEYWORDS

Webtoon; Collaboration; Communication; Human-centered AI; Creativity Support; Interactive System; User Interface; GAN; Interview; Usability Study

ACM Reference Format:

Hyung-Kwon Ko, Subin An, Gwanmo Park, Seung Kwon Kim, Daesik Kim, Bohyoung Kim, Jaemin Jo, and Jinwook Seo. 2022. We-toon: A Communication Support System between Writers and Artists in Collaborative Webtoon Sketch Revision. In *The 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22)*, October 29–November 2, 2022, Bend, OR, USA. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3526113.3545612>

1 INTRODUCTION

Webtoon, a compound word of web and cartoon, is a digital form of comics. In contrast to the traditional paper-printed comics, webtoons have relatively short (e.g., weekly) release schedules and are fully colored, occasionally embellished with dynamic contents, such as music or animation. With the advancement and proliferation of mobile phones, the webtoon industry has rapidly grown, and is becoming more complicated and systematic, driven by professional authors and agencies. Modern webtoons are created through collaboration between two specialized authors, namely a writer and an artist; the writer composes the plot and the dialogues between characters while the artist turns them into illustrations. Such a partnership between a writer and an artist brings many advantages as it enables them to specialize their abilities for higher quality webtoons as well as allows those with incomplete skill-sets to join the field.

As the collaboration between writers and artists becomes more commonplace, communication plays an essential role in successful webtoon publication. The communication usually starts from the writers as they convey what they envision to the artists, such as the appearance of the main characters. The easiest and the most direct form of communication would be to draw sketches. However, we observed many writers actually resort to text descriptions, even for subtle and nuanced details that are hard to describe with words as they are understandably not proficient in drawing. Such an indirect

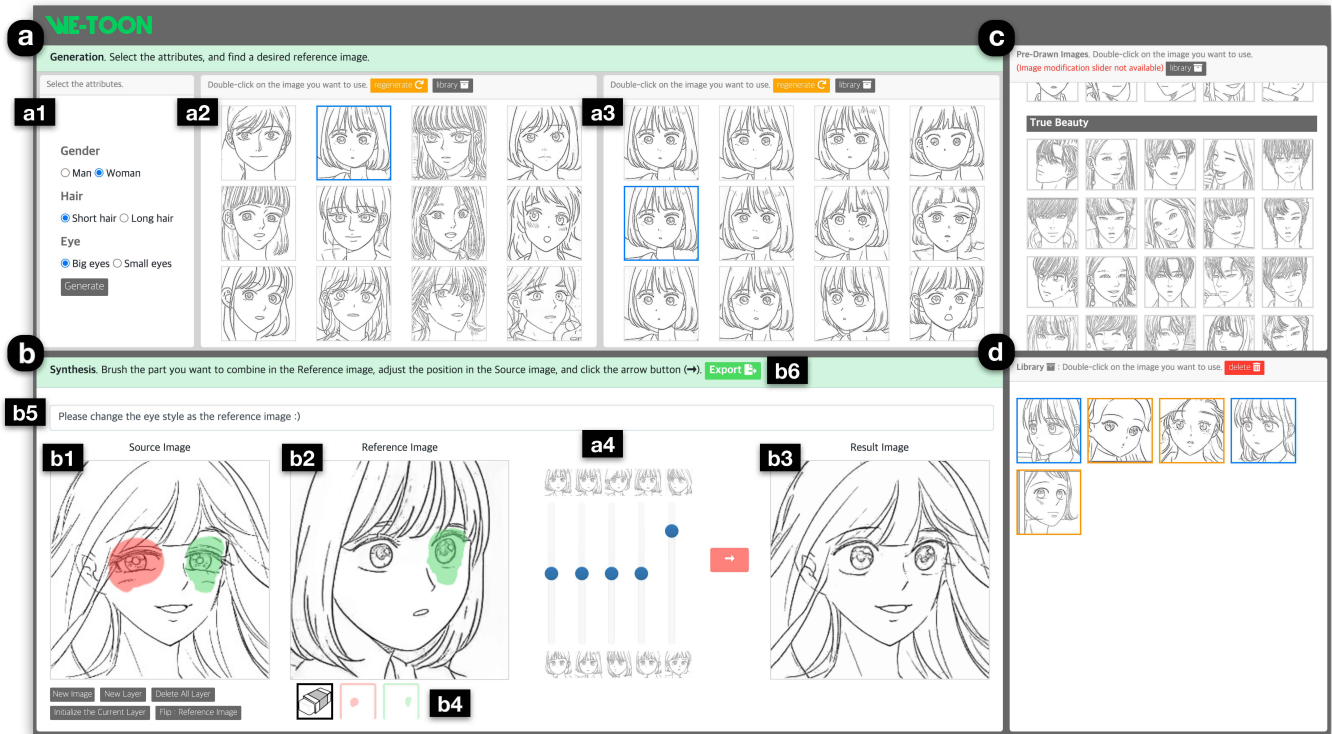


Figure 1: We-toon allows the writers to make clear revision requests to the artists. (a) The writers (a1) select attributes of a character and (a2) generate reference images. They can use (a3) image perturbation and (a4) fine-tuning to further tailor them. (b) The writers perform the image synthesis using (b1) the image drawn by the artists and (b2) the reference image to get (b3) the result image. (c) We-toon allows the writers to use pre-drawn images as references as well. (d) The generated images and pre-drawn images can be saved for later use.

communication can result in undesirable changes and additional revision iterations.

We conducted a formative study to understand: 1) how the authors collaborate (i.e., collaborative webtoon creation pipeline) and 2) which stages of the collaboration are most challenging to them. We had nine in-person interviews with four professional webtoon authors, each lasting for thirty to sixty minutes. From the interviews, we could understand how the writers and artists collaborate to create webtoons and summarize the result as a pipeline consisting of six main stages (Figure 2). Then, we identified two stages in the pipeline—1) character design and 2) draft revision—where maximum help is required according to the interviews. These two stages often contain numerous revision cycles, wherein the writers ask the artists to revise the images until they reach a consensus. Here, the writers convey their revision intents by sending reference images found on the internet along with text descriptions. For instance, a writer sends a set of reference images and specifies the exact part of each image—such as the shape of eyes or hairstyle—for the artist to refer to when revising. This indirect way of communication leads to miscommunications, because, in many cases, it is hard to elaborate the nuanced artistic feelings with keywords and commonly-found references.

The main contributions of our work are summarized as following:

- We performed a formative study with four professional authors to identify the collaborative webtoon creation process between the two roles, i.e., artists and writers;
- We implemented a GAN-based system, called *We-toon*, that allows image generation and synthesis for enhanced collaboration experience between the authors during sketch revision;
- We report a quantitative user study conducted with 24 professional webtoon authors conducted to investigate how webtoon authors could improve their communication during the sketch revision process using *We-toon*.

We-toon (Figure 1) aims to help the writers prepare a diverse set of reference images and synthesize them on the images drawn by artists to provide a more detailed and specific feedback. To ensure that we have the images best suited for webtoons, 236,449 actual webtoon character images are provided by Naver Webtoon¹. We manually filtered and preprocessed them into 47,233 head-cropped grayscale images (Figure 3). These preprocessed images were then

¹<https://www.webtoons.com>

used to train a state-of-the-art image generation model, StyleGAN2-Ada [12], which was used to generate entirely new sets of reference images on the fly. We also leveraged StyleMapGAN [15] to synthesize the reference images into any user-given image, allowing the writers to have a greater control over the references. In addition, when we design the user interface to deliver the feedback (i.e., revision request) to the artists, we take into account the artists' perspectives to help them clearly understand the writers' intentions. A revision request comprises of a textual description, brushed region, and a synthesized image that indicates "what", "where", and "how" to revise the image, respectively (Figure 4).

To evaluate our system, we conducted a user study with 24 webtoon authors (twelve writers and twelve artists) who had prior experiences in collaborative work. We gave them tasks to generate a character and draw a scene containing it (Figure 5). The experimental results confirmed that *We-toon* supports a more effective form of communication capable of relaying clearer requests for both the writers and artists. Moreover, the writers gave higher scores to the final revision in terms of all three metrics, i.e., appearance, harmony, and detail, when our system was involved in the process. The results indicate that *We-toon* improves the output quality of the refining process in practice. We further report the implications for AI-assisted communication support systems.

