ShadowMagic: Designing Human-AI Collaborative Support for Comic Professionals' Shadowing

Amrita Ganguly
George Mason University
Fairfax, Virginia, USA

Chuan Yan
George Mason University
Fairfax, Virginia, USA

John Joon Young Chung
Midjourney
San Francisco, California, USA

Tong Steven Sun
George Mason University
Fairfax, Virginia, USA

Yoon Kiheon
Pusan National University
Busan, South Korea

Yotam Gingold
George Mason University
Fairfax, Virginia, USA

Sungsoo Ray Hong
George Mason University
Fairfax, Virginia, USA

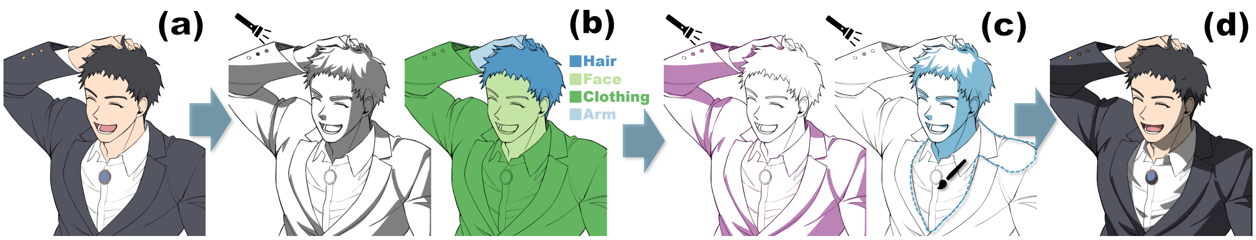

Figure 1: ShadowMagic workflow: ShadowMagic accepts a flat color layer and a line drawing layer as input (a). ShadowMagic's backend generates shadow suggestions based on a user's light direction choice (b, left). ShadowMagic's backend segments the flat regions into "semantic" segments, such as face, hair, or clothing (b, right). ShadowMagic's frontend lets users filter by semantic region. Users can choose to either adopt the AI-suggested shadows (c, pink shadows on the left, a user decided to use suggestions on clothing and arms) or apply additional edits (c, blue face and clothing shadows on the right, drawn by a user). A final outcome combines pink and blue shadows (d). This refinement can increase shadowing efficiency while providing sufficient control for a user.

ABSTRACT

Shadowing allows artists to convey realistic volume and emotion of characters in comic colorization. While AI technologies have the potential to improve professionals' shadowing experience, current practice is manual and time-consuming. To understand how we can improve their shadowing experience, we conducted interviews with 5 professionals. We found that professionals' level of engagement can vary depending on semantics, such as characters' faces or hair. We also found they spent time on shadow "landscaping"—deciding where to put big shadow regions to make a realistic volumetric presentation—while the final results can dramatically vary depending on their "staging" and "attention guiding" needs. We found they would accept AI suggestions for less engaging semantic parts or landscaping, while they would need to have the capability to adjust details. Based on our observations, we built ShadowMagic that (1) generates AI-driven shadows based on typically used light directions, (2) enables a user to selectively choose the results depending on the semantics, and (3) allows users to finish shadow areas by themselves for further perfection. Through a summative evaluation with 5 professionals, we found that they were significantly more satisfied with our AI-driven results than a baseline. We also found ShadowMagic's "step by step" workflow helps participants more easily adopt AI-driven results. We conclude by providing implications.

CCS CONCEPTS

• **Human-centered computing** → **Interactive systems and tools**; • **Computing methodologies** → *Graphics systems and interfaces*; **Artificial intelligence**.

UIST '24, October 13–16, 2024, Pittsburgh, PA, USA

ACM ISBN 979-8-4007-0628-8/24/10
https://doi.org/10.1145/3654777.3676332

The first and the second author contributed equally. The third author was not involved in technical pipeline development. The last author serves as the corresponding author.

KEYWORDS

Human-AI collaboration, Comic Shadowing, System for Professionals

ACM Reference Format:

Amrita Ganguly, Chuan Yan, John Joon Young Chung, Tong Steven Sun, Yoon Kiheon, Yotam Gingold, and Sungsoo Ray Hong. 2024. ShadowMagic: Designing Human-AI Collaborative Support for Comic Professionals' Shadowing. In *The 37th Annual ACM Symposium on User Interface Software and Technology (UIST '24), October 13–16, 2024, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3654777.3676332

1 INTRODUCTION

The comic industry has experienced a shift from paper-based to digital across the recent decades [46]. This transition has brought numerous changes by which comics are colorized, presenting challenges and opportunities for researchers and comic professionals [49, 53]. The digital comic colorization process generally follows the step-by-step stages of line drawing, flatting, shadowing and lighting, background, and special effects [53]. Through shadowing, comic professionals create depth, natural volume for objects, and a sense of realism. Mastery in shadowing can draw readers' attention to a character's emotion, actions in a scene, and atmosphere, ensuring that the storytelling remains clear and effective for readers [56]. As readers' expectations of high-quality comics evolve, comic professionals face adapting their expertise to meet these demands.

Since shadowing takes a crucial role in delivering a story to readers, past studies have advanced techniques for generating shadows [19, 26, 49, 57] and relevant techniques in comic colorization [6, 50, 52]. However, to date, professionals still heavily rely on basic features in mainstream image editing software, such as brushes, lassos, and erasers [1, 7, 18, 44], making shadowing still manual and labor intensive [53]. Professionals value their specific and contextual control on shadowing, leading them to prefer manual brushing and erasing over AI-driven automation [53].

This work aims to understand how new AI-driven shadowing support can naturally be applied to comic professionals' current shadowing workflow and practically help them. In doing so, we first conducted a formative study with five comic professionals (S1). In S1, we observed how they create shadows. We learned their shadowing workflow and characteristics of "good" shadows. Further, we listened to their thoughts on applying current state-of-the-art AI shadowing suggestions and how AI can be designed differently to improve their shadowing practice. We learned the following. First, professionals may have different levels of engagement in shadowing depending on semantic regions. For instance, they feel more engaged when adding shadows on faces as slight line changes can cause a different feeling, while they may feel it tedious to shadow hands or spiky hair. Second, they spend their energy in "landscaping" shadows—deciding where they will put large shadow regions—for realistic presentation. This process is perceived as repetitive, which also can be assessed based on a certain level of "norms". Third, what makes shadowing creative is that shadowing is not merely a device for depicting normative volume, but also a device for adding different emotions of characters (i.e., "staging") or attracting a reader's eyes to a specific object (i.e., "attention guiding") depending on professionals' intention. With different staging and attention-guiding intentions, desirable outcomes can be dramatically varied even given the same line-drawing scene, which makes it challenging to define ground truth. We found participants' willingness to adopt AI's suggestions to be higher if they perceive "that part" of the shadowing is tedious or repetitive.

Based on the S1 results, we designed ShadowMagic, that (1) provides reasonable quality AI-driven shadows at the beginning to reduce the effort for landscaping and (2) enables a user to selectively apply AI suggestions depending on semantics. ShadowMagic also lets a user customize the final outcome to effect their staging and attention-guiding needs. ShadowMagic leverages two AI models; one predicts initial shadowing suggestions that provide different results depending on typically used light directions, and the other one segments the semantic boundary with the 5 categories of hair, face, clothing, arms, and other objects. To support the case where a user wants to finish manually, ShadowMagic provides brush and eraser features. With 5 comic professionals, we evaluated our backend models' shadow quality in an experimental study and the ShadowMagic's quality through interviews after 1 day of use (S2). We found that professionals perceived the quality of shadows built by our AI model to be significantly higher than the baseline. We also found that ShadowMagic's interaction modality can help professionals to imagine why and how they can adopt AI-driven suggestions in their shadowing workflow.

