



Differential Diffusion: Giving Each Pixel Its Strength

Supplementary Material

Eran Levin¹  and Ohad Fried² 

¹Tel Aviv University, Israel

²Reichman University, Israel

We present additional results (Figures 1, 4 to 6, 9, 10, 17, 18 and 21), extend the discussion on experiments from the main paper (Figures 19 and 20), provide additional experiments that show the usability of our framework (Figure 24), and specify additional technical details for reproducibility (Section 12).

We provide further details on the “skipping” optimization in Section 1. We provide an extended discussion on strength fans in Section 2. We provide additional comparisons for soft-inpainting in Figure 1. In figure 10 in the main paper we show a comparison between our method and other mask-based methods. In Figure 4 we show that these results have not been caused by outlier seeds. We show that for different seeds, our method produces better results. We show more examples of the control that change maps introduce in Figures 5 to 7.

In Section 5.1 of the main paper we discuss the limitations of traversing multiple times across the latent space. We show the degradation in the image quality in Figure 8. We provide additional comparisons to the baselines described in this section in Figure 9. In Section 5.2 of the main paper we use 1,000 pairs of images to measure the applied change map. We show an ablation to this amplification in Figure 10, where we compare it to a single pair measurement, and also show the spatial bias of the LPIPS similarity maps [ZIE*18].

In the main paper, we demonstrate that DiffEdit [CVSC22] fails to comply with the given edit. We strengthen this demonstration with additional outputs in Figure 11. We show that our algorithm is compatible with many samplers in Figure 16. We show more examples of edits in Figure 17. In the main paper, we describe the usage of Gaussian blur for achieving the soft inpaint application. We provide additional examples at Figure 18, we also visualize the effect of different radii at Figure 12. In the main paper, we have outlined the structure of the user study; in Figure 20 we provide representative questions from the study. We created demos for our framework, and it has been integrated into several software suites by the community as can be seen in Section 10.

1. Running Time of Skipping

We measured our model with and without skipping, using gradient maps with different minimum values, averaging results over

100 different inference processes. In the main paper, we detail the modification of the original algorithm to integrate the skipping optimization; the explicit skipping algorithm is in Algorithm 1, and we visualize the runtime as a function of the minimum value in the change map in Figure 19. A clear linear correlation exists between these two variables. Our measurements (Table 1) demonstrate the utility of the skipping optimization, by saving up to 89 percent of the running time.

2. Strength Fan and Other visualization tools

The main paper demonstrates the ability of our framework to create a “strength fan”, a tool to estimate the effects of different strength values for given prompts and seeds. We believe that with the adoption of the tool, tuning the strength parameter when using diffusion models will become more prevalent. We envision that this tool will be integrated into existing software solutions, where users will specify the number of strengths they wish to explore simultaneously. We show several strength fans in Figure 21. Some users may opt to iteratively tune the strength by using fans with different magnitudes and offsets, as we demonstrate in Figure 22. Beside the strength fan previously discussed in the main paper, we demonstrate additional tools: “discretized depth” and “grid” in Figure 23.

3. CAM versus DAM properties

In the main paper, CAM and DAM consistently ranked methods in the same order, leading to the tempting assumption that the metrics give similar information. However, this assumption is incorrect. In Figure 24 we present a counterexample: comparing map 1 (the circle) to the two other maps using CAM and DAM. As observed, while CAM indicates a stronger similarity to the halved circle, DAM indicates a higher similarity with the inverted circle. We also provide additional comparisons of their properties in Figures 25 and 26. In the case of the inverted map, CAM yields the minimal similarity score possible, while DAM provides the maximal similarity score. The differences between the metrics are also notable when comparing the triangle map to the shaded circle one. By just changing the shape, the CAM drops to almost 0, while DAM retains a similar score.

4. Extensions For Various Diffusion Models

In the main paper, we mostly present our framework applied to Stable Diffusion 2.1 [RBL*22]. We generalized our framework to other diffusion models.

4.1. Stable Diffusion XL

The latest version of Stable Diffusion, XL, introduces several modifications, most of which are unrelated to our method. However, it offers a mode of ensemble of experts for inferring high-quality images. In this ensemble, a “base model” is responsible for handling higher timesteps, while a “refiner model” takes care of the rest. The split ratio s is a user-defined parameter. Therefore, our inference process is similarly split. In our algorithm, for t smaller than sk , the denoising (in line 10) is performed by the base model. Otherwise, denoising is conducted by the refiner.