2 RELATED WORK

2.1 Systems to Support Multi-human Collaboration

Our goal is to promote an effective communication between two different roles (e.g., a writer and an artist). Therefore, we surveyed the systems to support multi-human collaboration reported in recent years. Such systems are crucial for human-computer interaction and computer supported cooperative works [17, 19, 23]. Thus, the studies on such systems are extensively spread out to varying tasks in different domains, including online discussion in writing documents [38], developing early-stage designs of three-dimensional (3D) objects [25] or user interfaces [29], and visualizing the collaborator's gaze in remote working situations [9]. The works exist within a variety of modalities and for different purposes, not bounded in virtual [20, 36] or physical [42] world to help people collaborate in their domains. For example, a digital chatbot was used for an effective task management [39], multiple sources of inputs (e.g., text, hand-drawn sketches, and photographs) for digital multimedia projects [43], and co-creative space prototypes for an exclusive learning experience [21].

A latest tendency is—as adopted by us in this study—to implement AI techniques to replace the human labor in an effective way and facilitate a better collaboration experience. Many researchers leveraged AI for diverse roles such as a collaborative teammate [31, 40], a mediator of teamwork and interpersonal communication [10, 18], and a social collaborator [22, 41]. In these studies, diverse tasks in multi-human collaborations were covered by helping users share ideas, inspire new insight, and organize thoughts. For example, Winder [16] linked audio tapes to visual objects to provide an effective communication between team members, who collaborated on visual documents in a remote and asynchronous

situation. MicroWriter [35] increased the productivity in collaborative writing by offering novel ideas and organizing them with labels. Cococo [33] revealed how AI can enrich, impede, or alter creative social dynamics during music composition tasks.

In summary, although extensive studies have been conducted in the field of systems in multi-human collaboration in varying domains, webtoon is a relatively new field that remains hitherto unexplored.

2.2 Image Generation, Editing, and Local Synthesis

There has been a growing interest in the generative adversarial networks (GANs), especially in handling image datasets, since their first emergence in 2014. The popularity and the image generation quality of GAN was dramatically increased with StyleGAN [13, 14], which achieved a more disentangled representation of the latent space with its novel architecture. Soon after, diverse methods to investigate the latent vector space were invented to manipulate the StyleGAN-generated images. For example, GANSpace [11] is an approach that applies PCA on the latent space, and uses the dominant eigenvectors for the image manipulation. Next, Shen et al. [32] proposed the closed-form factorization (SeFa), which is similar to GANSpace but the PCA was applied on the weight matrices of the affine transformation. The authors showed that SeFa enables more disentangled manipulation and is thus better for controlling a single semantic attribute compared to GANSpace.

Image synthesis is one of the primary applications of GANs. Suzuki et al. [34] proposed a spatial conditional batch normalization to collage local areas from multiple images into one source image. By interpolating between the source and the collage images, they showed that the areas to change semantically. Although there were a lot of research to tackle the same task [3, 5, 8], one critical defect shared among them was that they cannot handle the real images not included in the training set. Recently, based on StyleGAN, Kim et al. introduced a modified architecture called StyleMapGAN [15] that can edit a source image by locally transplanting the style from a reference image. With a simple brush interaction, the user can transplant local regions from diverse references and manipulate the semantic extent. Unlike the methods reported in previous studies, the performance of StyleMapGAN, when applied to the real images, resulted in a competitive quality.

In our study, we combined the best from the above methods. We used StyleGAN2-Ada [12] and SeFa to allow the writers to generate and manipulate the reference images, respectively, instead of having to search for them on the web. Moreover, we leveraged StyleMapGAN to help the writers synthesize local styles and revise specific parts of the images to produce visual feedback that can be more intuitive and clear to the artists than simple textual feedback.

3 FORMATIVE STUDY

To understand how webtoon authors collaborate, we conducted semi-structured interviews with four professional webtoon authors each of whom had approximately one, two, five, and ten years of experiences in publishing webtoons on global platforms (e.g., NAVER Webtoon Corp. that holds more than 167 million monthly active users in the global market across over 100 countries). All the

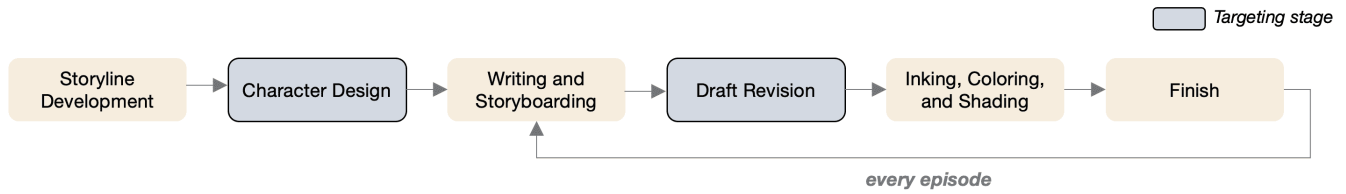


Figure 2: Collaborative webtoon creation pipeline, which starts with the storyline development and character design. The subsequent processes are repeated every week for serial publications.

authors had published at least one professional webtoon done in collaboration with at least one other author. We conducted thirty- to sixty-minute long semi-structured in-person interviews with each author. The interview questions mainly focused on three aspects: (1) how the collaborative webtoon creation proceeds in general; (2) which parts of the collaboration generally have the lowest communication effectiveness; and (3) how they handle the possible miscommunication in these stages. By analyzing interview results, we identified a general workflow of the collaborative webtoon creation process and established our system’s design objectives.

3.1 Collaborative Webtoon Creation Pipeline

Based on the data acquired from the interviews, we defined a collaborative webtoon creation pipeline (Figure 2) as follows. The pipeline consists of six main stages: (1) storyline development, (2) character design, (3) writing and storyboarding, (4) draft revision, (5) inking, coloring, and shading, and (6) finish. The first two stages (Stages (1) and (2)) are completed before releasing the first episode, whereas the latter stages (Stages (3) to (6)) proceed regularly, e.g., weekly, according to the publication schedule.

(1) Storyline Development. As a kick-off of the creation process, the writers develop a draft of the storyline; for example, they determine the genre, the background, the number of main characters, and what will happen to the characters. The ending of the webtoon may not be actually determined at this stage, since in many cases, the story remains open-ended and is elaborated later. This stage is usually executed solely by the writers without the cooperation of the artists.

(2) Character Design. After the storyline is developed, the writers and artists collaborate to design the characters. Firstly, the writers provide a textual description of a character, including the hair style, age, first impressions, and so on, to the artists. For the main characters, more internal and subtle details, such as background or quirks, may be set. Next, the writers search for a set of reference images in Google Images or Pinterest², which can be referred to by the artists during the design process. The quantity and quality of the reference images are both important because the images should give a comprehensive visual example of a character, and a single image is often not sufficient to convey what the writers imagined clearly. When the writers find a sufficient number of reference images, they send them to the artists with textual descriptions for further details.

Based on the reference images and textual descriptions, the artists start to draw a sketch and send it back to the writers for

confirmation. According to our interviews, the first draft rarely meets the writers’ expectations. To request a revision, the writers send more reference images and textual descriptions. After multiple back-and-forth processes, if all the authors consent to the sketch, then the artists draw a model sheet, i.e., a set of drawings of a single character depicting diverse facial expressions and poses. The model sheet is then used for future reference to keep the character design consistent throughout the publication.

(3) Writing and Storyboarding. In this stage, the writers formulate the details of a single serial publication that describe each scene with text to aid the artists in understanding the given situation. Once the detailed descriptions are provided, the artists create a storyboard, which includes, for example, the structure of each cut, word balloons, lines or dialogues, and the composition of the characters, objects, and scenery.

(4) Draft Revision. In this step, the authors collaborate to determine the detailed poses and facial expressions of the characters. Based on the storyboard, multiple iterations are performed, if needed, to identify the best drawing of each scene that suits the storyboard. According to the interviews, the maximum communication between the authors occurs at this stage, as they can have different mental models of the same scene. If the collaboration goes well, then the rough sketches are polished to a partially complete state.

(5) Inking, Coloring, and Shading. Inking (i.e., cleaning-up process of the sketches), coloring, and shading are performed sequentially by the artists. Although revising a cut is still possible in this stage, it requires a significantly longer time, which makes the modification nearly impossible in practice given the short publishing schedules.

