

COMFYGEN: PROMPT-ADAPTIVE WORKFLOWS FOR TEXT-TO-IMAGE GENERATION - SUPPLEMENTARY MATERIALS

Anonymous authors

Paper under double-blind review

1 ADDITIONAL RESULTS

In figs. 1 to 4 we show additional images generated using our method. figs. 1 and 2 show images generated using ComfyGen-IC for workflow selection, while figs. 3 and 4 show images created with flows from ComfyGen-FT.

These results showcase the generalizability of our approach to a wide range of prompts, from portraying unique artistic styles to portraying photo-realistic and imagined scenes.

2 ADDITIONAL IMPLEMENTATION DETAILS

2.1 COMFYGEN-IC

Here, we provide additional details on our ComfyGen-IC approach, including the prompts used to create the class labels, and those used for workflow selection.

To generate the set of category labels, we prompted Claude Sonnet 3.5 with the following:

“Given the following list of image prompts, generate a list of 20 labels that describe the key elements, styles, or themes of the images these prompts might produce. Focus on general descriptors like ‘people’, ‘photo-realistic’, ‘photo-artistic’, ‘nature’, ‘abstract’, etc.

Here are the prompts: [prompts]”

This produced the following list of category labels: ‘People’, ‘Photo-realistic’, ‘Photo-artistic’, ‘Fantasy’, ‘Sci-fi’, ‘Horror’, ‘Anime’, ‘Abstract’, ‘Surreal’, ‘Cyberpunk’, ‘Steampunk’, ‘Gothic’, ‘Digital art’, ‘Portrait’, ‘Nature’, ‘Landscape’, ‘Wildlife’, ‘Urban’, ‘Cosmic’, ‘Underwater’.

To assign labels to each text-to-image training prompt, we prompted the LLM with the following:

”Given the following image prompt and list of labels, select the most relevant labels that describe the key elements, styles, or themes of the image this prompt might produce. Provide only the selected labels, separated by commas.

Image prompt: [prompt]

Available labels: [labels]

Selected labels:”

The LLM assigned an average of 4.96 ± 1.33 labels per prompt, with a maximum of 10. For 2 prompts (0.4% of the training set), the LLM failed to assign any label and so they were discarded.

The full table of flows and scores amounted to roughly 80,000 tokens. To further reduce this, we filter out all flows which achieved a score below the median for every single category. This filters out a total of 125 (or roughly 40% of the flows).

The filtered table serves at the context which we provide to the LLM with the following flow selection prompt:

“[context]

Please classify the following prompt into one of the flows mentioned above:

[prompt]

Provide the flow ID and a brief explanation for your classification.”

As part of our investigation, we also ablated the use of 30 labels and different LLM labeling prompts, but were unable to improve results over this baseline.

2.2 COMFYGEN-FT

Our ComfyGen-FT models were finetuned using 4-bit quantization, with a batch size of 8 and 4 gradient accumulation steps. We used an AdamW optimizer with weight decay of 0.01. We further use 5 warmup steps, and a linear LR scheduler.

All experiments were conducted on a single NVIDIA H100 GPU.

During training and inference, we use the following instruction when prompting the LLM:

“Below is a prompt that describes an image a user wants to generate, and a numerical score describing the quality of an image. Please output a ComfyUI workflow in json format that will create an image with this score when given the prompt.

>>> Prompt:

[prompt]

>>> Score:

[score]

>>> Flow:”

3 IMAGE PROMPTS

Below, we provide the prompts used to synthesize all the large-figure images in the core paper and in section 1. In each table, we point to the image location using the figure number and a grid position. The grids are 0 indexed, and list the image location from top-to-bottom and left-to-right.

4 LIST OF MODELS APPEARING IN OUR WORKFLOWS

Recall that we augment human created flows by randomly swapping the models that they use. Below is a list of all models that appear in our dataset. All base models are on the scale of SDXL and below.

5 ADDITIONAL USER STUDY DETAILS

In fig. 5 we show an example question from our user study form. To create the forms, we randomly sampled 120 prompts for ComfyGen-FT and 120 prompts for ComfyGen-IC. We divided each group of 120 questions into 6 blocks of 20, and matched each block to a baseline. From this matching, we created 240 questions that pit one of our methods against one of the baselines. We discarded all questions for which at least one of the images contained nudity or overly sexualized content, leaving us with a total of 231 questions. Each question was initially answered by at least 3 users, but we discarded all responses from users that picked at least 2 answers where the chosen image had no relation to the input prompt (*e.g.*, a scenario where the prompt describes a cigar, one image shows a cigar, and the user picked an image that shows an empty street).