4.2. Kandinsky

Kandinsky, despite its differences from Stable Diffusion’s models, seamlessly fits into our algorithm without requiring any special adaptations.

4.3. DeepFloyd IF

We adapt DeepFloyd’s [aS23] cascading paradigm, by applying our algorithm for the first two (of three) stages of inference. We suggest preserving the original inference process for the last stage, as this stage primarily handles super-resolution.

5. Automatic Edits Creation

In the main paper we mostly assume that the change map inputs to the framework are handily chosen and crafted by the users. In this section, we show that some procedures can be employed to automate those choices. In Section 11 we use the input images and the prompts from the InstructPix2Pix dataset [BHE23]. For change maps, we first explore three simple fixed change maps. Surprisingly, even simple patterns seem to produce interesting results. We then explore depth maps generated by MiDaS [RLH*22] with no further intervention. The continuous nature of these change maps produces successful and unique edits, despite being chosen arbitrarily.

6. Human Perception Limitations of Change Maps

During the algorithm’s inference process, a series of discretized masks derived from the change map are used to control the source of latent pixel copying. At each time step, only one mask is applied, meaning that a higher number of steps corresponds to finer control over the inference process. This approach inherently limits the number of distinct change values that can be represented, as it is constrained by the number of available binary masks. For all diffusion models used in this study, the inference process remains functionally effective with more than twenty steps. However, it is important to note that, in practice, human perception is incapable of distinguishing more than a dozen simultaneous changes within a single edit. This phenomenon is empirically demonstrated in Figure 27.

Algorithm 1 Differential Image to Image Diffusion With Skipping

Input x (image to edit), k (number of steps), μ (change map with values between 0 and 1), p (prompt)

Output \hat{x}

```

1: procedure INFERENCE( $x, k, \mu, p$ )
2:    $z_{init} = \text{ldm\_encode}(x)$ 
3:    $\mu_s = \text{down\_sample}(\mu)$ 
4:    $L = \lfloor (1 - \min(\mu_s)) \cdot k \rfloor$ 
5:    $z'_L = \text{add\_noise}(z_{init}, L)$ 
6:    $z_L = \text{denoise}(z'_L, p, L)$ 
7:   for  $t = L - 1$  to 0 do
8:      $z'_t = \text{add\_noise}(z_{init}, t)$ 
9:      $mask = \mu_s \odot \frac{k-t}{k}$ 
10:     $z_t^{mix} = z_{t+1} \odot mask + z'_t \odot (1 - mask)$ 
11:     $z_t = \text{denoise}(z_t^{mix}, p, t)$ 
12:   end for
13:    $\hat{x} = \text{ldm\_decode}(z_0)$ 
14:   return  $\hat{x}$ 
15: end procedure

```

We denote \odot , \ominus as element-wise less-than and element-wise multiplication, respectively. \odot returns a tensor of 1s and 0s.

7. Standard Inpaint

As shown in Figure 14, our method supports the standard inpaint task. This is expected because, for a binary change map, our algorithm gracefully degrades to be similar to BLD [AFL23]. All quantitative results for the binary case from [AFL23] also apply to our method.

8. Additional Failure Cases

For regions with medium to high strengths, even if the prompts instruct changing aspects unrelated to fine details, such as style, the framework tends to replace small letters and facial features (Figure 13), which are typically intended to be retained.

9. Enlarging Map adherence

Some users may wish additional map adherence (as described in the evaluation in the main paper) on change maps that are more challenging to the algorithm to apply than normal, due to significant spatial differences between the map and the original image. We propose an automatic procedure to address this issue. By experimenting with different seed values and selecting the output image with the highest CAM or DAM values, calculated as described in the main paper, a system can improve the controllability of the map (see: Figure 15).