(6) Finish. In the last stage, the authors complete the details before uploading the episode. This stage includes background work, placing and completing the speech balloons, and spell-checking. For the background work, many webtoon authors retrieve a computer-rendered image using a 3D modeling software, such as SketchUp³, to draw buildings, landscape architectures, or interior design. When the background is completed, the artists settle the speech balloons at their appropriate positions considering the characters and background. Unless there are any spelling or drawing errors that do not align with the comprehensive composition, the artists send the episode to the platform editors for publishing on the web.

²<https://www.pinterest.com/>

³<https://www.sketchup.com/>

3.2 Design Objectives

Based on the collaborative webtoon creation pipeline, we identified two stages with the least communication effectiveness, viz. the character design and draft revision, and we set two design objectives (DO1 and DO2) to inform the design of *We-toon* for more effective communication in these stages.

DO1—Support the writers in retrieving the desired reference images: Finding the appropriate reference images is an essential step in character design and draft revision. Traditionally, the writers resorted to image search engines, such as Google or Pinterest, to obtain the references. However, this search process often provides images with wrong details, such as wrong hair length or head pose (e.g., yaw). This process can be streamlined by generating an image with user-defined details.

DO2—Support the writers in expressing their revision intent: The lack of drawing skills constrains the writers to use textual descriptions, which often results in ambiguous revision requests that confuse the artists. Synthesizing a reference image into the image drawn by the artists is a feasible approach to clarify the meaning of the revision requests even for those who do not have professional drawing skills.

4 SYSTEM: WE-TOON

To meet the design objectives indicated in Section 3, *We-toon* helps the writers modify an image given by the artists (i.e., the source image) by leveraging specific details from the reference images in two phases: **image preparation** and **image synthesis**. Hereafter, we use the term “**source image**” to refer to the image drawn by the artists and “**reference image**” to refer to the image generated by the writers. In the **image preparation** phase, the writers look for a reference image that would be synthesized into the source image (DO1). They can choose one from a collection of pre-drawn images (i.e., popular webtoon characters obtained from the web) in a classical but inflexible way, or the writers can browse the reference images dynamically generated by a GAN model, further enriching the search space. In the **image synthesis** phase, the writers can mix local regions of a reference image into the source image (DO2). To transfer the image details naturally, we built another GAN model using a webtoon image corpus.

4.1 Preparing Reference Images

To obtain reference images that will be mixed with the source image, the writers can either (1) generate an image or (2) choose one from the collection of pre-drawn images.

Generating Reference Images. For reference image generation, we built a GAN model (StyleGAN2-Ada) that generates an image using a 512-dimensional latent vector. Therefore, theoretically, there can be an infinite number of generated images. To help the writers explore such a large corpus of images efficiently with serendipity [4, 24, 37], we designed three dynamic querying interactions—**attribute filtering**, **two-level perturbation**, and **fine-tuning**—to refine the search scope successively.

As the first step of dynamic querying, the writers can choose from combinations of attributes that determine the overall appearance of the character in the reference images. For simplicity, we

support three binary attributes: gender (man/woman), hair length (short/long), and eye size (big/small) (Figure 1-a1).

Once the attributes of the character are chosen, the GAN model randomly generates twelve images corresponding to the combination (Figure 1-a2). Furthermore, we included different amounts of perturbation on the latent vector for each image to ensure that they have distinct styles. Clicking on an image designates it as a seed to generate twelve more images that are slightly modified from it, using small amounts of perturbation (Figure 1-a3). These variants tend to have differences in details such as facial expression, hairstyle, or gaze direction.

In addition to generating images using perturbation, fine-tuning is also provided for the writers to manipulate certain details, such as the head pose, directly. Specifically, each of the five dominant principal components in the latent space can be adjusted, via five sliders in the interface (Figure 1-a4), to change the latent representation of the image, resulting in an alteration of the key details, such as head pose or hair length [32]. Since the exact nature and extent of the manipulation are dependent on the latent space and are not pre-determined, two image labels are shown above and below each slider, suggesting “how” the reference image would change when moving the sliders.

Choosing from Pre-drawn Images. *We-toon* supports the use of pre-drawn images from our webtoon image corpus as the references (Figure 1-c). Notably, these pre-drawn images cannot be subjected to manipulations through fine-tuning as they are not generated by our GAN model.

Library. In our pilot tests, we found that the writers want to save the reference images (generated or pre-drawn) for future reference. The library serves as a temporary storage for such images, and the writers can add any images by double-clicking on them (Figure 1-d). The generated images are highlighted with blue borders, whereas pre-drawn images are marked in orange color.

4.2 Synthesizing Source and Reference Images

Once a reference image is selected during the preparation phase, the writers can synthesize the local regions of the reference image into the source image (i.e., the image drawn by the artist, Figure 1-b1) to revise the source image. Initially, the writers need to select the region on the reference image (Figure 1-b2) to mix into the source image using the brush interaction. The same region is automatically highlighted on the source image at the same time; however, if the two images are not aligned, then the writers can drag each region individually to align them. If multiple regions need to be mixed into the source image, then the writers can create different layers, each having one brush selection, to indicate the areas of interest (Figure 1-b4). The writers can add an additional text description (Figure 1-b5) to explain their intent in words. Next, the writers can export the resulting image (Figure 1-b3) with text as well as the brush-painted images (i.e., a revision request) by pressing the “export” button (Figure 1-b6).

4.3 Training GANs

We-toon employs two GAN models, viz. 1) StyleGAN2-Ada [12] for the image generation and perturbation, and 2) StyleMapGAN [15]

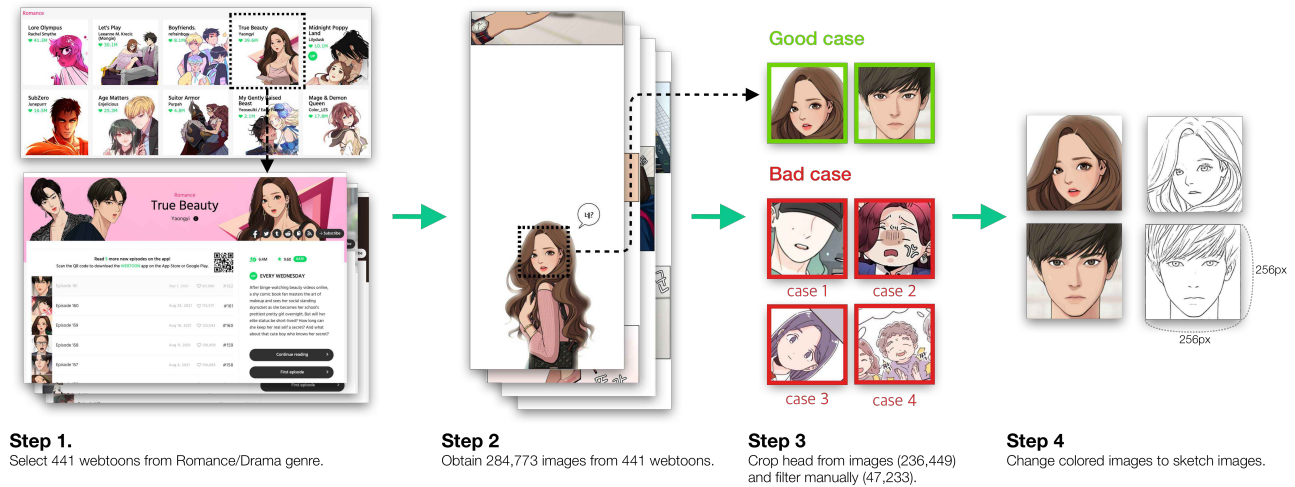


Figure 3: Webtoon dataset collection and preprocessing pipeline. We first select 441 romance and drama webtoons from the list (Step 1). Then we obtain 284,773 webtoon cut images (Step 2) from which we crop 236,449 head-cropped images and manually filter 47,233 images (Step 3). Lastly we change them to grey-scale sketch images (Step 4).

for synthesizing the source and reference images. In addition, *We-toon* applies one algorithm, viz. SeFa [32] for fine-tuning (i.e., the five sliders for fine-tuning an image).

Dataset. To train the models, we built a webtoon image corpus consisting of 47,233 256-by-256 face images. We did not use existing cartoon datasets [6, 7] since they mostly consisted of images from the Japanese cartoons (i.e., manga) that have different visual characteristics from the webtoons.