Figure 1: Curated images generated using ComfyGen-ICL. The list of prompts is available in the supplementary.

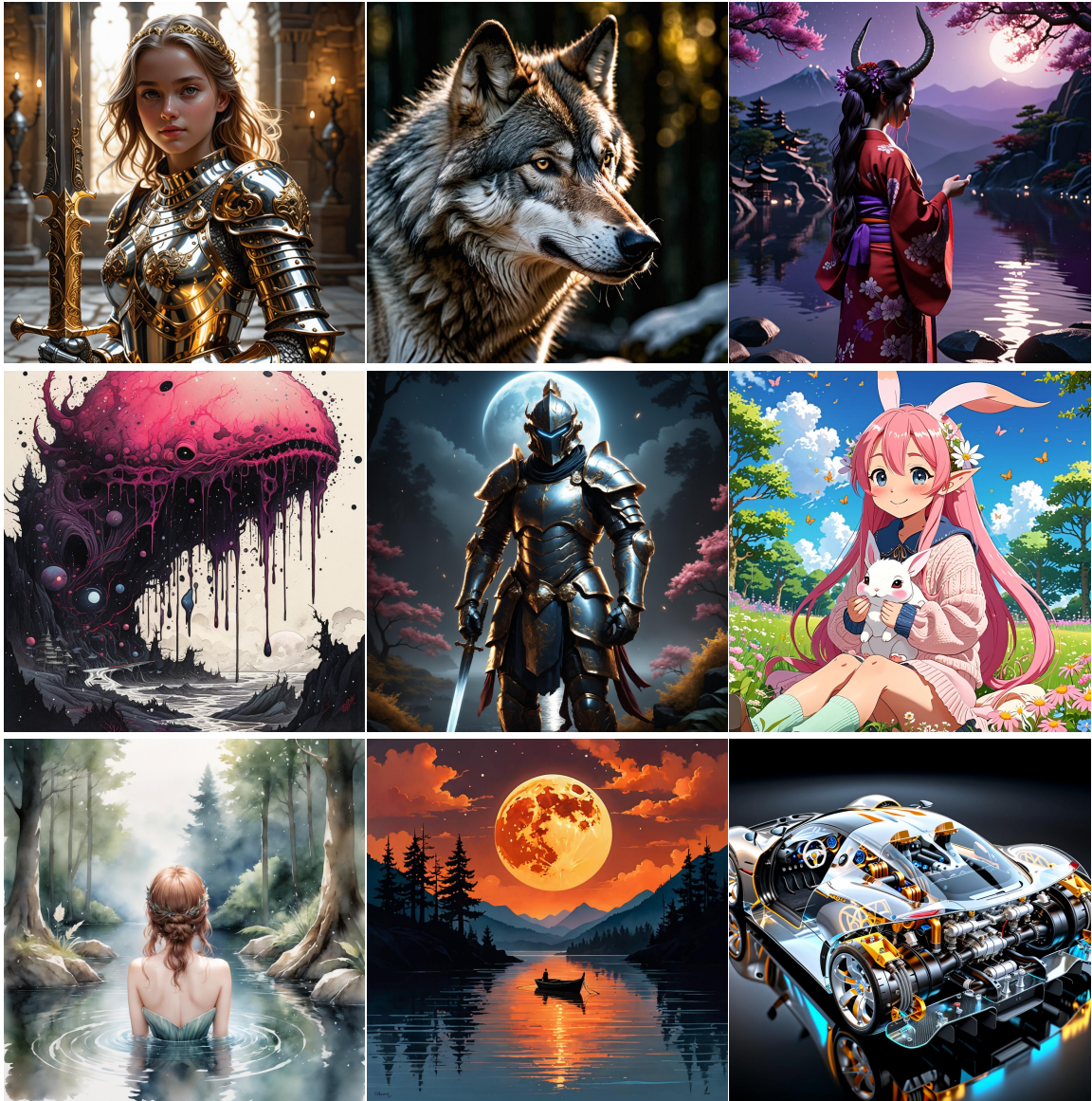


Figure 2: Curated images generated using ComfyGen-ICL. The list of prompts is available in the supplementary.