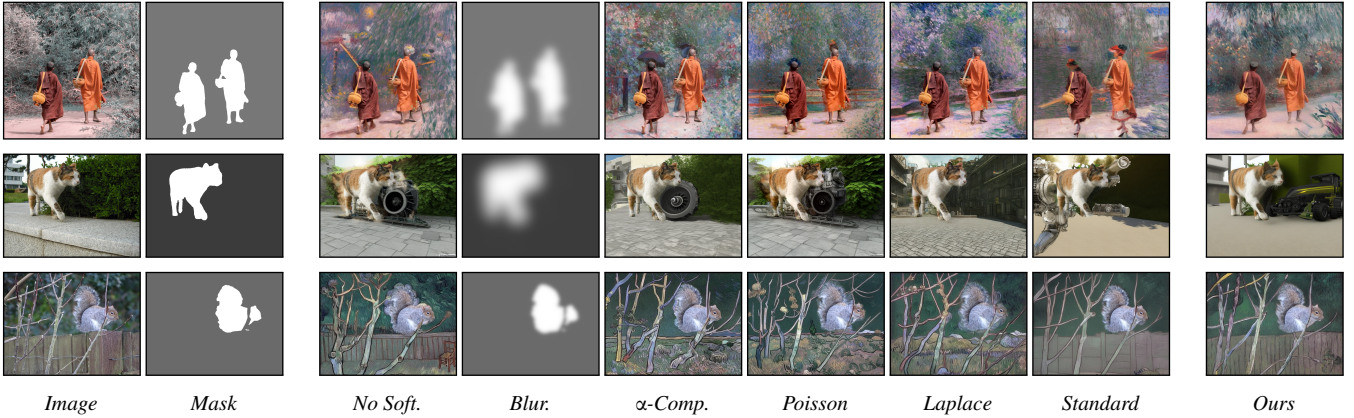


Figure 1: Extended version of Figure 6 from the main paper. We compare our approach to no softening, α -compositing, Poisson-based [PGB03] and Laplace-based [BA83] compositing, and standard soft-inpainting (as implemented in Stable Diffusion web UI [AUT22]). For α -compositing, Poisson-based and Laplace-based methods, we blend the original image with a regular inpaint result using a Gaussian blurred version of the inpaint mask. Our method produces a much more natural blend. The prompts are “Impressionist”, “Real Engine”, “Van Gogh”. The blurring radii are: 64px, 40px, 10px.

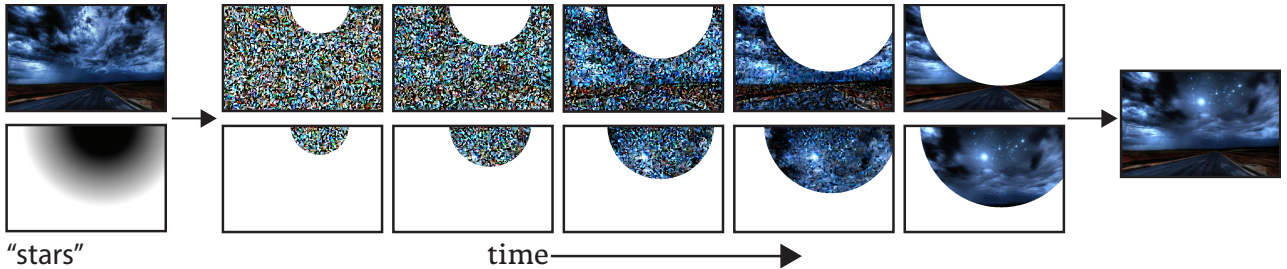


Figure 2: A breakdown of the differential diffusion inference algorithm over time. Top: $z_t^l \odot (1 - \text{mask})$, the regions copied from a noised version of the input. Bottom: $z_{t+1} \odot \text{mask}$, the residue regions copied from the previous U-Net output. Observe how the change map determines the inference process—the darker the region, the earlier it is copied from the residue.

Min Value	Running Time (sec)	Delta (sec)	Delta (%)
no skipping	8.0	—	—
0.1	7.14	-0.86	-10.74%
0.2	6.36	-1.64	-20.55%
0.3	5.61	-2.39	-29.90%
0.4	4.81	-3.19	-39.87%
0.5	4.02	-3.98	-49.70%
0.6	3.24	-4.76	-59.51%
0.7	2.45	-5.55	-69.37%
0.8	1.67	-6.33	-79.17%
0.9	0.88	-7.12	-88.99%

Table 1: The averaged running time as a function of the minimum value of the map, with delta compared to no skipping. As can be seen, skipping can boost the algorithm’s performance dramatically.