This work offers the following contributions:

- **S1 Findings:** S1 provides empirical insights into the highly context-dependent nature of comic shadowing workflows and general perception about how professionals perceive challenges in using existing AI-driven solutions.
- **ShadowMagic and S2 Results:** We introduce ShadowMagic, a novel system specialized in shadowing support for professionals. It incorporates two AI models and the workflow informed by S1. The S2 results provide the expected benefit when future researchers can expect that adopting similar workflow designs and techniques used in ShadowMagic.
- **Implication for design:** We provide implications for design that discuss how the notion of intermediate representation [53]—which predominantly argues how to develop AI's input and output in a "stage-by-stage" fashion based on a user's sequential workflow—can be extended to a non-sequential, non-deterministic, and "interaction-by-interaction" fashion within each stage.

2 RELATED WORK

In this review, we first cover studies that explain how AI, including generative AI, can support graphic professionals in general, as only a limited volume of research focuses specifically on enhancing professionals' shadowing experience. We then review computational approaches to creating shadows. Finally, synthesizing the two directions, we describe how more deeply observing comic professionals' shadowing workflow can provide new design opportunities.

AI for Professionals in Graphical Fields: Generative AI has impacted image generation and manipulation. It aids creative processes by rapidly visualizing concepts, generating assets, and creating image content from prompts [48]. The overall progress in AI techniques has made it possible to generate images from various types of data, reducing the manual effort required for exhaustive image capture [16]. Generative AI models like DALL-E 2 [33], Stable Diffusion [5], and Midjourney [55] create novel images from text by utilizing advanced deep learning approaches such as diffusion models to synthesize images across various styles and domains. Some studies investigated how designers use image generation models for early-stage visual design [8] and 3D design [54], highlighting their potential and limitations for architectural design practice. Beyond static images, emerging AI systems like Make-A-Video [43] for video generation and Magic3D [31] for 3D scene creation are enabling creativity support across multimedia formats. Despite challenges like intellectual property issues and biases in training data [61], generative AI's accessibility, versatility, and rapid advancements highlight its potential to transform creativity by aiding humans in innovative thinking and imagination.

While generative AIs are pushing boundaries, there are many other notable efforts to leverage non-generative AIs for creativity support in varying application areas, such as painting [3], 3D modeling [34], and visual arts [27]. Some work has explored co-creative agents that can improvise and collaborate with human users on creative tasks like drawing [14, 15]. As AI-driven tools become more prevalent, researchers have studied design principles for effective human-AI interaction (HAI) and collaboration [42]. The automation-control framework highlights the importance of defining proper roles for human agency versus AI automation [23]. DuetDraw, an interface for drawing with AI, emphasizes collaborative creativity. Its user study highlights the importance of detailed instructions and user-led interactions [36]. FlatMagic [53] reduces the effort of flatting in comic colorization, a labor-intensive stage, highlighting the benefits and challenges of carefully reflecting professionals' workflow in designing human-AI collaboration systems.

However, applying such principles to design field-ready creativity support tools for professionals is not without challenges [42]. Little prior work has specifically explored applying the automation-control paradigm to creative tools for comic shadowing by professional users [32]. While existing work showcases the state-of-the-art capabilities of generative AI and proposed technologies in graphics and artwork, a gap remains in addressing the unique needs of comic professionals.

Computational Shadowing Support: There are several computational approaches built for generating shadows in digital comics. Some offer no interactive user interface but exist solely as algorithms and focus on the technical aspects of shadow generation, employing various computational techniques to produce realistic shadows without direct user involvement. Others provide interactive systems, enabling users to generate shadows based on creative needs.

Algorithmic techniques focus on the technical aspects, using computational methods to create realistic shadows autonomously. For instance, Ramamoorthi et al. conducted a comprehensive gradient analysis of how lighting variation, surface reflectance, curvature, and soft shadows individually contribute to and combine in shading effects [38]. The authors of this paper [59] introduced an algorithm that generates digital painting lighting effects by estimating stroke density and mimicking artists' workflow. Another study used deep learning to generate detailed and artistic shadows from line drawings and lighting directions, respecting lines and space with sophisticated details and effects [60]. Colbert and Krivanek presented an efficient real-time rendering technique that combines BRDF importance sampling with environment map filtering to enable interactive viewing of objects with complex spatially-varying glossy materials under natural illumination [29]. DeCoro et al. introduced a shading rig [37] system where artists can pre-animate desired toon shading styles that automatically preserve the artistic direction under changing lighting at runtime. Going further, Sloan et al. proposed a method to capture custom artistic shading models from sampled artwork, allowing users to generate unique non-photo-realistic renderings that emulate the look and feel of particular artistic styles for depicting materials like skin, metal or paint [45].

There are dedicated shadowing systems that allow users to generate shadows based on their own preferences. Image editing softwares such as ClipStudio [7], Photoshop [1], and Sketchbook [44] offer general painting support that helps make shadows. Some approaches like 2DToonShade [26] semi-automatically generate shading and self-shadows for cel animation by applying simple yet effective algorithms directly to the 2D drawn artwork, providing an intuitive interface that stays close to the natural creative process. Other tools like SmartShadow [56] leverage deep learning to enable digital artists to draw shadows on line art through interactive brushes that allow scribbling to indicate shadow areas, precisely control boundaries, and consistently propagate directional shadows substantially accelerating the shadowing workflow. Beyond assistance tools, 2.5D modeling techniques [22] simulate 3D rotations from 2D vector art, automating in-betweening while rendering interactive 3D shading effects like Phong, cel shading, and environment mapping onto the 2D artwork. This bridging of 2D and 3D enables richer stylization while maintaining a hand-drawn aesthetic. Computational tools have also been developed for 3D shadow generation support. In Breslav et al. [4]'s approach, the desired 2D shadows cast by a 3D sculpture are specified as input, and geometric optimization computes a 3D shadow volume approximating those shadows. For voxel-based 3D graphics, congradient shading [12] recovers surface normals from the discrete voxel data to enable simple table-driven shading. Extending further into 3D non-photorealistic rendering [35], approaches have been proposed to map and transform 2D motion patterns onto 3D surfaces in a style reminiscent of hand-drawn art.

While generating shadows has been widely studied as a topic of research, we found little e ort has been made to apply a user-centered approach for supporting comic professionals' shadowing process. Recent literature in Human-AI Collaboration emphasizes the importance of deeply understanding the expert's work ow for generating the system used by them [23, 42]. This is because professionals' use of technology is highly speci c and contextualized. Without eliciting speci c requirements through dedicated e ort, the system may likely yield unsuccessful outcomes [53].

3 STUDY 1: FORMATIVE STUDY

We conducted Study 1 (S1) with comic professionals to gain a deeper understanding of their current practice and inform how we can design better human-AI collaboration in comic shadowing support. Our speci c Research Questions (RQs) are as follows:

RQ1. How do comic professionals create shadows? In particular, what is the common process that the professionals apply in creating shadows, what characteristics of shadow results make them desirable/good, and what are the di cult parts of creating shadows?

RQ2. How do comic professionals perceive using AIs in their shadowing work ow? This RQ includes their perception about the current state-of-the-art AI-driven shadows, and how the design of an interface can enable them to use the AI-driven shadow suggestions.

3.1 Methodology

We chose a semi-structured interview as a method to gain insights on our RQs. In recruiting participants, we applied convenience and snowball sampling strategies [18]. In doing so, we contacted comic companies specialized in creating digital comics and shared our participant criteria: rst, they must currently work in the comic industry and second, they must have more than 5 years of professional experience in colorizing digital comics. There were a total of 5 participants. Three of the participants identi ed as female and the rest as male. All received bachelor's degrees. Their ages ranged from 25 to 35. We conducted interviews on a rolling basis. Every interview was conducted online.