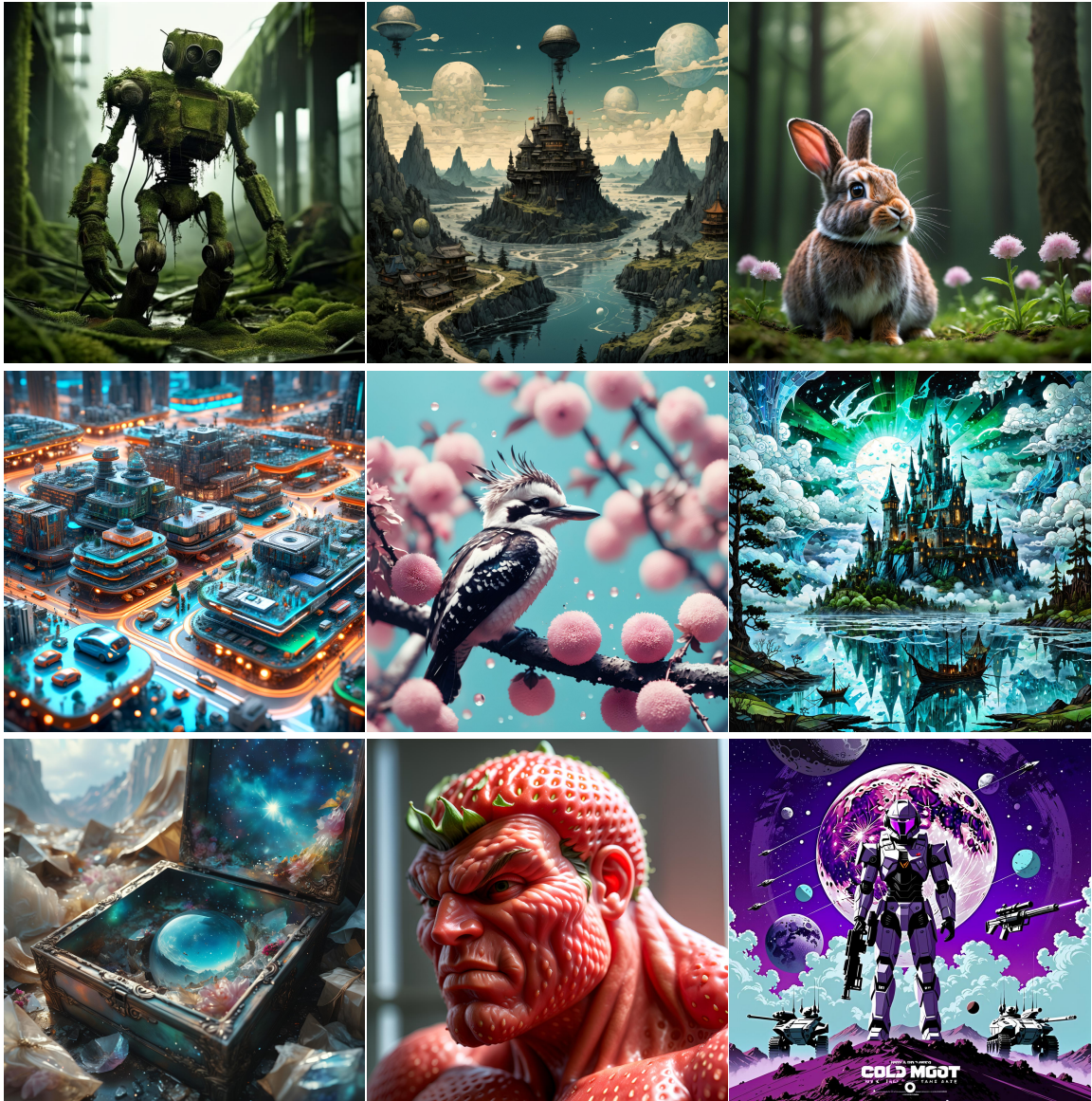


Figure 3: Curated images generated using ComfyGen-FT. The list of prompts is available in the supplementary.



Figure 4: Curated images generated using ComfyGen-FT. The list of prompts is available in the supplementary.