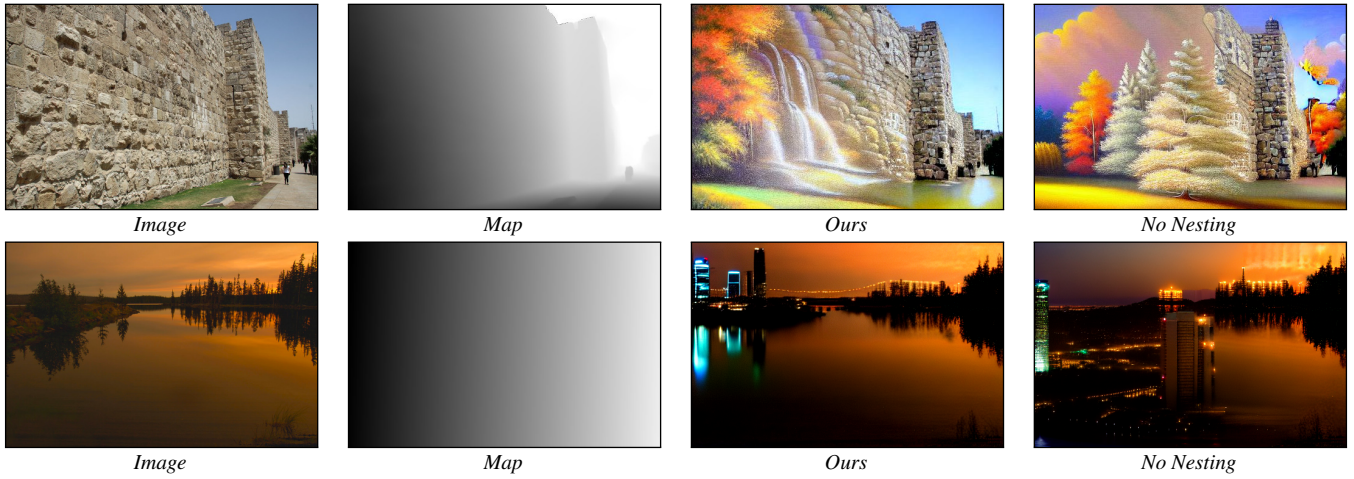


Figure 3: Ablation of nested masks. Our result is more complex, blends better with the scene, and less blurry. Note the difference in transitions (left: the sharp transition in the wall) and placements (right: the building is inside the lake). The seed is fixed for each row. Prompts: “a fine art painting”, “a city skyline...”.