To create a high-quality face dataset, we selected 441 romance or drama webtoons from Naver Webtoon, which is one of the largest webtoon service platforms in the world. From the raw images, we extracted the face images using Anime Face Detector⁴, which is developed based on a Faster-RCNN model [27]. After filtering out the non-face images using the face detector, the dataset was further refined manually, during which an image was discarded if it had one of the following problems (Figure 3-Step 3):

- Any part of the face is obscured by other objects, such as hats, jewelry, or other body parts.
- Certain parts of the face are overly distorted.
- The face is tilted by more than 45 degrees.
- The character is too young or too old.

To extract edges from fully-colored webtoon images and transform them to sketch images, we used a pre-trained model of LineDistiller⁵ (Figure 3-Step 4).

Model Details. StyleGAN2-Ada uses a 512-dimensional vector sampled from the normal distribution as a latent vector z . For the image generation, we sampled approximately 2,000 latent vectors, and labeled each to classify the combination of attributes (e.g., male/short hair/small eyes). We varied the noise level by multiplying z by a small constant (0 to 0.8).

The fine-tuning feature could use a total of 512 different vectors obtained through SeFa. However, we only employed the five

Table 1: Information of the webtoon dataset.

| | Romance | Drama | Total |
|--------------------------|---------|--------|---------|
| # of webtoons | 227 | 214 | 441 |
| # of obtained images | 208,221 | 76,552 | 284,773 |
| # of head cropped images | 183,459 | 52,990 | 236,449 |
| # of filtered images | 30,915 | 16,318 | 47,233 |

dominant principal components as they induce the largest semantic changes [32]. Each principal component vector is added to the latent vector z after multiplying it by the value corresponding to the slider location from -10 to 10. The StyleGAN-based generative models (e.g., StyleMapGAN [15]) exhibit a trade-off between distortion (e.g., identity preservation) and editability (e.g., semantic interpolation), depending on the resolution of the latent space (e.g., feature map) [28]. If the facial identity of the drawn image changes during the editing process, then a miscommunication may arise between the writers and the artists. To overcome this limitation, we used a 32×32 resolution feature map that guarantees identity preservation, although it has lower editability compared to the 8×8 or 16×16 resolution feature maps [15]. We trained the models from scratch with our own dataset using 4 Nvidia TITAN RTX GPUs.

5 EVALUATION

We conducted a user study with 24 professional authors (twelve writers and twelve artists) to answer the following three questions:

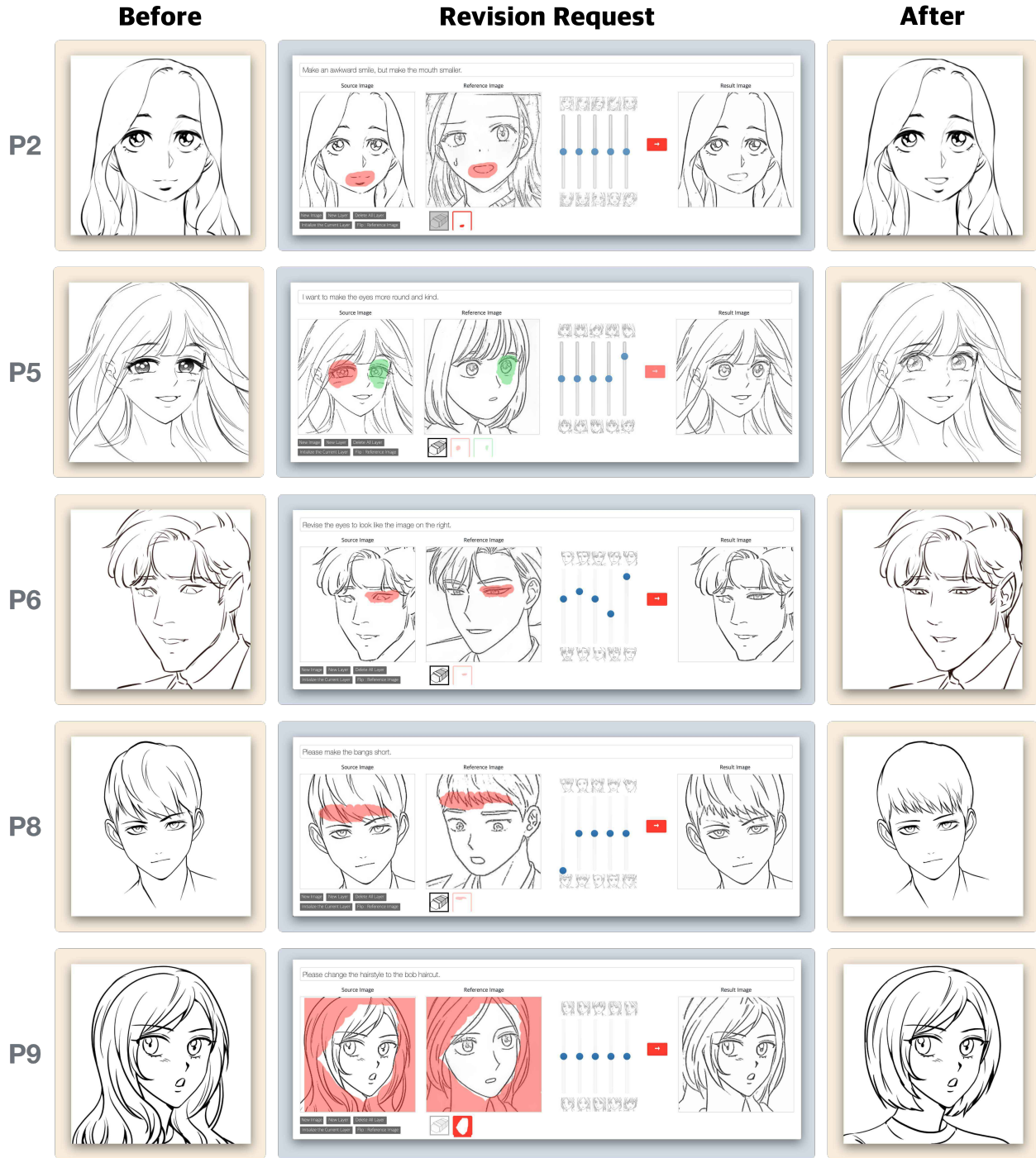
(Q1) Can the writers find or generate desired reference images? (DO1)

(Q2) Can the writers revise the images that are obtained from the artists using reference images? (DO2)

(Q3) Ultimately, can the authors communicate effectively during the sketch revisions using *We-toon*? (DO1 and DO2)

⁴<https://github.com/qhgz2013/anime-face-detector>

⁵<https://github.com/hepesu/LineDistiller>



The **artist** sends a draft.

The **writer** responds with a revision request.

The **artist** draws new image with appropriate modifications and resends it.

Figure 4: Five examples of revision process using We-toon in our user study. First, artists draw sketch image and sends it to writers (i.e., Before). Next, the writers modify the image to produce a comprehensive feedback (i.e., Revision Request), and sends it to the artists. The artists refer to the revision request and makes a proper modification (i.e., After).

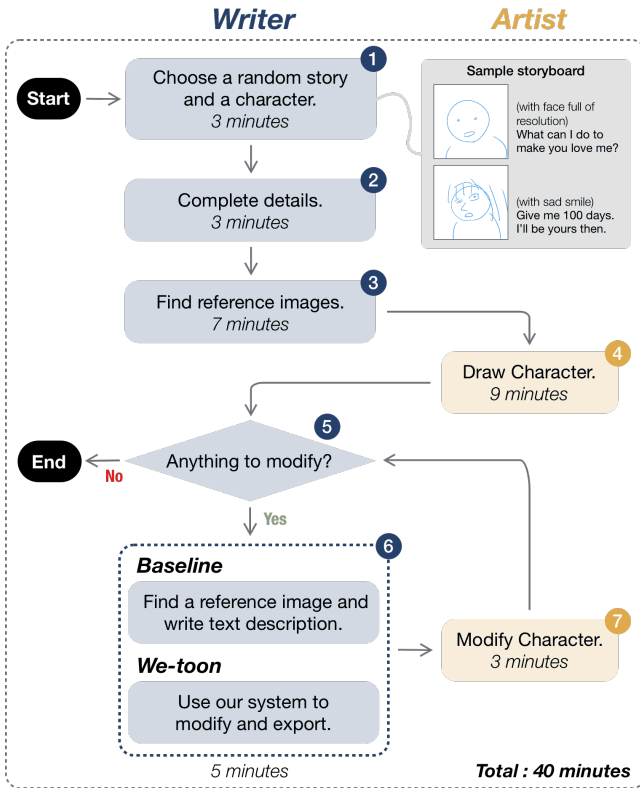


Figure 5: Task Design. (1) The writer selects a random story and a character. They are provided with a corresponding sample storyboard. (2) They complete the details of the character (e.g., personality and background) and (3) find reference images using search engines like Pinterest. (4) Once the details and reference images are given to the artist, they draw a character based on the information. (5) The writer could ask a revision until there is nothing to modify. (6) When asking a revision, in the baseline condition, the writer finds reference images using search engines to convey their revision intention, but in the *We-toon* condition, they use our system.