Figure	Prompt
Fig. 3 (0, 0)	“The entire observable universe in a single bottle, Dreamlike, Surreal landscapes, Mystical creatures, Twisted reality, Surreal still life, Extremely Detailed Oil Painting, glow effects, god rays, Hand drawn, render, 8k, cartoon, octane render, cinema 4d, blender, dark, atmospheric 4k ultra detailed, cinematic sensual, Sharp focus, humorous illustration, big depth of field, Masterpiece, colors, 3d octane render, 4k, concept art, trending on artstation, hyper realistic, Vivid colors, extremely detailed CG unity 8k wallpaper, trending on ArtStation, trending on CGSociety, Intricate, High Detail, dramatic, masterpiece, best quality, ultra-detailed, unreal engine, octane render, HDR”
Fig. 3 (0, 1)	“cottoncandy, In a whimsical candy kingdom, there stands a unicorn crafted from marshmallows, exuding childlike wonder and an air of mystery. Its body is as white and fluffy as a cloud, embodying the spirit of a playful sprite emerging from sweet dreams. Its head features a radiant rainbow candy gem that forms the iconic horn, shining with seven-colored brilliance, seemingly illuminating every innocent aspiration. The eyes are crafted from translucent hard candies, filled with tender yet inquisitive expressions. The mane consists of multi-hued strands of marshmallow fluff, each one dazzling like a rainbow, swaying gently in the breeze, whispering enchanting tales. Its tail resembles a billowy marshmallow cloud, soft and dreamy; when it gallops joyfully, it leaves behind a trail of vibrant marshmallow magic. Its hooves resemble those cast from creamy white chocolate, solid yet sweet, firmly planted on the ground made of colorful candies. This marshmallow unicorn stands tall, serving as both a tangible representation of children’s naive fantasies and a mystical presence brimming with magical allure within the fairy tale world, UHD, Extreme detail, natural light, volume light, fantasism, professional color, professional composition”
Fig. 3 (0, 2)	“Masterpiece double exposure of a girl silhouette blending with monochrome apocalypse aftermath and a colorful natural landscape in the underlying backdrop, sharp contrast, detailed crisp lines, in focus”
Fig. 3 (1, 0)	“A photo of a large man with beard riding a vespa wearing loriseg armor and helmet, red tunic”
Fig. 3 (1, 1)	“A hyper-realistic, ultra-detailed painting featuring a fantastical landscape with a village at the end of time. The scene has a perfect composition and chiaroscuro inspired by Rembrandt, with concept art elements by Mariusz Lewandowski, Louis Aston Knight, Karine Eibatova, John Howe, Jessica Rossier, Jarosław Janikowski, Inessa Garmash, Henriette Ronner-Knip, and Harumi Hironaka. The painting is vibrant, beautiful, painterly, detailed, and textural. Watercolor painting, masterpiece, best quality, hyper detailed, ultra realistic, 32k, RAW photo, landscape, fantasy, village at the end of time, perfect composition, chiaroscuro by Rembrandt, concept art, vibrant, beautiful, painterly, detailed, textural, artistic, Rembrandt, Mariusz Lewandowski, Louis Aston Knight, Karine Eibatova, John Howe, Jessica Rossier, Jarosław Janikowski, Inessa Garmash, Henriette Ronner-Knip, Harumi Hironaka., extremely detailed, swirling ink”
Fig. 3 (1, 2)	“true masterpiece, masterpiece cinematic lighting, cinematic shot from below, extremely detailed, high detail, hires textures, incredibly detailed, intricate details, photorealism, intricately designed, sea storm up inside a large transparent glass ball, drops outside the ball rainfall on floor background, HD32k, focus”
Fig. 3 (2, 0)	“no humans, scenery, closeup, branch, tree, leaf, nature, rain, outdoors, depth of field, droplets”
Fig. 3 (2, 1)	“fine art, oil painting, best quality, dark tales, illustration, each color adds depth, and the entire piece comes together to create a breathtaking spectacle of motion and tranquility., while the ball is adorned with an array of stripes in various hues. the figurine, while her right hand delicately holds a small, epic splash cover art in the van gogh style, starry sky, dan mumford, andy kehoe, 2d, flat, delightful, vintage, art on a cracked paper, patchwork, stained glass, fairytale, storybook detailed illustration, cinematic, ultra highly detailed, tiny details, beautiful details, mystical, luminism, vibrant colors, complex background”
Fig. 3 (2, 2)	“a ral-sun flying flower, dusty background, epic, heroic, bokeh, sharp detailed, hyperrealistic, amazing, macro photo, god rays, volumetric”

Table 1: List of prompts used for Fig. 3 in the core paper.