The interview was divided into two parts. First, we went through a series of questions built for understanding the following aspects: (Q1) their general process in creating the shadow, (Q2) the aspects that make good or bad shadow results, (Q3) the aspects they feel are tedious or enjoyable in creating shadows, and (Q4) their strategy to make the outcome more unique. Second, we demonstrated three AI-generated shadowing methods, ShadeSketch [60], SmartShadow [56], and PaintingLight [59], and heard their opinion on (Q5) what aspects they like or dislike, (Q6) the likelihood of applying the outcomes to their practice, and (Q7) what aspects must be handled to be used in their practice in the future. To provide a consistent interview experience across participants, we made a slide deck with every question and the AI-driven demos above. In addition, we prepared ve line drawings provided by other professionals in case participants wanted to demonstrate their method of shadowing. We shared the slide and line drawings before the interview. Every interview was video-recorded and then transcribed.

Two researchers performed qualitative analysis. They followed an iterative qualitative coding process [24, 39] that starts from coding and then connecting the repeating or related insights in developing analytic memos. They iterated coding and memos after each new participant. After nishing every interview, they built an a nity diagram and derived the general themes separately. In concluding their nal themes, they compared the two theme structures and discussed the commonalities and discrepancies. They continued multiple rounds of discussions until they reached a consensus.

3.2 Results

Roles of shadows: We found that desirable characteristics of shadows (Q2) are highly related to professionals' general shadowing process (Q1), and their strategy to make unique shadowing (Q4). Good shadowing depends on two primary factors: landscaping and context. Landscaping involves deciding where to place shadows to represent the realistic volume of objects. This is a fundamental requirement and objective in shadowing and is more amenable to automation. While context in shadowing can be related to multiple di erent aspects, we found participants generally agreed that shadowing has two major functions: staging, and attention guidance. These aspects are subjective, and professionals often engage deeply with these tasks, implying less enthusiasm for AI intervention in this area.

When adding shadows in designing the scene in a digital comic, *staging* means making the characters' internal emotions and their situations (P2), unmistakably clear by putting shadows that don't look realistic in natural light settings (P4). Even using the same line drawing, the ways comic professionals create shadows can make huge di erences in terms of presenting the character's internal state. P4 demonstrated the di erence through examples shown in Fig. 2 (a1) and (a2). She started explaining by putting the light source from the left top (see the light cone in Fig. 2 (a1)) then added the shadows that look realistic and natural. The character's gentle smile will make this scene to be kind and peaceful. However, as she put thick and unrealistic shade on her face, she became the character with conspiracy. These di erent ways of putting the shadow can be applied di erently to di erent mental statuses, such as anger, frustration, hate, forgiveness, love, happiness, boredom, and many more (P1).

Figure 2: Advanced functions of shadowing. Staging: (a1) Natural shadowing on the face expresses a kind greeting, while (a2) unnatural shadowing foreshadows a conspiracy. Guiding reader attention: (b1) Realistic shadowing draws readers' focus to the right character's possessions, while (b2) simplifying his possession's shadows causes a reader to focus more on the situation in which the left character kneels while the right threatens.

Another function is attention guiding. In general, people's gaze can go first to the object with detailed shadow. They may neglect or put their gaze later on the entirely shaded object. P2 mentioned that there are always objects of focus that comic professionals hope readers see first then other objects. She continued: Making readers engaging in 70 new scenes every week is a challenge. I cannot expect them to focus on every piece I draw. A few fans may get everything, but most won't. Across the 70 scenes, there is an attention priority. I put shadow based that priority to help readers get the overall plot and storyline using minimum attention. P4 gave another demonstration on this aspect through Fig. 2 (b1) and (b2). Fig. 2 (b1) shows a natural color where shadows are applied with the same degree of detail. Since the right character has several objects in his body, people's gaze may go to the details of the right character's possessions, such as throwing knives or wrinkles in clothing. Fig. 2 (b2) intentionally reduces the shadow details. She explains that now people's attention may go to understanding the situation between two people the left kneels and the right threatens rather than details of the right's possessions.

Overall, the way artists draw shadows to support staging and attention guidance indicates that expressing context is an intricate process that the artists largely enjoy, as it allows them to creatively influence the viewer's focus and emotional response. On the other hand, the basic goal of shadowing which is adding realistic volumes on objects is perceived as more common and tedious, involving repetitive tasks that require precision but offer less creative satisfaction. Artists spend substantial time on landscaping in shadowing, i.e., deciding where to create big base shadows to make realistic volumes. They prefer AI suggestions to tackle the common and tedious parts, such as ensuring consistent light sources and accurate shadow placement.

Where in the scene matter: We found the aspects of shadowing that make them feel tedious or enjoyable (Q3) are closely connected to their intention to adopt AI-driven suggestions. First of all, they felt tedious when they shadowed: (1) complex objects that require certain effort, such as clothing with complex patterns, wrinkles, or made with fur (P2), hands (P2), or curly or short hairs (P5), (2) objects that have not much of value from staging perspective, such as crowds in background (P1), and (3) objects they have shadowed repetitively in the past, such as hair except face (as subtle shadowing change in the face can heavily change the scene atmosphere). Additionally, shadowing large areas, or landscaping, which involves creating realistic volumetric representation, is also perceived as tedious. Participants noted that they spend a significant amount of time on this process, as it is crucial for achieving a natural and convincing presentation.

Not surprisingly, when we asked about their intention to adopt the AI's suggestion, our participants seemed open to adopting the AI's suggestion for those that they felt tedious. P3 mentioned: The parts that I feel tedious become labor. The parts that challenge me make me to be more creative. I hope AI can cut down my labor so that I can focus better on my creation.

About current AI shadowing suggestions: Regarding their thoughts on current AI-driven shadowing approaches (Q5) and aspects the current tools must handle to be used in the future (Q7), while every participant mentioned the recent results have remarkably improved compared to non-AI-driven shadowing approaches, they still feel uncomfortable using AI-driven suggestions mainly because of their insufficient quality. P3 mentioned: It requires more effort to revise than do it from scratch. Another reason that made it hard to consider applying AI-driven suggestions is related to the interface handles; in some cases, they found the existing handles are unnecessarily complicated. After seeing the interface that provides 360 degrees of light direction, for example, several participants mentioned that there is a set of light directions commonly used in digital comics. In other cases, they couldn't indicate their basic intention in generating shadows. In other cases, they felt it is not possible to revise the undesirable outcomes. One interesting insight we learned was that most approaches rely on line drawing only to create the suggestions. However, P2 and P4 mentioned that the shadowing is impacted by flat regions rather than line drawing, implying that it may be desirable to consider line drawing and flat results in generating shadow suggestions.

Common shadow settings in digital webcomic: There can be different styles in shadowing, such as a full gradient shadow, multiple steps, and a cell style. Our participants mentioned that AI-generated shadow can be specifically useful for cell-style shading mainly because this type of shadowing is the most frequent form of shadow applied in the comic industry due to its straightforwardness, simplicity, and efficiency in production. P3 mentioned: *Cell shadow is the most frequent style of shadow we use. Cell shadow is used every time, even when working on the product with full gradient*

3.3 Design Requirements

S1 results found that a good shadow depends on the *context*. A line drawing cannot determine the ideal shadow. Rather, a good shadow receives a reader's attention from the previous scene and passes the attention to the next scene based on a professional's staging and attention-guiding strategy. In that sense, defining ground truth in comic shadowing is nearly impossible. In applying the strategy, we observed that they added natural shadows on some objects while applying unnatural shadows on other objects, as shown in Fig. 2 (a2) and (b2). Connected to this observation, depending on the type of objects in the scene, a professional's intention to adopt AI-driven suggestions may vary they may shadow some objects manually while being more open to applying AI's suggestions for less important ones. Finally, we found applying AI-driven suggestions is not without challenges, and the main reasons are insufficient suggestion quality and mismatched interface handles. Based on our observation, we drove the following Design Requirements (DRs):

- DR1. Improving the initial quality of AI-driven suggestions must be the starting point of designing a human-AI collaboration system for comic shadowing.
- DR2. To support staging, attention-guiding, and selective AI-driven suggestion adoption, the system can help professionals to toggle the AI's suggestions depending on semantics in the scene, such as hands, face, and hair.
- DR3. In providing light direction control, simplify it by following widely used light direction settings in the comic industry.
- DR4. Follow the golden rule of Human-AI Collaboration: Provide a method to manually recover the shadows when a user is not satisfied with AI's suggestion.