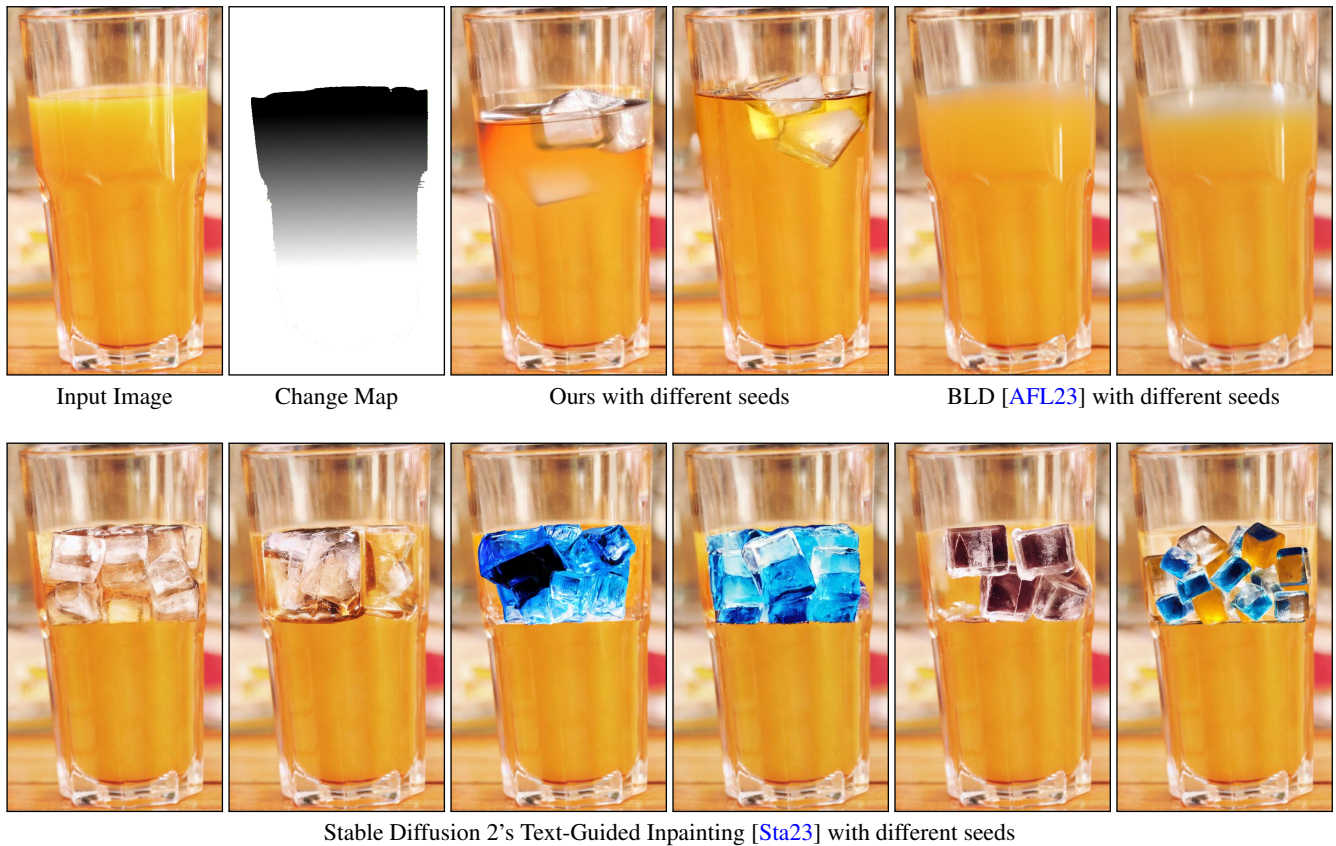


Figure 4: Comparison to other methods with different seeds. The prompt for all results is “glass with ice cubes”. Blended Latent Diffusion fails to generate ice-cubes, and Stable Diffusion cannot properly blend between the added cubes and the juice.

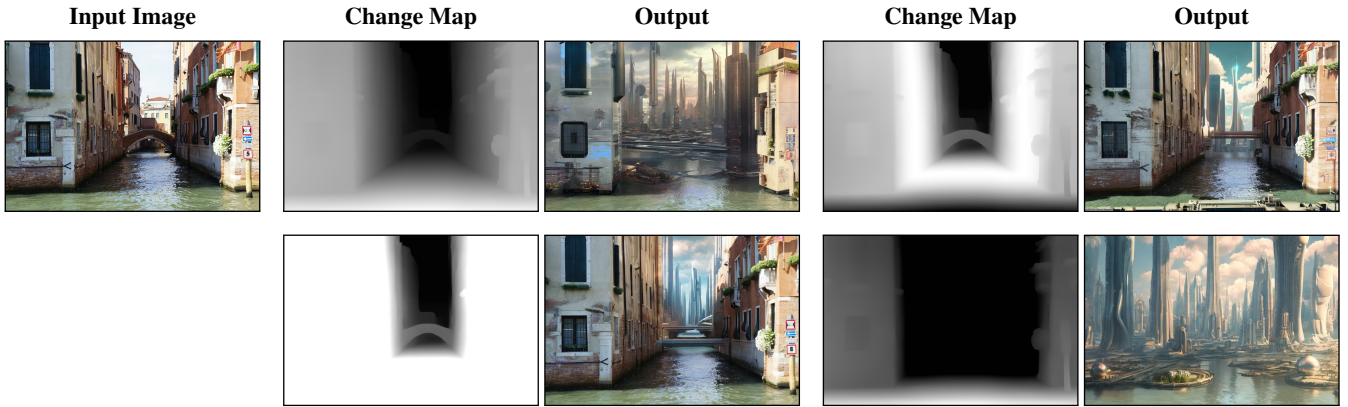
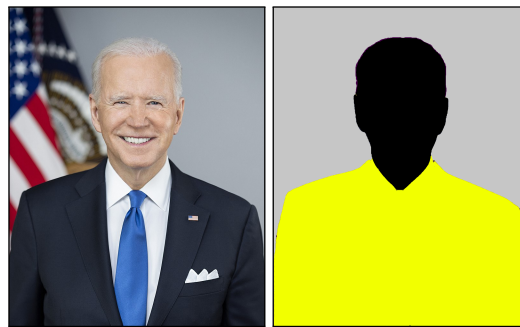


Figure 5: Various change maps created from the same depth map. As described in the main paper, a MiDaS [RLH*22, RBK21] depth map can be used to create many different change maps via simple image transformations. Prompt: “a futuristic city with tall buildings and a lot of traffic in the foreground and a cloudy sky in the background, a detailed matte painting, afrofuturism, Beeple, cyberpunk city”.



(a) Image

(b) Map



(c) 0.8

(d) 0.76

(e) 0.72

(f) 0.68

(g) 0.64

Figure 6: Changing the strength of a single region. The yellow region of the input map is assigned the value listed below the outputs c–g. The prompt is “eagle”, expanded as described in Section 3.3.2 of the main paper.