Figure	Prompt
Supp. Fig. 1 (0, 0)	“2D anime style, elegant girl floating in a vortex of city lights and digital elements, deep and expressive eyes, vibrant blue and purple colors, intricate details, magical and futuristic ambiance, highly detailed background with swirling buildings and lights, complex and dynamic scene”
Supp. Fig. 1 (0, 1)	“A close-up portrait of an alien being, with luminescent skin and eyes that hold centuries of wisdom, set against the backdrop of their advanced, technology-filled habitat., 90th photos from photo album, film, vintage style, flash from camera, hyper realistic, fine textures, high quality textures of materials, volumetric textures, natural textures, natural colors, correct white balance, color correction, dehaze, clarity”
Supp. Fig. 1 (0, 2)	“masterpiece, best quality, high quality, intricate, absurdres, very aesthetic, no humans, landscape, outdoors, mountain tops, wind, windy, wind lines, clouds, above clouds, cliff, wind magic, aurora, ultra wide angle shot, cinematic style, highly detailed, extremely detailed, sharp detail, majestic, shallow depth of field, movie still, soft light, circular polarizer, colorful, wallpaper, professional illustration, anime”
Supp. Fig. 1 (1, 0)	“adventurous cute funny lizard, donned a tiny explorer’s hat and carried a mini backpack, walk in wild colorfully jungle, detailed scales, by Jean-Baptiste Monge, Gilles Beloeil, Tyler Edlin, Marek Okon, Pixar, 8k, album art, comic style, golden ratio, perfect composition, a masterpiece, trending on artstation, extreme close up, shot from below . High dynamic range, vivid, rich details, clear shadows and highlights, realistic, intense, enhanced contrast, highly detailed”
Supp. Fig. 1 (1, 1)	“Amazing detailed photography of a cute adorable samurai kitten holding Katana with 2 paws, Cherry Blossom Tree petals floating in air, high resolution, piercing eyes, lifelike fur, Anti-Aliasing, FXAA, De-Noise, Post-Production, SFX, insanely detailed and intricate, hypermaximalist, elegant, ornate, hyper realistic, super detailed, noir coloration, serene, 16k resolution, full body”
Supp. Fig. 1 (1, 2)	“bg_imgs, portrait, wallpaper, colorful, highres, absurdres, huge filesize, fantasy, foreshortening, black dress, Extremely gorgeous metal style, Metal crown with ornate stripes, Various metals background, Sputtered molten iron, floating hair, Hair like melted metal, Clothes made of silver, Clothes with gold lace, flowing gold and silver, everything flowing and melt, flowing iron, flowing silver, lace flowing and melt, best quality, masterpiece, illustration, an extremely delicate and beautiful, extremely detailed, CG, unity, 8k wallpaper, Amazing, finely detail, masterpiece, best quality, official art, extremely detailed CG unity 8k wallpaper, absurdres, incredibly absurdres, huge filesize, ultra-detailed, highres, extremely detailed, beautiful detailed girl, extremely detailed eyes and face, beautiful detailed eyes, light on face”
Supp. Fig. 1 (2, 0)	“4n1v3rs3, 4n1v3rs3, 2, Create an art deco inspired illustration of A scenic lighthouse perched on a rocky coastline, overlooking the turquoise waters of Cala Ratjada. Include vibrant pastel colors, sleek lines, and a retro summer atmosphere. The style should be reminiscent of vintage travel posters with a modern twist”
Supp. Fig. 1 (2, 1)	“lighting Style, dim light, low light, dramatic light, partially covered in shadow, award winning photography, RAW photo, Hyperrealistic, beautiful African woman, wearing traditional outfit, vibrant colors, intricate patterns, detailed textures, soft natural lighting, ethereal glow, participating in a traditional ceremony, serene and dignified expression, traditional headwrap, detailed jewelry, lush green background, realistic shadows, fine details in fabric and skin, ultra-quality, cultural richness, ceremonial setting, glowing skin, traditional makeup, colorful beads, elegant pose, cultural heritage, ceremonial attire Detailed natural skin”
Supp. Fig. 1 (2, 2)	“perfect flower in the nature”

Table 2: List of prompts used for Fig. 1 in the supplementary.