5.1 Study Design

We designed a comparative study to understand if and how *We-toon* facilitates effective communication between the authors during the collaborative sketch revision process. The study design included one independent variable, Interface (**baseline** and *We-toon*). In both the conditions, the writers searched for three reference images using the existing search engines (Google and Pinterest) to request the artists to draw an initial sketch image. For the subsequent revision cycles, the writers were asked to continue using the search engines (Google or Pinterest) to find reference images for communication with artists, in the **baseline condition**. In contrast, in the *We-toon condition*, they could generate and synthesize the reference images on their own using our system. We used a within-subjects design, i.e., writers experienced both the interface conditions. We

counterbalanced the order of the interface conditions to prevent ordering effects.

Participants. We recruited 24 professional authors (twelve writers and twelve artists) who had prior experiences in collaborative work. They were recruited through online webtoon communities (e.g., social media platform) and compensated approximately \$43.00 for a two-hour-long study. We randomly paired the writers and the artists, resulting in twelve pairs of authors, and each pair of authors performed the designated tasks independently from the other pairs.

Task. Seeking external validity, we designed a collaborative character design task where the writer-artist pair collaborated to sketch a single webtoon scene (Figure 5). Considering that all the authors we interviewed in the formative study did collaborative works remotely, the two authors were physically isolated from each other at different locations and they communicated only using a commercial messenger, reflecting the real-life conditions. We kept the authors' identities anonymous to each other to avoid any potential bias.

The task consisted of multiple sub-tasks, each done by either the writer or the artist (Figure 5). The task started with a character design step (Figure 5-(1)). The writer was given a randomly selected storyboard (among two) containing two scenes: one for a male character and the other for a female character. The writer was asked to choose one of them to design. We provided descriptions of the characters, such as name, age, and occupation, and background, to the writers. Owing to time constraints, the writers were instructed to start from preset stories instead of freely writing their own stories.

After choosing a character to design, the writer set up detailed attributes of the character, such as personality, background (e.g., a personal trauma like the death of a child), and appearance (i.e., long hair). The writer was asked to articulate each attribute in one or two sentences that were delivered to the artist (Figure 5-(2)).

Next, the writer searched three reference images on Google or Pinterest that resembled the character in mind well (Figure 5-(3)). During this stage, the writers, even under the *We-toon* condition, were instructed to use the existing search engines as the initial reference images could be pictures of real people or illustrations. We set the number of reference images to three following the actual practices in the field.

Once the writer decided on the reference images, the images and the descriptions about the character were sent to the artist via a messenger. Then, based on this information, the artist drew a sketch and sent it back to the writer (Figure 5-(4)).

In the formative study, we observed that most writers tend to request revisions using only a single reference image. The reason being, not only is it hard to find appropriate references, but sending multiple references for revisions can result in confusion and miscommunication as the artists become unclear of what to focus on. Therefore, given the initial sketch from the artist, we allowed the writer to only send a single additional reference image to request a revision on specific aspects of the character (Figure 5-(5)). The selection of this additional reference image differed according to the interface condition (Figure 5-(6)). In the baseline condition, the writer was asked to search a reference image using the existing search engines, whereas in the *We-toon* condition, the writer used *We-toon* to modify the sketch directly. Then, the

reference image, along with a simple explanation of the revision, was sent to the artist.

Based on the revision request, the artist revised the sketch and sent it back to the writer (i.e., an iteration, Figure 5-(7)). The authors were not allowed to send extra images or messages other than the ones we mentioned during the study to control the number of messages for communication. Therefore, even if the writer’s revision request was ambiguous, the artist could not ask the writer for clarifications; instead, they were instructed to report such cases to us. The revision process was repeated for up to 40 minutes until the writer was satisfied with the sketch.

Apparatus. The writers were provided with a laptop (Macbook Pro 2020) and a mouse. The artists were provided with a tablet (12.9-inch iPad Pro) and a digital pen (Apple Pencil); however, they could use their own devices, if desired. The artists used Clip Studio Paint⁶, which is one of the most popular programs used by webtoon artists, for drawing the sketches.

Procedure. After signing a consent form, the participants were briefly introduced to the overall procedure and task for three minutes. Then, the writer watched a 10-minute tutorial video about *We-toon*. The authors performed the collaborative character design task twice: once under the baseline condition, and next, using *We-toon*. Each task took up to 40 minutes, and we randomized the order between the conditions. The participants had a 10-minute break between the two tasks. Before performing the task using *We-toon*, the writer could freely use the system for 5 minutes to practice. The duration of the entire session was approximately 120 minutes.

Interview. We conducted surveys and semi-structured interviews (before the experiment and after finishing each condition) to assimilate the participants’ experiences about the collaboration task as well as the feedback on our system (Table 2). To quantify the writers’ satisfaction with the revised image, we used the following evaluation criteria: (1) *Appearance*: how well the character’s external conditions were reflected in the final image (e.g., age); (2) *Harmony*: how well the character in the final image was in harmony with the given scene; (3) *Detail*: how well the character’s attributes, which were set up by the writers, were reflected in the final image. The writers answered each criterion using a nine-point Likert scale.

5.2 Quantitative Results

In this Section, we refer to the participants with an abbreviated form. For example, we represent the first pair as P1; the writer of the first pair is denoted as W1, while the artist is denoted as A1. We used a significance level (α) of 0.05 for all the statistical tests.

We used the non-parametric Wilcoxon signed-rank test to analyze our data. For Q1, all the writers, except W3, agreed that *We-toon* can produce appropriate reference images ($M=7.0$, $SD=1.6$) (Table 2: Q1). Similarly, eight among the twelve writers agreed that the system could synthesize sample images properly ($M=6.6$, $SD=1.4$) (Table 2: Q2). A statistically significant difference was observed as ten among the twelve writers reported that *We-toon* conveys the revision intention more intuitively and effectively ($M=7.5$, $SD=1.7$), compared to the baseline ($M=5.9$, $SD=2.1$) ($z = 2.46$, $p < .05$) (Table 2: Q3a). For Q3b, the writers were satisfied with the final image produced using *We-toon*, compared to the baseline, in terms of all

the three metrics (e.g., appearance, harmony, and detail) (Table 2: Q3b). While the baseline scored 6.3, 6.7, and 6.6, *We-toon* exhibited 7.5, 7.2, and 7.3 on an average. There was a statistically significant difference in appearance ($z = 2.00$, $p < .05$), but not in harmony and detail. The artists also claimed that *We-toon* aided in understanding the revision intention more clearly and accurately ($M=7.8$, $SD=1.1$), compared to the baseline ($M=6.3$, $SD=2.2$), showing a statistically significant difference ($z = 2.08$, $p < .05$) (Table 2: Q3c). The average number of revision requests and the total time needed for the revision were both lower with *We-toon* condition. Specifically, *We-toon* condition required 4.6 minutes with 1.6 revisions on average, while the baseline condition did 6.2 minutes with 1.8 revisions on average (see Table 4). In summary, the writers could convey the revision intentions more intuitively and effectively using *We-toon*, and the artists could understand the intentions more clearly and accurately when *We-toon* was used. In addition, the overall appearance of the final image was better when *We-toon* was involved.

5.3 Qualitative Results

After each task, we had a semi-structured interview to further understand the authors’ experiences. We requested the detailed reasoning behind their answers on each survey question (Table 2). We also asked which one was the most useful in the revision cycles. The writers had six functions (i.e., image generation, image perturbation, pre-drawn images, fine-tuning, image synthesis, and revision request) and the artists had three parts (i.e., text description, local brushed region, and synthesized image) to rank. In this section, we summarize the prevalent responses to these interview questions.