4 SHADOWMAGIC

Our approach aims to improve AI's ability to generate high-quality shadows for comic colorization workflows, while giving professionals control over shadow results through an end-to-end interface. Our system performs two steps to generate an initial segmented shadow suggestion for user editing. Each input contains a line layer and a flat layer. We first predict its shadow using ControlNet [58] and its semantic segmentation using YOLOv5 [28], as shown in Fig. 3 (a). Next, we extract a binary shadow from the ControlNet prediction and cut it into multiple shadow layers based on the semantic segments from YOLOv5, as shown in Fig. 3 (b). Finally, we present all the layers in our web-based editing interface for further refinement, as shown in Fig. 3 (c).

The front-end system allows professionals to adjust shadow areas at a semantic level, providing flexibility to apply shadows according to their creative vision and requirements. Based on our findings in S1, we designed a streamlined front-end that removes unnecessary and undesirable controls such as the need to specify light direction in a continuous (e.g. 360°) manner. Furthermore, we follow the principle of human-AI collaboration [42], where AI suggestions can be controlled by the professionals. Professionals have the option to select which parts of the AI-generated results to adopt, and the system fully supports manual editing if desired. To facilitate this manual editing, we offer basic brush and eraser tools, along with undo and redo functionality.

4.1 AI models

The original dataset to train the AI models was created by professionals at a single company who made shadows dedicated to this project. After S1, they first prepared 1,000 existing flat-colored line drawings made previously. Of these, they made 2 directions (left and right) of shadows for 960 drawings and 4 directions for 40 drawings. This was because left and right had more variability which required more examples for generating results to meet our quality standard while top and bottom directions were fairly consistent. Tab. 1 shows some high resolution examples from the dataset. More examples, in low resolution, are available on GitHub[1]

4.1.1 *Initial shadow prediction*. Although the comic/illustration quality generated by current diffusion models can appear close to the professional level, these generic image generation models cannot fulfill specific generation tasks, such as adding shadows based on a given light direction (see Fig. 5 (a2) and (a3), where the result in (a2) is generated by StableDiffusion XL[2] with Anime style and the result in (a3) is generated by Reproduction[3].

To obtain a shadow prediction with acceptable quality (DR1), we guide the prior from pre-trained stable diffusion models on a large-scale dataset to our shadowing task, as shown in Fig. 5 (a4), (a5), and (a6).

[1] https://github.com/Amrita537/ShadowMagicUIST24
[2] https://huggingface.co/docs/diffusers/en/using-diffusers/sdxl
[3] https://civitai.com/models/118729?modelVersionId=144778

Figure 3: System framework: ShadowMagic contains a ControlNet and a YOLOv5 ne-tuned with our dataset. (a) For each input containing a at layer and a line layer, we predict 4 variant grayscale shadowing results with the same given light direction and its semantic segmentation mask. (b) We then extract and crop the shadows into sub-shadow regions depending on semantic segment labels (e.g., hair or face). Finally, a user will pick one of the suggestions and re ne it through our front-end.

Figure 4: Example of Shadow Size Shrinking: (a1 and b1) Shows the initial shadow, (a2 and b2) shows the shadow shrink after 4 5 clicks, (a3 and b3) shows the shadow shrink after 8 9 clicks.

To achieve this target, we created our own training dataset as shown in Fig. 5 (b1) to (b5). We collected 342 images created by professionals from a digital comic company. 41 of them contained ground truth shadows from all four light directions (left, right, top, and back) widely used by professionals (S1). The remaining 301 images contained ground truth shadows from only the left and right light directions.

Based on this dataset, we generated the conditioning images as a blend of the line layer and the at layer, as shown in Fig. 5 (b1). The corresponding shadowed target images were blends of the ground truth shadow and line layers, as shown in Fig. 5 (b2 right). In this way, the network is trained not to retain the color information of the input image and focus solely on the task of predicting shadows. This also enhances the quality of the output in the subsequent shadow extraction step. As text-to-image di usion models can be guided with natural language prompts, we also used the prompt, add shadow from [light direction] light and remove color .

We formulated the shading task as a conditional image-to-image translation task and trained a ControlNet based on our constructed training set. We chose Divineelegancemix[4] as the base ControlNet model and trained on an Nvidia A100 GPU for 48 GPU hours.

When the trained model is used for the inference of shadows from input images, we upscale the output image, since the model output's resolution is 1024 1024 typically much smaller than professional input images. We used Real-ESRGAN [51] to up-sample the model's output by 4 and then resize it to match the input image. We extract binary shadows by thresholding the output to 0.5 (assuming pixel values in the range 0 to 1).

We allow users to control the size of the generated shadows via a simple but e ective iterative shrinking algorithm. We apply a 5 5 Gaussian lter to smoothly erode the binary shadows and then re-threshold. Applied naively, this approach would cause shadows to peel away from line drawing edges, which typically correspond to occluding contours and hence shadow discontinuities. To prevent this, we preserve the shadow pixel values near the line drawing edges. We cache every shrunken shadow after users click the shrink button, so that the user can expand the shadow by restoring from the cached results. Examples of shadow shrinking can be seen in Fig. 4.

4.1.2 Character semantic segmentation . We allow users to selectively adopt shadow suggestions (DR2). We do this by segmenting shadows according to scene semantics. We decompose this into two sub-problems: 1) accurately identifying semantic segment labels and 2) predicting precise semantic segment contours. We considered ve semantic labels (S1): hair, face, clothing, arm, and object. The schema was guided by S1, with professionals mentioning they pay more attention to faces and arms and less to hair or common, everyday

[4] https://civitai.com/models/6174?modelVersionId=48473

Figure 5: Shadowing results comparison & dataset examples. Upper row: (a1) An input image. (a2) and (a3) Shading results from two o -the-shelf di usion-based models, StableDi usion XL and Reproduction, respectively. (a4) Shading results from our model. (a5) Blending our model's shading results with the input clearly shows our model's better visual quality. (a6) ShadowMagic allows artists to decrease the shadow size. Bottom row: (b1) One training example with shadow ground truth images from the (b2) right, (b3) left, (b4) top, and (b5) back light directions. (b2, right) The training target during our ne-tuning.

clothing. Beyond these four classes, we added an additional category labeled other to account for elements that do not t neatly into the primary categories. To perform semantic segmentation, we ne-tuned the YOLOv5 segmentation model on a dataset we constructed for this task. The dataset images are created by professionals from the same digital comic company, resulting in 614 images. We annotated semantic labels of these images with an annotation tool we developed (Fig. 7). The design of the annotation interface was informed by past annotation user interface studies [2, 9, 11].

The YOLOv5 model outputs bounding boxes of detected objects with labels and masks that approximately identify the semantic segments. To obtain more accurate contours, we keep the labels identi ed from YOLOv5 and replace the approximate masks with precise ones extracted from the input at color images. Namely, we replace an approximate mask with the union of all at regions with a high degree of overlap (Fig. 7b1, b2, and b3).