Figure	Prompt
Supp. Fig. 2 (0, 0)	“amazing quality, masterpiece, best quality, hyper detailed, ultra detailed, UHD, perfect anatomy, in castle, girl knight, holding a sword, dazzling, transparent, polishing, 1 arm up, waving sword, gold armor, glowing armor, glowing eyes, full armor, shine armor, dazzling armor, extremely detailed”
Supp. Fig. 2 (0, 1)	“closeup award winning photo of wolf, perfect environment, extremely detailed, dark shot”
Supp. Fig. 2 (0, 2)	“detailed, aesthetic, 8k unreal engine photorealism, ethereal lighting, purple, nighttime, darkness, surreal art, fantasy, glowing, night, dark environment, AyameNewYears, horns, long hair, side, red kimono, floral print, hair flower, sash, haori, scenery, ink, mountains, water, trees, full body, reflection, light, arm behind back, sandals, from side”
Supp. Fig. 2 (1, 0)	“Japanese Ink Drawing, Ink Dripping Drawing, horror-themed space-themed Surrealism, cartoon style, concept art surreal, anatomy, anatomical drawings, human body, fantastical organs, space demon, sharp teeth, wide grimace, masterpiece, best quality, highly detailed, sharp focus, dynamic lighting, vivid colors, texture detail, particle effects, storytelling elements, narrative flair, 16k, UE5, HDR, subject-background isolation . digital artwork, illustrative, painterly, matte painting, highly detailed, expressive, dramatic, organic lines and forms, dreamlike and mysterious, Surrealism . cosmic, celestial, stars, galaxies, nebulae, planets, science fiction, highly detailed . eerie, unsettling, dark, spooky, suspenseful, grim, highly detailed, ink drawing, dripping ink, ink drawing, inkwash, Japanese cartoon style, japanese torii”
Supp. Fig. 2 (1, 1)	“a mystical, and high-resolution image of an anthropomorphic Elysium Knight in Radiant Darkness, The character should be depicted in a manga cover style with wealthy portraiture and poster art elements. Volumetric lighting, The image should feature rich colors and high contrast, focusing on the best quality, official art, and a beautiful and aesthetic appearance. wearing Sci-fi clothes and enhanced Astronaut suit, and standing in a cowboy shot pose. The character should have tattoos, muscular abs, shoulder armor, mystic Ethereal moon forest lake in the background. Emphasize a samurai theme with a detailed background and floral elements.”
Supp. Fig. 2 (1, 2)	“ova, background 2D anime, anime screencap, a pastel-colored character with long, flowing pink hair, large horns, and elf-like ears, wearing a cozy cable-knit sweater. The character is sitting in a beautiful, flower-filled meadow, surrounded by butterflies and small woodland creatures. She is gently petting a small bunny on her lap, smiling softly. The background is detailed with various types of flowers, trees, and a bright blue sky. The lighting is natural and bright, enhancing the cheerful and peaceful atmosphere. The camera shot is a medium shot, capturing the character and the detailed meadow.”
Supp. Fig. 2 (2, 0)	“watercolor art painting, watercolor, 1 forest queen, solo, sitting in the water, vintage anime aesthetic, back view, illustration, turn back her face, mysital, detailed eyes”
Supp. Fig. 2 (2, 1)	“the image portrays a tranquil scene of a boat floating gently on the water, surrounded by an expansive landscape. the moon, full and glowing with a warm, reddish orange hue, casts a mystical ambiance over the entire scene. its reflection shimmers off the surface of the water, adding to the serene atmosphere. in the distance, mountains loom under the moon’s soft glow, their peaks partially obscured by the low hanging clouds. they appear majestic yet gentle, as if watching over the peaceful night below. trees line the shore in the foreground, their silhouettes faintly visible against the darkening sky. this picturesque setting evokes a sense of calm and tranquility, inviting viewers to take a moment and appreciate the beauty of nature. it is a symphony of colors and shapes, each element working harmoniously together to create a visually captivating and emotionally soothing composition.”
Supp. Fig. 2 (2, 2)	“spectacular digital rendering of a rear view of a transparent hyper car revealing internal mechanical components such as engine, car chassis, suspension, and internal wiring, detailed textures, accurate lighting and shadows, 8k quality, intricate patterns, high-definition, glossy finish, vivid reflections, perfect lighting, showroom”

Table 3: List of prompts used for Fig. 2 in the supplementary.