The writers preferred the GAN-generated images over the pre-drawn images for the references. In our experiment, we found that most writers (11 among 12) indicated that image generation was much useful than the pre-drawn images. We asked the writers to rank the functions of *We-toon* in order of usefulness. Seven writers ranked the image generation as the most useful function, and five writers placed the pre-drawn images at the bottom (Table 3). Since the evaluation for both the functions were positioned on the extreme opposites, we interviewed each author to perform a deeper analysis of their responses.

We identified three reasons for the writers’ preference in using the generated images as references. Firstly, they found the attribute filtering function useful. For example, W10 said, “*It was hard for me to find the styles of characters with the pre-drawn images, but the detailed categorization in random image generation helped me to reduce the efforts required to find proper references*”. The attribute filtering required less effort and time by the writers to find their desired reference images, instead of searching through a large number of images.

Secondly, the images generated in real-time were diverse and of sufficiently high quality to be used as reference images without further work from the writers. For example, W1 said “*I was able to find diverse reference images that were generated instantly and constantly with high fidelity*”. Since most writers (11 among 12) answered that *We-toon* could generate diverse images, which they could utilize as references (Table 2), we can conclude that using only

⁶<https://www.clipstudio.net/>

Table 2: Quantitative results on our questions: 1) image generation, 2) image synthesis, and 3) the satisfactory level of the task outcome and communication within revision cycles. * indicates a statistically significant difference with $p < .05$.

| Interviewee | Question | Baseline comparison | Method | Score | | |
|-------------|----------|---|-----------------------|-----------------------|------------|-----|
| | | | | M | SD | |
| Writers | Q1 | Produce reference images | No | <i>We-toon (Ours)</i> | 7.0 | 1.6 |
| | Q2 | Synthesize sample images | No | <i>We-toon (Ours)</i> | 6.6 | 1.4 |
| | Q3a | *Convey intention intuitively and effectively | Yes | Baseline | 5.9 | 2.1 |
| | | | | <i>We-toon (Ours)</i> | 7.5 | 1.7 |
| | Q3b | *Appearance | Yes | Baseline | 6.3 | 1.7 |
| | | | | <i>We-toon (Ours)</i> | 7.5 | 1.9 |
| | | | | Baseline | 6.7 | 2.1 |
| | Q3b | Harmony | Yes | <i>We-toon (Ours)</i> | 7.2 | 2.3 |
| | | | | Baseline | 6.6 | 1.7 |
| Q3b | Detail | Yes | <i>We-toon (Ours)</i> | 7.3 | 1.8 | |
| | | | Baseline | 6.3 | 2.2 | |
| Artists | Q3c | *Understand intention clearly and accurately | Yes | <i>We-toon (Ours)</i> | 7.8 | 1.1 |

Table 3: Ranks for each function by the writers. The number in each cell refers how many people gave that rank for each function. Functions are sorted in a manner that the one gets more higher ranks located in the leftmost column.

| Rank | Image Generation | Fine-tuning | Image Synthesis | Pre-drawn Image | Image Perturbation | Revision Request |
|-----------|------------------|-------------|-----------------|-----------------|--------------------|------------------|
| 1 (Best) | 7 | 3 | 2 | 0 | 0 | 0 |
| 2 | 3 | 3 | 2 | 2 | 1 | 1 |
| 3 | 1 | 3 | 4 | 0 | 3 | 1 |
| 4 | 0 | 2 | 3 | 2 | 4 | 1 |
| 5 | 1 | 1 | 1 | 3 | 2 | 4 |
| 6 (Worst) | 0 | 0 | 0 | 5 | 2 | 5 |

GAN-generated images could be enough for the whole revision process.

Lastly, the writers were worried about plagiarism arising from using the pre-drawn images, while not a single author voiced similar concerns about using the GAN-generated images. Specifically, two writers (W11, W12) expressed their concerns about using the pre-drawn images even for communication purposes, while acknowledging that using the generated images are free from such issues. W12 said, "Webtoon authors are very sensitive about the plagiarism issue as it can destroy their professional career at once. I know it is only for communication purpose, but I was a little apprehensive in using the pre-drawn images just in case. However, I do not think there is such an issue when it comes to generated images".

The writers found fine-tuning was useful. Many writers were satisfied using fine-tuning function of *We-toon*; 6 among the 12 writers evaluated it as the most or the second most useful function (Table 3). We could summarize the writers' preference with several reasons. First, *We-toon* can convey revision intentions on the poses of the characters. Frequently, it is hard to find reference images that have the desired pose using image search engines (e.g.,

Pinterest). However, since our fine-tuning can help find diverse yaws and pitches through a simple interaction, writers can create the poses they want with little effort. W1 said, "Since I could generate and manipulate the image into diverse poses, it was very practical and took much less time than using a search engine," indicating that *We-toon* offers convenient pose editing. Likewise, this logic can be applied to finding specific facial expressions as well. For example, when directing a scene, authors sometimes require a character to show a mixed feeling on their faces (e.g., a sad smile). W2 said, "I believe *We-toon* can also support directing a scene in collaborative creation, as it helps me to find diverse poses and facial expressions which I can use for the communication," which confirms that *We-toon* can handle diverse real-world cases. Lastly, it provides a visual form of rough sketches depicting the writers' imagination. For instance, W3 said, "Sometimes I do not know how to describe, in words, the image I imagine; however, the fine-tuning helped me visualize the images by creating with different styles dynamically".

On the other hand, the writers were less likely to use image perturbation, compared to fine-tuning. Since the image perturbation produced a random change without specific guidance on the

Table 4: The total and average revision time taken by the artists (minutes).

| Method | | <i>M</i> | <i>SD</i> |
|-----------------------|--------------------|------------|-----------|
| Baseline | Total time (min) | 6.2 | 2.6 |
| | Revision # | 1.8 | 0.6 |
| | Average time (min) | 3.6 | 1.4 |
| <i>We-toon (Ours)</i> | Total time (min) | 4.6 | 2.2 |
| | Revision # | 1.6 | 1.0 |
| | Average time (min) | 3.1 | 1.1 |

semantics, the writers could not expect what was going to change, thus resulting in an infrequent usage.

The writers could take advantage of the image synthesis.

As shown in Table 2, all the participants asserted that *We-toon* could properly support the image synthesis task on average (W2, W3, W5, W6, W7, W10, W11, W12) or at least help to some degree (W1, W4, W8, W9). Two writers (W7, W11) ranked the image synthesis as the most useful function, and two writers (W2, W4) ranked it as the second most useful function (Table 3).

The writers were pleased to use the image synthesis, because it enabled easy image revisions for those who do not have proficient drawing skills. W7 stated, "*We-toon can be a wonderful solution to those who are not experienced in visualizing their thoughts on characters, since it is important to express an abstract concept into an obvious visual form,*" implying that our system can be used by writers with different levels of proficiency in drawing.

The artists found the revision requests generated by *We-toon* effective and convenient. Since *We-toon* provides a comprehensive feedback consisting of a combination of 1) text description, 2) local brushed region, and 3) the synthesized image, the artists could clearly understand "what" to fix, "where" to fix, and "how" to fix at once. A1 stated, "*Using We-toon in a revision cycle was intuitive and direct. The brushed regions let me recognize where to fix, and the modified image was expressive enough to give a clear idea of how to revise the original image,*" appraising the integrated feedback from *We-toon*. We further asked artists to rank the three parts of the feedback. Eight among the twelve artists (A2, A4, A5, A6, A7, A9, A10, A11) ranked the synthesized image as the most useful among the three (Table 5). They liked it because it enabled a visual communication for revision with a single image. A8 said, "*The writers often send a bunch of images including celebrities and webtoon characters, but We-toon uses a single image that contains the writer's intention for a revision which makes the process simple and concise,*" implying that *We-toon* improved the way artists had been working during the revision process. As the feedback of *We-toon* supports an example-based communication with a single image, the artists could understand the revision intention more effectively.

6 DISCUSSION

In this section, we discuss the implications of using AI-assisted communication support systems, and report our plans for future work as well as the limitations of the study.