4.2 ShadowMagic User Interface

We present ShadowMagic 's front-end (Fig. 6), which helps visualize the AI generated shadows and enables users to modify AI-generated shadow results. The system allows users to control shadow generation by selecting the right direction and reviewing the suggested shadows. If users are not fully satis ed, they can manually edit shadows using brush and eraser tools (DR4). To support selective adoption of AI suggestions, the system enables professionals to toggle AI suggestions based on scene semantics, such as hands, face, and hair (DR2). Additionally, ShadowMagic simpli es light direction control by following widely used settings in the comic industry. There are four light directions: left, right, top, back. For each direction, ShadowMagic o ers four shadow suggestions (DR3). ShadowMagic is Implemented as a web application with HTML, CSS, Bootstrap, and JavaScript.

Speci cally, the user starts shadowing by opening a PSD le (F1). The image appears on the canvas (F5) with line drawing and at layers listed in a Layers panel (F9). The user can request ShadowMagic to suggest shadows with the speci ed light directions. This produces four candidates shadows (F6). The user can customize the shadows by changing the opacity using a slider and shadow size using the expand (+) and shrink (-) buttons (F7). Initially, the user can only shrink the shadows. The user can also toggle the part selector to expand and see the ve semantic segments ShadowMagic provides (F7, F11). The user can check and uncheck checkboxes to turn shadows on and o for regions by name (F11). The user can also manually edit the shadow, either by drawing brush strokes (F3) or erasing (F4). The brush and eraser have simple undo and redo functionality placed beside them on the UI (F3, F4). Once satis ed, users bookmark the current shadow layer to save it (F8). Bookmarked shadows appear in a list (F12). Their visibility can be toggled with a click. Users can save the current shadows and all bookmarked shadows layers by choosing Save from the File menu (F1, extended).

In addition to the above-mentioned functionality, ShadowMagic also includes basic canvas interactions (F2), such as panning and zooming.

Figure 6: ShadowMagic UI layout. Left: initial state. Right: extended state

Figure 7: (a1 and a2) Annotated contour and layer masks, respectively, using our annotation tool. (b1) YOLOv5 object detection and instance segmentation results. (b2) Flat based segmentation contours. (b3) Overlap identi cation and at contour merging.

5 SUMMATIVE STUDY

To evaluate the e ectiveness and e ciency of ShadowMagic, we conducted a summative study (S2) with ve professionals who work in a digital comic production company. The goals of S2 are described in the following Research Questions (RQs):

RQ1. Quality of Shadow Suggestions How do professionals perceive the quality of the shadows generated through ShadowMagic? Improving the quality of the shadows seemed to be the priority goal for developing practical and useful shadowing solutions. In measuring the quality of our shadow generation results, we aim to quantitatively evaluate it by comparing it against a state-of-the-art AI-driven method [60] (S2A).

RQ2. Quality of Interaction How do professionals perceive the quality of the interaction provided through ShadowMagic? In measuring the quality of interaction, we will measure it qualitatively through an interview after using ShadowMagic for 1 day (S2B).

5.1 Participants

We recruited ve professionals with years of experience in digital comic generation. There were no overlap between S1 and S2 participants. In recruiting participants, we used the same method as S1. Consequently, four manager-level professionals and one 3rd-year employee in webcomic companies agreed to test our AI-generated results and ShadowMagic interaction. Their ages range from 24 to 41 (= 30). They have an average of two years of experience in the cartoon industry, and three of them have over two years of experience in comic shadowing. Their companies' core products are webcomic production (# = 4) and post-colorization (# = 1).

5.2 Experimental Study (S2A)

In this study, we compared the perceived quality of our method ($M_{4G?}$)'s outputs to those from a state-of-the-art baseline (M_{10B4}), the rst open-sourced framework to auto-generate shadows using generative adversarial network [60].

5.2.1 Method

We measured participants' perceived quality of the shadowing generated by two methods through a survey. We rst generated 48 shadowing examples using the two shadowing methods (M_{10B4} and $M_{4G?}$) applied to our comic dataset for a fair comparison. The dataset consists of unique sketches with equally distributed light directions (left, right, top, and back), in which 24 sketches were from the test data of the baseline model, and 24 were collected by our team, which were not used for training our model. Then, we showed each example (shu ed to reduce any order e ect) and asked participants to rate their perceived quality of shadows. Speci cally, we asked to answer on a 7-level Likert scale from Strongly Disagree (1) to Strongly Agree (7) regarding the statement This example is a high-quality shadow. Since the basic role of a shadow is to present a realistic volume, by "high-quality" we refer to how realistic the shadows look.

Figure 8: Distributions of perceived quality ratings (S2A) regarding two methods (M_{10B4} and $M_{4G?}$). The red vertical lines represent the average ratings.

5.2.2 Results. When comparing M_{10B4} with $M_{4G?}$, participants gave higher ratings on the shadowing examples created by our $M_{4G?}$ (M = 3.06, SD = 1.34) than M_{10B4} (M = 2.64, SD = 1.25). Based on the Mann-Whitney U-test, we found a signi cant di erence between the two methods' ratings (? = 0"019Ÿ 0"05). The Mann-Whitney U-test was used since the rating data was not normally distributed (the Shapiro-Wilk normality test showed a signi cant ?-value Ÿ 0"0001). The Fig. 8 shows the rating distribution for perceived quality.

As shown in Fig. 8, we can identify an improvement in the perceived shadowing quality when using ShadowMagic. The number of records with very poor shadow quality (1-2 ratings) was reduced by nearly 1/3 and moved to higher quality perceived by the comic professionals, particularly more improvements toward the 3-4 ratings than the 5-6 ratings. No perfect-quality examples of shadowing (7 ratings) were created by either method, but the following best-quality examples (6 ratings) were only perceived in our experimental condition.

5.3 Expert Interviews (S2B)

5.3.1 Method. We prepared the ve links that each participant can use ShadowMagic separately. To facilitate their usage, we also prepared 19 PSD les; each had a at layer and a line drawing layer. Aside from the 19 PSD les, we informed our participants that ShadowMagic can generate results using their own les. We provided the following guidelines: (1) the le must have two layers; in terms of naming the layers, one must include at and another one must have line , (2) suggest using les that don't exceed 2000 px height or width for fast shadow generation. We asked them to use ShadowMagic for at least 1 day. We gave them 5 days before the interview. After the usage of the system, we conducted a 1-hour interview with our participants.

We divided the interview into three parts: (Q1) General perception about ShadowMagic, (Q2) Their assessment of the quality of initial shadow, and (Q3) Their perception of applying generative AI-driven results depending on the semantic parts separately. Across every part, we asked about their perceived pros and cons and how each feature can a ect their e ciency and e ectiveness productivity. In Q1, we additionally asked about ShadowMagic's balance between automation and control. Finally, we asked what possible directions that could better support their shadowing process.

5.3.2 Results. Every participant shared the bene t of using ShadowMagic as boosting their task e ciency. Generating AI-driven suggestions can help them to plan the outline of the shadows faster. P4 commented: De ning the high-level shadows of a big volume takes a lot of e ort and time in the shadowing process. Seeing the shadow volume helped me quickly set up the coarse level of shadows. P1 commented: If the suggestions are reasonable, using ShadowMagic may save 70% of my time. While participants were generally favorable about the way they experienced our shadows, they mentioned that the tool has a certain level of uncertainty related to poor AI suggestions. P5 commented: I like the fact that I have multiple candidates. I can generally nd at least 1 shadow that I can use to work on. That being said, I had some cases that I didn't like any of the suggestions.

When it fails to generate good shadow suggestions. There were diverging opinions on our manual re nement tools. When the AI's suggestion was not good enough, participants mentioned that using our tools to touch up the result was not easy, mainly due to the poor quality of our brush and eraser. P2 commented: ShadowMagic doesn't provide a pressure sensitive brush. But this feature is essential. So, I prefer to download the intermediate results from here and then nish up the shadow using Photoshop. P4 and P5 also mentioned the necessity of improving the brush for better control. On the other hand, P2, P3, P4, and P5 shared that the semantic segmentation can counterbalance the weakness of the brush function. Even if the results are not all good, they found some segments to still be useful, like clothing or hair. They could show and hide the good and bad regions, saving their time and e ort and helping them make better results. P5 commented: I loved that feature, it can help me only select the part that I like. It enables both e cient and e ective shadowing.