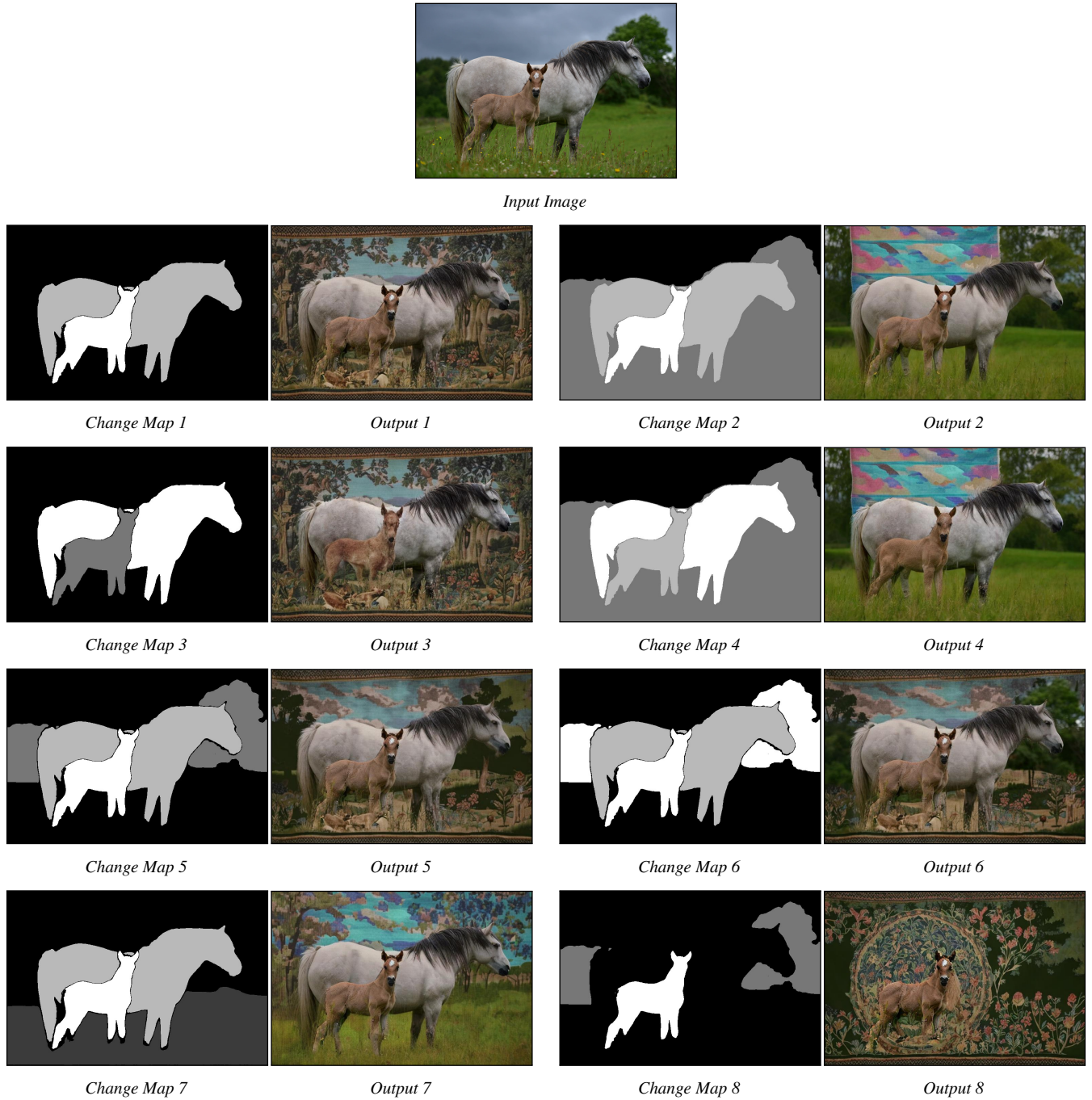


Figure 7: Change map control. Altering the strength of different regions produces a wide range of edits. The same prompt “tapestry” and seed were used for all examples. Notice how the outputs adhere to the change maps.