Table 5: Ranks for each part by the artists. The number in each cell refers how many people gave that rank for each part.

| Rank | Synthesized Image | Local Brushed Region | Text Description |
|-----------|-------------------|----------------------|------------------|
| 1 (Best) | 8 | 2 | 2 |
| 2 | 1 | 6 | 5 |
| 3 (Worst) | 3 | 4 | 5 |

6.1 Implications of AI-assisted Communication Support Systems

Clarity first, quality next. Previous studies were focused on adopting the AI to outperform its baseline in terms of the output quality [1]. However, our experimental results demonstrate that AI can work at its best when the clarity matters more than the quality, as it can prevent the miscommunication via a low-cost short-duration process. Up until now, even the state-of-the-art AI may generate only partially complete images than the real artworks. However, we think that it can still support the creators in diverse domains to communicate effectively. *We-toon* provides functions such as GAN-based **image generation** (Figure 1-a2) and **fine-tuning** (Figure 1-a4) so that the writers can actualize their visions via simple interactions. The output reference images were partially complete with a relatively low quality; however, the writers liked them more than the higher-quality **pre-drawn images** (Figure 1-c) (Table 3). We could elicit such a satisfaction by concentrating on the prior objective (i.e., clarity).

Specifically, some of the revision requests were extremely ambiguous when conveyed with text, e.g., W2 wanted a more petite feeling, W3 asked a post-adolescent feel, and W11 requested a carefree vibe. However, *We-toon* enabled a more straightforward revision by providing an image synthesis result. A11 said, "*Using the images of celebrities often made the revision abstract and dubious, but We-toon enabled a clearer way to request a revision,*" which indicates that the incomplete drawings generated by *We-toon* allows users to communicate more clearly. In this process, quality was a less significant factor, as the main aim was to visualize the writers' intended revision quickly. By providing drawing abilities to the writers, both the authors could bridge their differences in thoughts, thereby making the revision process clearer and faster. W7 said "*What we need is a reference containing our revision intention, but not a high-quality image. We-toon enabled it with a simple interaction in a short time.*"

Providing predictable mapping helps the users greatly. In our study, the writers found attribute filtering helpful, because it provided control over the inherently random image generation. This enabled the writers to restrict the browsing space according to their requirements, thereby lowering the cost of finding the appropriate references. However, the users wanted even more control and predictability by asking for more attributes that are well suited to the target domain's recent trend and preferences. W3 said, "*Although I know the references exist for communication purposes, I did not want to choose one that feels old-fashioned,*" which indicates

that the users wanted a detailed categorization that matches the attributes prioritized by the users in that domain.

Similarly, to provide predictable mapping, we added two image labels above and below each slider in **fine-tuning** function (Figure 1-a4) to preview how the target image would change by moving the sliders. Most of the writers enjoyed using it because they could match their expectation on the manipulated image in advance. In contrast, we did not put any semantic label on the random changes occurred in the **image perturbation** function (Figure 1-a3), which many writers felt uncomfortable and were unwilling to use. In short, when a predictable response from the AI model is provided, the users believe the model is under control and can fully exploit its functions as they wish as per their desires.

AI can be used to circumvent around plagiarism concerns.

Using references is often necessary in communication between collaborators as requesting a revision without them can seriously hinder the speed and quality of the conversation. However, searching and referring to existing works can risk the authors being subconsciously influenced by them, violating copyright issues, and unintentionally taking advantage of others' intellectual properties.

The result presented in Section 5.3, i.e., the authors did not worry about plagiarism issue when using the GAN-generated images, is interesting as it contradicts the previous claim [26], where the authors suggested that people have a negative bias toward AI-generated images; they prefer human artworks to AI-generated ones. As shown in our evaluation, we presumed that people do not have perception bias when they are used for communication purposes. Therefore, we suggest that utilizing AI for generating references in the future can be a viable solution to circumvent the copyright issue. However, since the plagiarism of using AI-generated contents has not been discussed extensively, the issue is still an open question. We concede that it is debatable, and there is a room for extra discussion in the future.

AI can help start and mediate communications. Through the interviews, we found that some writers (W4, W5) felt pressurized while directly altering the images drawn by the artists. It is often an implicit rule not to step on or disrespect the other author's area of expertise. However, we found that using *We-toon* mitigate this gap, since the writer can shift the responsibility to the AI, which can induce an active communication. W4 said, "*There are some artists that do not let others touch on their drawings. Using We-toon can mediate this issue, as I do not touch the image with my own hands but the system concocts it independently,*" which indicates that the writers can reduce the burden of editing images by shifting the responsibility to the system.

Moreover, in our experiment, many participants indicated that *We-toon* seems to be useful for producers, who work in between the writer and the artist to mediate their relationship to ensure a successful overall production (W3, W6, and W8). W3 said, "*I also work as a producer who manage the ideas and relationships between the writer and artist, and I always had difficulties mediating the miscommunication between them. I think not only the writer-artist pair but the writer-producer pair can also use We-toon for an effective communication,*" which implies that utilizing AI can supplement as well as the work of producers, allowing the creators to initiate communications more easily. In the future, if properly addressed, we think AI can be utilized as an excellent proxy for

direct communication between the writers and the artists, as it is currently done by the webtoon agencies or managements.

6.2 Limitations and Future Work

Although *We-toon* supports an effective communication between authors as well as diverse real-world cases of sketch revisions, certain limitations exist, which will be addressed in a future study.

Two authors commented that they wished to modify the whole image, not just small parts, but it was hard to utilize *We-toon* as it does not support direct semantic manipulation of the source images. For example, W8 wanted to make the character look a bit older, and even after requesting revisions three times through *We-toon*, she gave a very low appearance score of 3 to the final image. We think that manipulating global semantics of images using state-of-the-art image manipulation techniques such as HyperStyle [2] would be of help in this regard.

We-toon only focuses on close-up shots of faces and a fixed set of characters and expressions. When developing *We-toon*, we focused on primarily romance/drama genres, which account for the largest proportion (46%) of works on the Naver Webtoon platform. In the formative study, we found that the close-up shots are the most common scene composition and the authors put more emphasis on the facial expressions than other features in these genres. Thus, we trained the GAN models on the things authors cared the most (i.e., faces and hair styles) in the domain. That being said, it is possible to expand *We-toon* to support different types of characters, different body parts, or other features of webtoons if corresponding datasets are collected. In the action genre, for example, poses and movements are the most important features to consider. To support this, we could try using a recent model, InsetGAN [30], which generates a full body of characters.

Lastly, we categorized the images manually in the attribute filtering. This will become even more extensive when the system requires more categories. We plan to automate the process with a self-supervised clustering scheme such that the data can be separated without labels and do not require manual classification anymore.

7 CONCLUSION

We presented a communication support system—*We-toon*—for helping webtoon authors with different skill sets to collaborate. A thorough investigation was performed via a formative study, and the comprehensive collaborative webtoon creation pipeline was defined. We identified two stages (i.e., character design, and draft revision), in which ineffective communication is frequent, and accordingly designed *We-toon* to help the writers in retrieving the desired reference images and expressing their revision intents through image generation and synthesis. We conducted a user study with 24 professional webtoon authors, and the corresponding response analysis indicated that our system can support a better communication between the collaborators (i.e., the writers and the artists), thus providing an improved experience during the revision cycles. We further discussed the implications of using AI-assisted communication support systems.

ACKNOWLEDGMENTS

This work was supported in part by the National Research Foundation of Korea (NRF) grants funded by the Korea government (MSIT) under Grants NRF2019R1A2C2089062 and NRF-2019R1A2C1088900, and in part by the Hankuk University of Foreign Studies Research Fund. The ICT at Seoul National University provided research facilities for this study.