It seemed that the perceived usefulness of the system was heavily a ected by the quality of the initial shadow. While some insights are anecdotal, participants shared their detailed observations regarding the patterns they perceived that ShadowMagic generates good or bad shadows. The common sentiment was that ShadowMagic can create shadows for big volumes well. However, there are artifacts when looking into the details. This meant that using our tool helps plan a volume's shadows or serves as potential reference imagery while presenting details with reasonable quality that artists can x with a few strokes.

However, it is not possible to use the results directly for production. P1 commented: I found ShadowMagic generates very good quality for clothing that takes up a big area on the canvas. But I cannot use the results made in small areas, such as ngers. Another insight we learned was about the light direction. Every participant mentioned that they especially liked the backlit results (P1, P2, P4, and P5). Participants generally liked the results from the left or right light direction. However, they found the light direction is not always consistent.

P4 mentioned: In some cases, the light suggestions confuse me because 80% of the shadows show the light coming from the left, while 20% seems to come from the right. In general, we conclude that further features must be devised to iterate the shadow suggestions depending on the level of detail (i.e., big chunks of volumetric shadows vs. small and detailed lines) and local light direction changes.

5.4 Discussion

We identi ed the following feature-level comments, raised during interviews, that are worth considering when building a more full- edged , nearly deployable tool:

About shadowing styles: We designed ShadowMagic to provide shadowing support for cel-styled shadowing and four light directions. We intentionally aimed for a narrow scope of binary shadowing to balance wide applicability and technical feasibility. A few participants mentioned the usefulness of binary shadowing as an essential technique that can be applied in nearly every shadowing scenario. At the same time, they mentioned that binary shadows can serve as a foundation to be mixed and layered with more complex shadowing techniques, such as multi-level shadows or gradient shadows. In layering di erent shadowing styles in one scene, they mentioned the importance of harmony. These capabilities would provide a more advanced and functional design that can practically bene t professionals, yet require more technical advances.

About light directions: Providing four light directions helped participants cover the majority of scenes that they will likely encounter. While less frequent, trendy shadowing styles are sometimes used for making unique compositions, such as zenith lighting, nadir lighting, or selective lighting [17]. Some participants discussed making plug-and-play designs for di erent shadowing styles can be a reasonable way of advancing the design. As the guidance input to the di usion model can be interpolated in its vector space [10], future work could be able to provide direction control of ner granularity than four directions.

About more customization on staging and attention guidance: ShadowMagic is designed to provide creative control for the parts that the artists are likely to be more engaged in (i.e., attention guidance, staging, or semantic regions where subtle change can make a dramatic e ect) while automating the parts that are perceived as tedious and repetitive (i.e., shadow landscaping, or repetitive semantic regions). However, one possible alternative version of the tool would even have AI make suggestions for those decisions, such as suggesting the light direction based on the staging needs.

About small but non-trivial features: Finally, there were some comments related to system usability. Several participants were satis ed with the overall quality of the shadows, but they explained occasional frustration when our models didn't accurately predict semantic regions. Additionally, while not expected, implementing a pressure-sensitive brush tool seems to be essential in real-world practice. ShadowMagic had a limited capability to realize this feature due to platform limitations.

6 IMPLICATIONS FOR DESIGN

Based on the ndings in S1 and S2, we rst discuss future research opportunities that reside in interactive, human-AI collaborative comic shadowing. Next, we discuss our re ection on how Intermediate Representations the foundational concept we leveraged in developing ShadowMagic can be applied di erently to bene t future design research into professional work ows within comics and beyond.

Beyond ShadowMagic The S1 and S2 results recognized the following problems in shadowing that require non-trivial research e ort:

- Shadow Style Variations and Combinations Creating binary shadows is a foundational part of shadowing, but it may only be the beginning for certain desirable outcomes. Participants in S1 and S2 explained the usefulness of providing shadow styles beyond binary. Future work can aim to support various shadowing styles, such as multi-level shadows and gradient shadows, with corresponding controls. Realizing this direction will require the creation of more generalizable shadow models that can adapt to a wide range of artistic styles and preferences. Another challenge involved in this direction is to develop a technical foundation that allows shadow layers with di erent styles to communicate with each other and make harmonized results.
- Plug & Play for Personalized Shadowing : In the S2 results, participants expressed a desire to create customized AI recommendations tailored to their preferences and settings. This direction is aligned with studies that show the usefulness of providing a capability for users to determine the AI's behavior [20, 21, 25, 40, 47]. One example is incorporating various lighting directions as discussed in Section 5.4. Another example is o ering AI-driven controls for directing attention or providing information. In future research, understanding the underlying elements of personalized AI recommendations and testing di erent design variations of each element can signi cantly bene t professionals.
- Interaction Modality: Direct Manipulation, Verbal Description, or Hybrid? : While generative AI technologies introduced interesting opportunities, the main entry that current users can leverage is through verbal description. Such a descriptive way of guiding interaction can impose daunting challenges for professionals in integrating the technology into their work ows, as the majority of them are familiar with direct manipulation designs where they can draw something in 2D and see results [41]. The interaction modality of direct manipulation and verbal description introduces challenges and opportunities for future human-AI collaboration researchers working on visual creativity support.

Beyond stage-by-stage Intermediate Representation. This work was largely inspired by the notion of Intermediate Representation [53] that explains how scoping input and output of AI by looking into a user's work ow stage-by-stage at, shadow, lighting can lead to a successful human-AI collaboration artifact. In particular, by splitting each stage in scoping AI-driven automation's input and output based on an expected user group's sequential work ow, the AI can provide the automation that a professional can adopt or revise [53]. Our work furthers this line of inquiry in terms of how designers can apply AI within each stage, in an *interaction-by-interaction* manner; i.e., in each stage, artists may have di erent levels of desire to accept AI suggestions. For instance, using AI-driven automation on landscaping or shadowing tedious semantic regions may likely be received well while important semantic regions or intricate staging representations may not. Intermediate representation, the notion of de ning the scope of AI automation by not only observing the stage-wise work ow but also investigating contextual preferences, resides within each stage.

Adopting Generative AI in Professional Workflows: While generative AI technologies introduced opportunities in various applications, they were often designed without considering professionals' work ows [30]. They were often considered as an opportunity to expand the creation experience to casual users lacking expertise. Supporting professionals is challenging in the sense that they are productivity-driven and have a speci c set of tools they are already familiar with. To develop a system for professionals, we rst tried to consider the user's work ow and where we can provide adequate support with the capability of generative models. In our case, shadowing was the target. Once we identi ed the target of support, we closely collaborated with practitioners to gather a dataset from the user's work context. We collected the data pairs consisting of pre-shadowing images and shadowed ones, with which we could ne-tune existing text-to-image di usion models to be aligned to the professionals' task contexts. While this approach introduced a model that outperformed the baseline model, ndings from S2A indicate that there is still room for improvement. One bottleneck for improving the model is collecting the dataset at scale, as creating datasets with experts can be expensive. We can potentially overcome this issue by data augmentation: we can generate image data with the current model checkpoint, lter high-quality data instances with practitioners, and further ne-tune the model with the ltered data. This approach would incur less cost, as evaluating shadow quality would require less human e ort than creating shadows.

7 CONCLUSION

This work introduced ShadowMagic, a human-AI collaboration system designed to support comic professionals in the shadowing stage of comic creation. Through studies with professionals, we identi ed opportunities for using ne-tuned AI models to generate initial shadow suggestions that can be selectively re ned through an interactive user interface. In an evaluation, professionals rated shadows from our AI model higher in quality than a baseline and provided positive feedback on ShadowMagic's modality for integrating AI assistance into their practices. Our work demonstrates how contextualized datasets can netune powerful generative AI for speci c professional work ows, pointing toward human-centered approaches that augment rather than replace human expertise across creative disciplines.