Figure	Prompt
Supp. Fig. 3 (0, 0)	“A photo of a abyssal destroyed robot covered in moss, post apocalyptic city, lush overgrowth, by Luis Royo, by Greg Rutkowski, dark, gritty, intricate, volumetric lighting, volumetric atmosphere, concept art, cover illustration, octane render, trending on artstation, 8k, dynamic pose”
Supp. Fig. 3 (0, 1)	“by Mattias Adolfsson and Satoshi Kon, hyper realistic medium full shot photo of a otherworldly landscape, oppulent, glamorous, masterful, poster art, bold lines, hyper detailed, expressive, award winning, dark limited color palette, high contrast, depth of field, intricate details, masterpiece, best quality, rim lighting, looking at viewer, dynamic pose, wide angle panoramic view”
Supp. Fig. 3 (0, 2)	“close up photo of a rabbit, forest, haze, halation, bloom, dramatic atmosphere, centred, rule of thirds, 200mm 1.4f macro shot”
Supp. Fig. 3 (1, 0)	“closeup of a circuit board city with very small futuristic cars, small flying vehicles, small robots and small humans, ultra hd, realistic, vivid colors, highly detailed, UHD drawing, pen and ink, perfect composition, beautiful detailed intricate insanely detailed octane render trending on artstation, 8k artistic photography, photorealistic concept art, soft natural volumetric cinematic perfect light”
Supp. Fig. 3 (1, 1)	“Cross-Processing by Artur Amijewski and Rodelio Astudillo, award winning, kookaburra, aesthetic of symbolism with bubbling atmosphere, skybox / skydome, well-defined edges, creative tour de force with meticulous details, morganite pink and mauve colors”
Supp. Fig. 3 (1, 2)	“detailed ink, pen and ink, mail art, best quality, detailed epic ice transparent ethereal otherworldly ghost castle in the blue sky, clouds, smoke, fog, detailed landscape, ghost figures, lake, boat, green forest, detailed flying dragon at the sky, detailed scales, warm lights, glittering, Craola, Dan Mumford, Andy Kehoe, 2d, flat, art on a cracked paper, patchwork, stained glass, cute, adorable, fairytale, storybook detailed illustration, cinematic, ultra highly detailed, tiny details, beautiful details, mystical, luminism, vibrant colors, complex background”
Supp. Fig. 3 (2, 0)	“ethereal fantasy concept art of masterpiece, best quality, RAW macro photo of just some garbage that someone put in a box/frame . magnificent, celestial, ethereal, painterly, epic, majestic, magical, fantasy art, cover art, dreamy”
Supp. Fig. 3 (2, 1)	“detailed realistic close up of a strawberry shaped like a muscular man, sitting, natural light”
Supp. Fig. 3 (2, 2)	“minimalist, cinematic, movie poster, cold colors, purple theme, standing, weapon, outdoors, sky, cloud, gun, moon, helmet, robot, ground vehicle, motor vehicle, multiple planets, planet, spacecraft”

Table 4: List of prompts used for Fig. 3 in the supplementary.