REFERENCES

- [1] Rinat Abdrashitov, Fanny Chevalier, and Karan Singh. 2020. Interactive Exploration and Refinement of Facial Expression Using Manifold Learning. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 778–790.
- [2] Yuval Alaluf, Omer Tov, Ron Mokady, Rinon Gal, and Amit Bermano. 2022. HyperStyle: StyleGAN Inversion With HyperNetworks for Real Image Editing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 18511–18521.
- [3] Yazeed Alharbi and Peter Wonka. 2020. Disentangled Image Generation Through Structured Noise Injection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [4] Paul André, Jaime Teevan, and Susan T Dumais. 2009. From x-rays to silly putty via Uranus: serendipity and its role in web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2033–2036.
- [5] David Bau, Jun-Yan Zhu, Hendrik Strobelt, Bolei Zhou, Joshua B. Tenenbaum, William T. Freeman, and Antonio Torralba. 2019. GAN Dissection: Visualizing and Understanding Generative Adversarial Networks. In *Proceedings of the International Conference on Learning Representations (ICLR)*.
- [6] Gwern Branwen, Anonymous, and Danbooru Community. 2020. Danbooru2019: A Large-Scale Anime Character Illustration Dataset. <https://www.gwern.net/Crops#figures>. <https://www.gwern.net/Crops#figures> Accessed: August 29, 2021.
- [7] Spencer Churchill. 2019. Anime face dataset. <https://www.kaggle.com/splcher/animefacedataset> Accessed: August 29, 2021.
- [8] Edo Collins, Raja Bala, Bob Price, and Sabine Süsstrunk. 2020. Editing in Style: Uncovering the Local Semantics of GANs. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [9] Sarah D'Angelo and Darren Gergle. 2018. An eye for design: gaze visualizations for remote collaborative work. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [10] Jeffrey T Hancock, Mor Naaman, and Karen Levy. 2020. AI-mediated communication: definition, research agenda, and ethical considerations. *Journal of Computer-Mediated Communication* 25, 1 (2020), 89–100.
- [11] Erik Härkönen, Aaron Hertzmann, Jaakko Lehtinen, and Sylvain Paris. 2020. GANSpace: Discovering Interpretable GAN Controls. In *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin (Eds.), Vol. 33. Curran Associates, Inc., 9841–9850. <https://proceedings.neurips.cc/paper/2020/file/6fe43269967adbb64ec6149852b55cc3e-Paper.pdf>
- [12] Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. 2020. Training Generative Adversarial Networks with Limited Data. In *Proc. NeurIPS*.
- [13] Tero Karras, Samuli Laine, and Timo Aila. 2019. A Style-Based Generator Architecture for Generative Adversarial Networks. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 4396–4405. <https://doi.org/10.1109/CVPR.2019.00453>
- [14] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and Improving the Image Quality of StyleGAN. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [15] Hyunsu Kim, Yunje Choi, Junho Kim, Sungjoo Yoo, and Youngjung Uh. 2021. Exploiting Spatial Dimensions of Latent in GAN for Real-time Image Editing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [16] Tae Soo Kim, Seungsu Kim, Yoonseo Choi, and Juho Kim. 2021. Winder: Linking Speech and Visual Objects to Support Communication in Asynchronous Collaboration. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 453, 17 pages. <https://doi.org/10.1145/3411764.3445686>
- [17] Robert Kraut, Carmen Egido, and Jolene Galegher. 1988. Patterns of contact and communication in scientific research collaboration. In *Proceedings of the 1988 ACM conference on Computer-supported cooperative work*. 1–12.
- [18] Yi-Chieh Lee, Naomi Yamashita, and Yun Huang. 2020. Designing a chatbot as a mediator for promoting deep self-disclosure to a real mental health professional. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–27.
- [19] Sheng Feng Li and Andy Hopper. 1998. A framework to integrate synchronous and asynchronous collaboration. In *Proceedings Seventh IEEE International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises (WET ICE '98)(Cat. No. 98TB100253)*. IEEE, 96–101.
- [20] Zhengqing Li, Theophilus Teo, Liwei Chan, Gun Lee, Matt Adcock, Mark Billinghurst, and Hideki Koike. 2020. OmniGlobeVR: A collaborative 360-degree communication system for VR. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 615–625.
- [21] Duri Long, Takeria Blunt, and Brian Magerko. 2021. Co-Designing AI Literacy Exhibits for Informal Learning Spaces. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–35.
- [22] Kai Lukoff, Taoxi Li, Yuan Zhuang, and Brian Y Lim. 2018. TableChat: mobile food journaling to facilitate family support for healthy eating. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–28.
- [23] Pascal Molli, Hala Skaf-Molli, Gérard Oster, and Sébastien Jourdain. 2002. Sams: Synchronous, asynchronous, multi-synchronous environments. In *The 7th international conference on computer supported cooperative work in design*. IEEE, 80–84.
- [24] Xi Niu, Fakhri Abbas, Mary Lou Maher, and Kazjon Grace. 2018. Surprise me if you can: Serendipity in health information. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12.
- [25] Cecil Piya, Vinayak, Senthil Chandrasegaran, Niklas Elmqvist, and Karthik Ramani. 2017. *Co-3Deator: A Team-First Collaborative 3D Design Ideation Tool*. Association for Computing Machinery, New York, NY, USA, 6581–6592. <https://doi.org/10.1145/3025453.3025825>
- [26] Martin Ragot, Nicolas Martin, and Salomé Cojean. 2020. Ai-generated vs. human artworks. a perception bias towards artificial intelligence?. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems*. 1–10.
- [27] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In *Advances in Neural Information Processing Systems*, C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, and R. Garnett (Eds.), Vol. 28. Curran Associates, Inc. <https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf>
- [28] Daniel Roich, Ron Mokady, Amit H Bermano, and Daniel Cohen-Or. 2021. Pivotal Tuning for Latent-based Editing of Real Images. *arXiv preprint arXiv:2106.05744* (2021).
- [29] Ugo Braga Sangiorgi, François Beuvsens, and Jean Vanderdonck. 2012. User interface design by collaborative sketching. In *Proceedings of the Designing Interactive Systems Conference*. 378–387.
- [30] Axel Sauer, Katja Schwarz, and Andreas Geiger. 2022. StyleGAN-XL: Scaling StyleGAN to Large Diverse Datasets. *arXiv preprint arXiv:2202.00273* (2022).
- [31] Isabella Seeber, Eva Bittner, Robert O Briggs, Triparna De Vreede, Gert-Jan De Vreede, Aaron Elkins, Ronald Maier, Alexander B Merz, Sarah Oeste-Reiß, Nils Randrup, et al. 2020. Machines as teammates: A research agenda on AI in team collaboration. *Information & management* 57, 2 (2020), 103174.
- [32] Yujun Shen and Bolei Zhou. 2021. Closed-Form Factorization of Latent Semantics in GANs. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 1532–1540.
- [33] Minhyan (Mia) Suh, Emily Youngblom, Michael Terry, and Carrie J Cai. 2021. AI as Social Glue: Uncovering the Roles of Deep Generative AI during Social Music Composition. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 582, 11 pages. <https://doi.org/10.1145/3411764.3445219>
- [34] Ryohei Suzuki, Masanori Koyama, Takeru Miyato, and Taizan Yonetsuji. 2018. Collaging on Internal Representations: An Intuitive Approach for Semantic Transfiguration. *CoRR* abs/1811.10153 (2018). <http://arxiv.org/abs/1811.10153>
- [35] Jaime Teevan, Shamsi T. Iqbal, and Curtis von Veh. 2016. *Supporting Collaborative Writing with Microtasks*. Association for Computing Machinery, New York, NY, USA, 2657–2668. <https://doi.org/10.1145/2858036.2858108>
- [36] Balasaravan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*. 182–195.
- [37] Alice Thudt, Uta Hinrichs, and Sheelagh Cappendale. 2012. The bohemian bookshelf: supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on human factors in computing systems*. 1461–1470.
- [38] Sunny Tian, Amy X Zhang, and David Karger. 2021. A System for Interleaving Discussion and Summarization in Online Collaboration. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW3 (2021), 1–27.
- [39] Carlos Toxtli, Andrés Monroy-Hernández, and Justin Cranshaw. 2018. Understanding chatbot-mediated task management. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–6.
- [40] Dakuo Wang, Elizabeth Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. 2020. From human-human collaboration to human-ai collaboration: Designing ai systems that can work together with people. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–6.

- [41] Christine Wolf and Jeanette Blomberg. 2019. Evaluating the promise of human-algorithm collaborations in everyday work practices. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–23.
- [42] Te-Yen Wu, Jun Gong, Teddy Seyed, and Xing-Dong Yang. 2019. Proxino: enabling prototyping of virtual circuits with physical proxies. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 121–132.
- [43] Zhenpeng Zhao, Sriram Karthik Badam, Senthil Chandrasegaran, Deok Gun Park, Niklas LE Elmqvist, Lorraine Kisselburgh, and Karthik Ramani. 2014. skWiki: a multimedia sketching system for collaborative creativity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1235–1244.