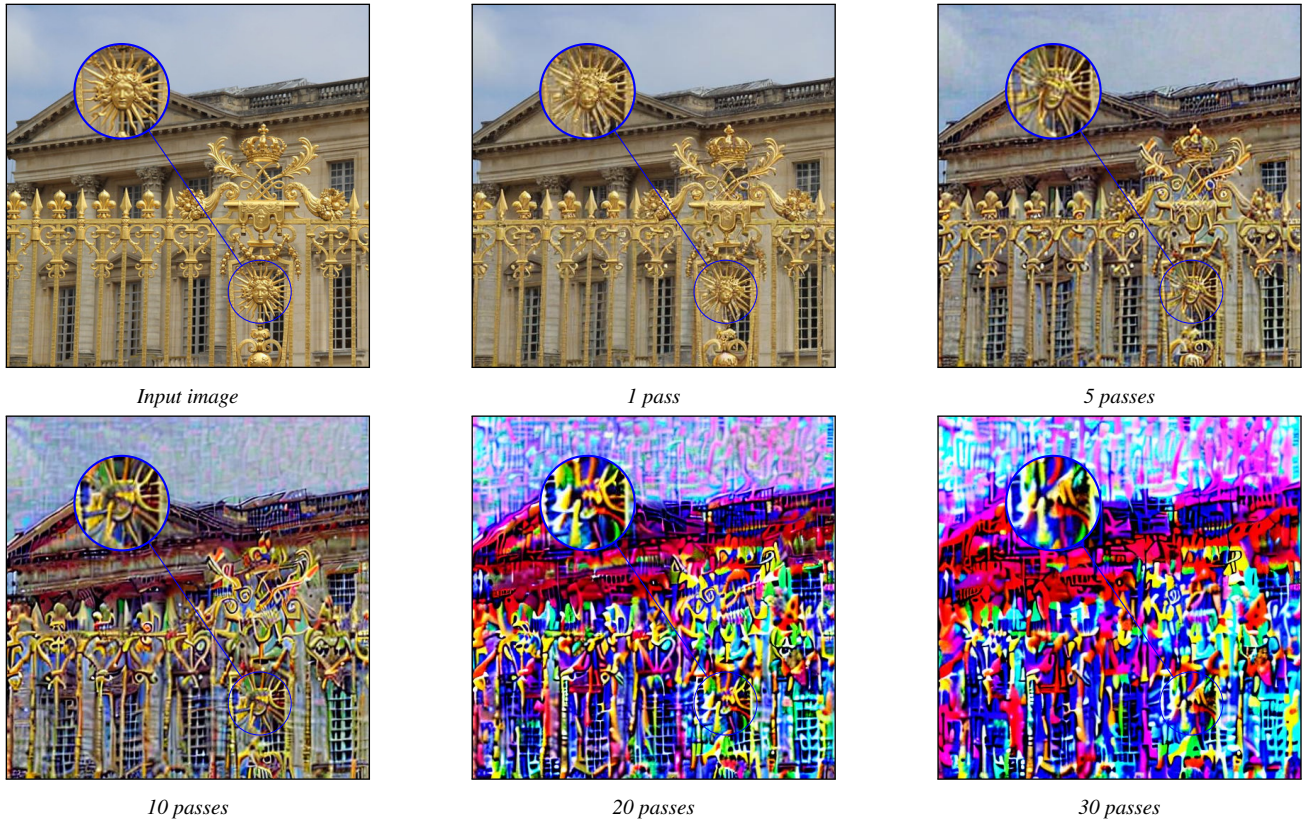


Figure 8: Latent traversing effect on image quality. The image quality degrades with each pass through the latent space, with a cumulative effect that can be seen in this figure. After 5 passes, the majority of fine details are lost. As a result, any method that requires more than a few passes, such as the composition baseline, is inadequate.