Figure	Prompt
Supp. Fig. 4 (0, 0)	“Once upon a time, organized theory, systematic scientific papers, elegant, figuration, sharp lines, black and white colors, beautiful, trending on artstation, volumetric lighting, by Guy Denning, no text, colorized drawings, black and white, artistic”
Supp. Fig. 4 (0, 1)	“celestial promenade, the horror of stars, twilight sunshine, heaven help us, cityscape masterpiece, realistic, best quality, cosmic horror, bright horror”
Supp. Fig. 4 (0, 2)	“highly detailed and hyper realistic photo, by Alena Aenami, by Archibald Thorburn, by Daniele Afferni, a breathtaking otherworldly landscape with psychedelic colors, in the style of monochromatic silhouette reflection, limited dark palette, unusual dark colors, faded colors, atmospheric haze, highly dramatic cinematic lighting, motion blur, film grain, professional, excellent composition, finest details, maximized details, ultimate detail level, masterpiece, best quality”
Supp. Fig. 4 (1, 0)	“refreshing, vibrant glowing coconut juice drink, dew drops, refreshing, in the style of a product hero shot in motion, dynamic magazine ad image, photorealism, sleep and mystical elements around the background”
Supp. Fig. 4 (1, 1)	“The art of Origami, Paper folding, Swan on a lake, Amazing colours, Intricate details, Painstaking Attention to Details, UHD”
Supp. Fig. 4 (1, 2)	“The devil with glowing eyes ingenious opus magnum by Michael Vincent Manalo and Elvira Vigna, pastel oil painting, high key with glistening light, cosmic and direful and dramatic atmosphere, zestful quinacridone nickel azo gold and electric blue color swatch, raw details, crisp details, bokeh lights, analog photo, full body, tall height, against the background volcano mouth in the background, blue sky, epic clouds, oil on canvas, art by Jeremy Mann, god rays, dramatic light, art by Eduard Wilhelm Pose and Master of the Holy Blood and Ivan Bilibin, oxygen-rich air and sheltered atmosphere, snowy, aesthetic of hard - edge painting, symmetry and balance, pioneering unparalleled masterwork with superior details, skin texture, masterpiece, top quality, best quality, official art, highest detailed, atmospheric lighting, cinematic composition, complex multiple subjects, 4k HDR vaporwave style, cyberpunk, vibrant, neon colors, highly detailed, Leica Q2 with Summilux 35mm f/1.2 ASPH, clear face, Ultra High Resolution, wallpaper, 8K, Rich texture details, hyper detailed, detailed eyes, detailed background, dramatic angle”
Supp. Fig. 4 (2, 0)	“In a lush, tropical rainforest, a vibrant, exotic bird like a parrot perches amidst a kaleidoscope of flowers, with dappled sunlight filtering through the dense canopy. This scene is captured in the style of Marco Lumiere, known for his vivid and lively color palettes and a slightly impressionistic touch, highlighting the vitality of nature and the majestic beauty of the bird in its natural habitat.”
Supp. Fig. 4 (2, 1)	“wooden arch bridge, river, crowd, mount fuji, boats, cherry tree, scenery, outdoors, best quality, highly detailed”
Supp. Fig. 4 (2, 2)	“zentangle Flickering grass beneath a lone smart light bulb that casts an eerie glow in the otherwise serene sunlit meadow, as if waiting for something to emerge from the shadows. . intricate, abstract, monochrome, patterns, meditative, highly detailed”

Table 5: List of prompts used for Fig. 4 in the supplementary.

Base models and refiners:

- AethervseLightning v10
- AlbedobaseXL v13
- AnimateXL v30
- AnythingXL
- crystalClearXL ccXL
- DreamshaperXL turboDpmppSDEKarras
- EnvyhyperdriveXL v10
- GleipnirV0.3
- JibMixXL v9 “BetterBodies”
- JuggernautXL v9 Rdphoto2Lightning
- LeosamsHelloworldXL v70
- Proteus v03
- RealismEngineSDXL v10
- RealvisXL v40 BakedVAE
- RealvisXL v40 LightningBakedVAE
- SDXL Base 1.0 0.9VAE
- SDXL Base 1.0
- SDXL Refiner 1.0 0.9VAE
- SDXL Refiner 1.0
- SDVN7 - NijiStyleXL v1
- SSD-1B
- TurbovisionXL SuperFastXL V431BakedVAE
- Stable Cascade
- Pixart-Σ

LoRAs and embeddings:

- easynegative
- bad-hands-5
- nfixer
- Add-Detail XL
- EpicF4nta5yXL
- AnimeTarot
- JuggCineXL2
- LCM LoRA SSD-1B
- LCM LoRA SDXL
- LogoRedmond
- MJ52 v2.0
- MJ52
- PerfectEyesXL
- Pixel-Art-XL v1.1
- Ral-Dissolve-SDXL
- SDXL Glass
- SDXLFaetastic v24
- Sinfully Stylish SDXL

- Werewolf SDXL
- WowifierXL v2
- XL more art-full-beta1

Super resolution and face restoration models:

- 4x NMKD Superscale - SP 178000 G
- 4x UltraSharp
- RealESRGAN x2 plus
- codeformer
- GFPGAN v1.4

VAEs:

- SharpSpectrum VAEXL
- SDXL VAE fp16 fix
- SDXL VAE

Below is a text prompt describing an image that we want to generate, and two text-to-image model outputs corresponding to this prompt. Please select the result that you prefer, taking into account both the quality of the image and its adherence to the text prompt:

Text prompt: "a magical map leading to a mysterious land, torn and worn, sitting on a wizard's desk, spilled multicolored ink, bottles, flasks, high quality, imagination, 8k, fantasy art"



☐ A



☐ B

Figure 5: An example question from our user study.