ACKNOWLEDGMENTS

This research was generously supported by NSF #2128867.

REFERENCES

[1] Adobe. 2021. Adobe Photoshop. https://www.adobe.com/products/photoshop/. Accessed: 2021-09-04.
[2] Zinat Ara, Hossein Salemi, Sungsoo Ray Hong, Yasas Senarath, Steve Peterson, Amanda Lee Hughes, and Hemant Purohit. 2024. Closing the Knowledge Gap in Designing Data Annotation Interfaces for AI-powered Disaster Management Analytic Systems. In *Proceedings of the 29th International Conference on Intelligent User Interfaces*. 405 418.
[3] Luca Benedetti, Holger Winnemöller, Massimiliano Corsini, and Roberto Scopigno. 2014. Painting with Bob: assisted creativity for novices. In *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM New York, NY, USA, New York, NY, USA, 419 428.
[4] Simon Breslav, Karol Szerszen, Lee Markosian, Pascal Barla, and Joëlle Thollot. 2007. Dynamic 2D patterns for shading 3D scenes. *ACM Trans. Graph.* 26, 3 (jul 2007), 20 es. https://doi.org/10.1145/1276377.1276402
[5] Hanqun Cao, Cheng Tan, Zhangyang Gao, Yilun Xu, Guangyong Chen, Pheng-Ann Heng, and Stan Z. Li. 2024. A Survey on Generative Di usion Models. *IEEE Transactions on Knowledge and Data Engineering* (2024), 1 20. https://doi.org/10.1109/TKDE.2024.3361474
[6] CELSYS. 2020. Clip Studio Paint Instruction Manual. https://www.clip-studio.com/site/gd_en/csp/userguide/csp_userguide/500_menu/500_menu_edit_autocolor.htm. Accessed: 2023-02-06.
[7] CELSYS. 2021. Clip Studio Paint. https://www.clipstudio.net/. Accessed: 2021-09-04.
[8] DaEun Choi, Sumin Hong, Jeongeon Park, John Joon Young Chung, and Juho Kim. 2024. CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 1055, 25 pages. https://doi.org/10.1145/3613904.3642794
[9] Minsuk Choi, Cheonbok Park, Soyoung Yang, Yonggyu Kim, Jaegul Choo, and Sungsoo Ray Hong. 2019. Aila: Attentive interactive labeling assistant for document classi cation through attention-based deep neural networks. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1 12.
[10] John Joon Young Chung and Eytan Adar. 2023. PromptPaint: Steering Text-to-Image Generation Through Paint Medium-like Interactions. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology* (<conf-loc>, <city>San Francisco</city>, <state>CA</state>, <country>USA</country>, </conf-loc>) *(UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 6, 17 pages. https://doi.org/10.1145/3586183.3606777
[11] John Joon Young Chung, Jean Y Song, Sindhu Kutty, Sungsoo Hong, Juho Kim, and Walter S Lasecki. 2019. E cient elicitation approaches to estimate collective crowd answers. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1 25.
[12] Daniel Cohen, Arie Kaufman, Reuven Bakalash, and Samuel Bergman. 1990. Real time discrete shading. *The Visual Computer* 6, 1 (1990), 16 27. https://doi.org/10.1007/BF01902626
[13] John W Creswell and Cheryl N Poth. 2016. *Qualitative inquiry and research design: Choosing among ve approaches*. Sage publications.
[14] Nicholas Davis, Chih-PIn Hsiao, Kunwar Yashraj Singh, Lisa Li, Sanat Moningi, and Brian Magerko. 2015. Drawing apprentice: An enactive co-creative agent for artistic collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*. 185 186.
[15] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Lisa Li, and Brian Magerko. 2016. Empirically studying participatory sense-making in abstract drawing with a co-creative cognitive agent. In *Proceedings of the 21st*

International Conference on Intelligent User Interfaces. 196 207.
[16] Mohamed Elasri, Omar Elharrouss, Somaya Al-Maadeed, and Hamid Tairi. 2022. Image generation: A review. Neural Processing Letters 54, 5 (2022), 4609 4646.
[17] Eridey. [n. d.]. Guide to Drawing Shadows. https://www.clipstudio.net/how-to-draw/archives/163236 Accessed: 2024-06-06.
[18] Krita Foundation. 2012. Krita. https://krita.org/ Accessed: 2021-08-11.
[19] Sébastien Fourey, David Tschumperlé, and David Revoy. 2018. A Fast and E cient Semi-guided Algorithm for Flat Coloring Line-arts. (2018).
[20] Yuyang Gao, Siyi Gu, Junji Jiang, Sungsoo Ray Hong, Dazhou Yu, and Liang Zhao. 2024. Going beyond xai: A systematic survey for explanation-guided learning. Comput. Surveys 56, 7 (2024), 1 39.
[21] Yuyang Gao, Tong Steven Sun, Liang Zhao, and Sungsoo Ray Hong. 2022. Aligning eyes between humans and deep neural network through interactive attention alignment. Proceedings of the ACM on Human-Computer Interaction 6, CSCW2 (2022), 1 28.
[22] Joao Paulo Gois, Bruno AD Marques, and Harlen Costa Batagelo. 2015. Interactive shading of 2.5 D models.. Graphics Interface 89 96.
[23] Je rey Heer. 2019. Agency plus automation: Designing arti cial intelligence into interactive systems. Proceedings of the National Academy of Sciences 116, 6 (2019), 1844 1850.
[24] Sungsoo Ray Hong, Jessica Hullman, and Enrico Bertini. 2020. Human factors in model interpretability: Industry practices, challenges, and needs. Proceedings of the ACM on Human-Computer Interaction 4, CSCW1 (2020), 1 26.
[25] Sungsoo Ray Hong, Jorge Piazentin Ono, Juliana Freire, and Enrico Bertini. 2019. Disseminating Machine Learning to domain experts: Understanding challenges and opportunities in supporting a model building process. In CHI 2019 Workshop, Emerging Perspectives in Human-Centered Machine Learning. ACM
[26] Matis Hudon, Mairéad Grogan, Rafael Pagés, Jan Ond°ej, and Aljo²a Smoli¢. 2019. 2DToonShade: A stroke based toon shading system. Computers & Graphics: X 1 (2019), 100003. https://doi.org/10.1016/j.cagx.2019.100003
[27] Jennifer Jacobs, Joel Brandt, Radomír Mech, and Mitchel Resnick. 2018. Extending manual drawing practices with artist-centric programming tools. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 1 13.
[28] Glenn Jocher, Alex Stoken, Jirka Borovec, NanoCode012, Ayush Chaurasia, TaoXie, Liu Changyu, Abhiram V, Laughing, tkianai, yxNONG, Adam Hogan, lorenzomammana, AlexWang1900, Jan Hajek, Laurentiu Diaconu, Marc, Yonghye Kwon, oleg, and Francisco Ingham. 2021. ultralytics/yolov5: v5.0 - YOLOv5-P6 1280 models, AWS, Supervise.ly and YouTube integrations (v5.0). https://doi.org/10.5281/zenodo.4679653
[29] Jaroslav K°ivánek and Mark Colbert. 2008. Real-time shading with ltered importance sampling. In Computer Graphics Forum, Vol. 27. Wiley Online Library, 1147 1154.
[30] Nahyun Kwon, Tong Steven Sun, Yuyang Gao, Liang Zhao, Xu Wang, Jeeeun Kim, and Sungsoo Ray Hong. 2024. 3DPFIX: Improving Remote Novices' 3D Printing Troubleshooting through Human-AI Collaboration Design. Proceedings of the ACM on Human-Computer Interaction 8, CSCW1 (2024), 1 33.
[31] Chen-Hsuan Lin, Jun Gao, Luming Tang, Towaki Takikawa, Xiaohui Zeng, Xun Huang, Karsten Kreis, Sanja Fidler, Ming-Yu Liu, and Tsung-Yi Lin. 2023. Magic3d: High-resolution text-to-3d content creation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 300 309.
[32] Joseph Lindley, Paul Coulton, and Miriam Sturdee. 2017. Implications for adoption. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 265 277.
[33] Gary Marcus, Ernest Davis, and Scott Aaronson. 2022. A very preliminary analysis of DALL-E 2. arXiv:2204.13807 [cs.CV]
[34] Justin Matejka, Michael Glueck, Erin Bradner, Ali Hashemi, Tovi Grossman, and George Fitzmaurice. 2018. Dream lens: Exploration and visualization of large-scale generative design datasets. In Proceedings of the 2018 CHI conference on human factors in computing systems. 1 12.
[35] Niloy J. Mitra and Mark Pauly. 2009. Shadow art. ACM Transactions on Graphics 28, 5, 156:1 156:7. https://doi.org/10.1145/1661412.1618502
[36] Changhoon Oh, Jungwoo Song, Jinhan Choi, Seonghyeon Kim, Sungwoo Lee, and Bongwon Suh. 2018. I Lead, You Help but Only with Enough Details: Understanding User Experience of Co-Creation with Arti cial Intelligence. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (<conf-loc>, <city>Montreal QC</city>, <country>Canada</country>, </conf-loc>) (CHI '18). Association for Computing Machinery, New York, NY, USA, 1 13. https://doi.org/10.1145/3173574.3174223
[37] Lohit Petikam, Ken Anjyo, and Taehyun Rhee. 2021. Shading Rig: Dynamic Art-directable Stylised Shading for 3D Characters. ACM Trans. Graph. 40, 5, Article 189 (sep 2021), 14 pages. https://doi.org/10.1145/3461696
[38] Ravi Ramamoorthi, Dhruv Mahajan, and Peter Belhumeur. 2007. A rst-order analysis of lighting, shading, and shadows. ACM Trans. Graph. 26, 1 (jan 2007), 2 es. https://doi.org/10.1145/1189762.1189764
[39] Johnny Saldaña. 2015. The coding manual for qualitative researchers. Sage.
[40] Aécio Santos, Sonia Castelo, Cristian Felix, Jorge Piazentin Ono, Bowen Yu, Sungsoo Ray Hong, Cláudio T Silva, Enrico Bertini, and Juliana Freire. 2019. Visus: An interactive system for automatic machine learning model building and curation. In Proceedings of the Workshop on Human-In-the-Loop Data Analytics. 1 7.
[41] Ben Shneiderman. 1983. Direct manipulation: A step beyond programming languages. Computer 16, 08 (1983), 57 69.
[42] Ben Shneiderman. 2020. Human-centered arti cial intelligence: Three fresh ideas. AIS Transactions on Human-Computer Interaction 12, 3 (2020), 109 124.
[43] Uriel Singer, Adam Polyak, Thomas Hayes, Xi Yin, Jie An, Songyang Zhang, Qiyuan Hu, Harry Yang, Oron Ashual, Oran Gafni, et al. 2022. Make-a-video: Text-to-video generation without text-video data. arXiv preprint arXiv:2209.14792 (2022).
[44] Sketchbook. 2021. Sketchbook - For everyone who loves to draw. https://www.sketchbook.com/. Accessed: 2021-09-04.
[45] Peter-Pike J Sloan, William Martin, Amy Gooch, and Bruce Gooch. 2001. The lit sphere: A model for capturing NPR shading from art. In Graphics interface, Vol. 2001. Citeseer, 143 150.
[46] Statista. 2024. Sales value of digital comics in Japan from scal year 2014 to 2023. https://www.statista.com/statistics/1119858/japan-e-comics-market-size/. Accessed: 2024-07-23.
[47] Tong Steven Sun, Yuyang Gao, Shubham Khaladkar, Sijia Liu, Liang Zhao, Young-Ho Kim, and Sungsoo Ray Hong. 2023. Designing a direct feedback loop between humans and Convolutional Neural Networks through local explanations. Proceedings of the ACM on Human-Computer Interaction 7, CSCW2 (2023), 1 32.
[48] Chaitanya Krishna Suryadevara. 2020. Generating free images with OpenAI's generative models. International Journal of Innovations in Engineering Research and Technology 7, 3 (2020), 49 56.
[49] Daniel Sýkora, John Dingliana, and Steven Collins. 2009. LazyBrush: Flexible Painting Tool for Hand-drawn Cartoons. Computer Graphics Forum 28, 2 (2009), 599 608. https://doi.org/10.1111/j.1467-8659.2009.01400.x arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2009.01400.x
[50] TaiZan. 2021. Petalica Paint: AI-powered Automatic Colorization. https://petalica-paint.pixiv.dev/index_en.html Accessed: 2023-02-06.
[51] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. 2021. Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data. In 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). IEEE, Montreal, BC, Canada, 1905 1914. https://doi.org/10.1109/ICCVW54120.2021.00217
[52] Naver Webtoon. 2022. Webtoon AI Painter. https://ai.webtoons.com/painter Accessed: 2023-02-06.
[53] Chuan Yan, John Joon Young Chung, Yoon Kiheon, Yotam Gingold, Eytan Adar, and Sungsoo Ray Hong. 2022. FlatMagic: Improving Flat Colorization through AI-Driven Design for Digital Comic Professionals. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 1 17.
[54] Chengzhi Zhang, Weijie Wang, Paul Pangaro, Nikolas Martelaro, and Daragh Byrne. 2023. Generative Image AI Using Design Sketches as input: Opportunities and Challenges. In Proceedings of the 15th Conference on Creativity and Cognition (<conf-loc>, <city>Virtual Event</city>, <country>USA</country>, </conf-loc>) (C&C '23). Association for Computing Machinery, New York, NY, USA, 254 261. https://doi.org/10.1145/3591196.3596820
[55] Chenshuang Zhang, Chaoning Zhang, Mengchun Zhang, and In So Kweon. 2023. Text-to-image Di usion Models in Generative AI: A Survey. arXiv:2303.07909 [cs.CV]
[56] Lvmin Zhang, Jinyue Jiang, Yi Ji, and Chunping Liu. 2021. SmartShadow: Artistic Shadow Drawing Tool for Line Drawings. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). 5391 5400.
[57] Lvmin Zhang, Chengze Li, Edgar Simo-Serra, Yi Ji, Tien-Tsin Wong, and Chunping Liu. 2021. User-Guided Line Art Flat Filling with Split Filling Mechanism. In IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
[58] Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Di usion Models. In 2023 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, Paris, France, 3813 3824. https://doi.org/10.1109/ICCV51070.2023.00355

[59] Lvmin Zhang, Edgar Simo-Serra, Yi Ji, and Chunping Liu. 2020. Generating digital painting lighting e ects via RGB-space geometry. ACM Transactions on Graphics (TOG) 39, 2 (2020), 1 13.
[60] Qingyuan Zheng, Zhuoru Li, and Adam Bargteil. 2020. Learning to shadow hand-drawn sketches. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 7436 7445.
[61] Mi Zhou, Vibhanshu Abhishek, Timothy Derdenger, Jaymo Kim, and Kannan Srinivasan. 2024. Bias in Generative AI. arXiv:2403.02726 [econ.GN]

8 APPENDIX

Tab 1 presents three high-resolution examples of comic at and line images in the rows, along with corresponding shadow variations from four di erent light directions in the columns. Additional low-resolution examples are available on GitHub[5]

[5]https://github.com/Amrita537/ShadowMagicUIST24