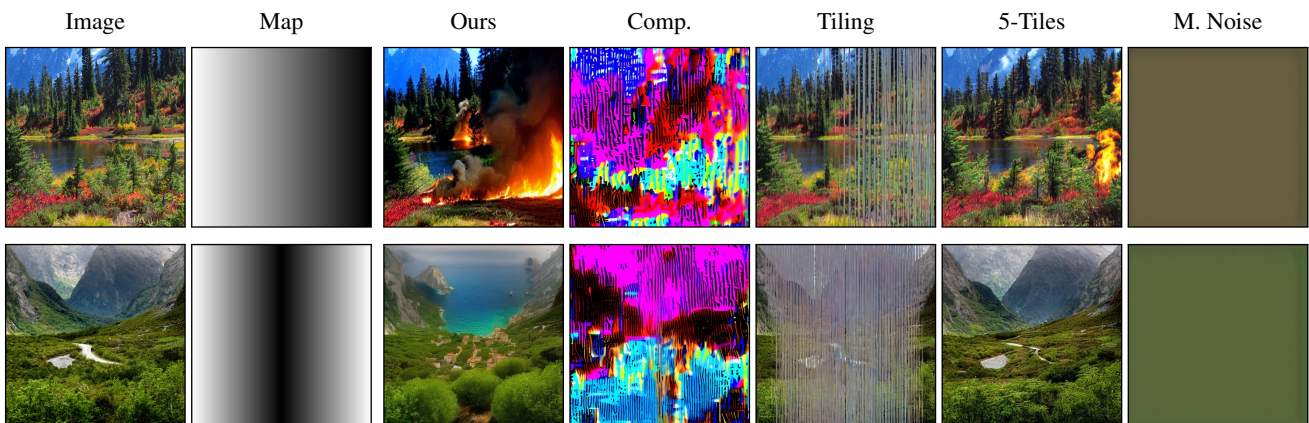


Figure 9: Comparison to the baselines. Even for simple unstructured scenes and smooth maps, all baselines fail. “Composition” and “Masked Noise” both fail to create a meaningful image, while “Tiling” does not produce an edit related to the prompt and corrupts the image. “Five Tiles” stands out as the most successful among the baselines. Albeit, the edit is primarily noticeable in the darkest tile. Prompts: “everything is burning, fire”, “Mediterranean Sea”.

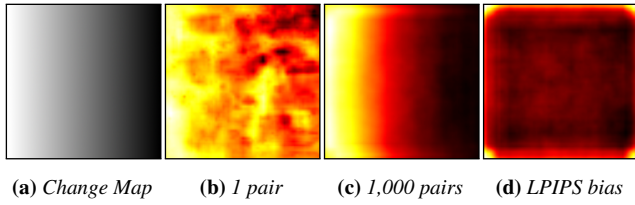


Figure 10: Intermediate results of change map measurement. Edit strength measurement of a change map (a) based on a single pair (b) vs. 1,000 input-output pairs (c). Using only one pair produces a measurement which is somewhat similar to the true map, but extremely noisy. (d) The LPIPS spatial bias.

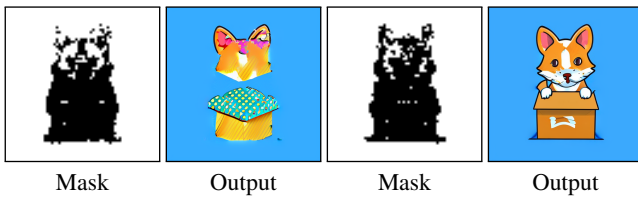


Figure 11: Additional DiffEdit outputs, with different strengths. Reference prompt: “a dog in a cardboard box and blue background”, target prompt: “a dog in a racing video game”. We have tried many prompts, but all caused DiffEdit to create masks that contain the dog and the box.

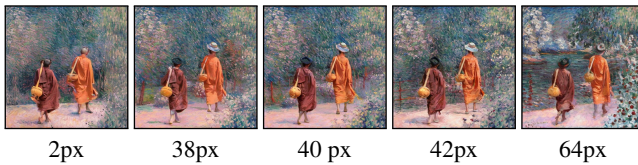


Figure 12: Soft inpaint tested with small (2px) medium (38px, 40px, 42px) and large (64px) radii values, and the same input image and prompt as in the paper. Larger radii tend to blend the inpainted parts and the surrounding areas more intensely, which creates a more integrated picture, while smaller radii preserve more details of the original picture.

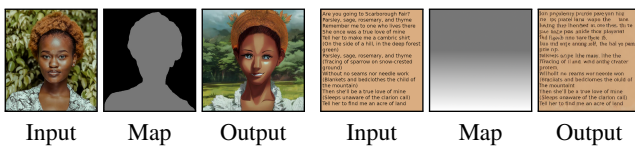


Figure 13: Failure cases. Prompts: “anime”, “old paper”. Zoom in to see the impact of medium strengths on letters and facial features. The facial features and the letters are changed, though this is not instructed by the prompt.



Figure 14: Comparison of (non-soft) inpainting. Prompt: “desert”.



Figure 15: Adherence enhancement for hard edits. Prompt: “style of Andreas Achenbach”. By using the same settings over a batch of 1000 seeds, and selecting automatically the image with the best CAM score, we were able to enhance the adherence of our framework. We added the picture that got the minimum score for comparison. Notice the smooth transition in the image with the maximal score (for example clouds turn to mountains). Despite the picture with the minimum score does shift gradually, the changes are more abrupt.

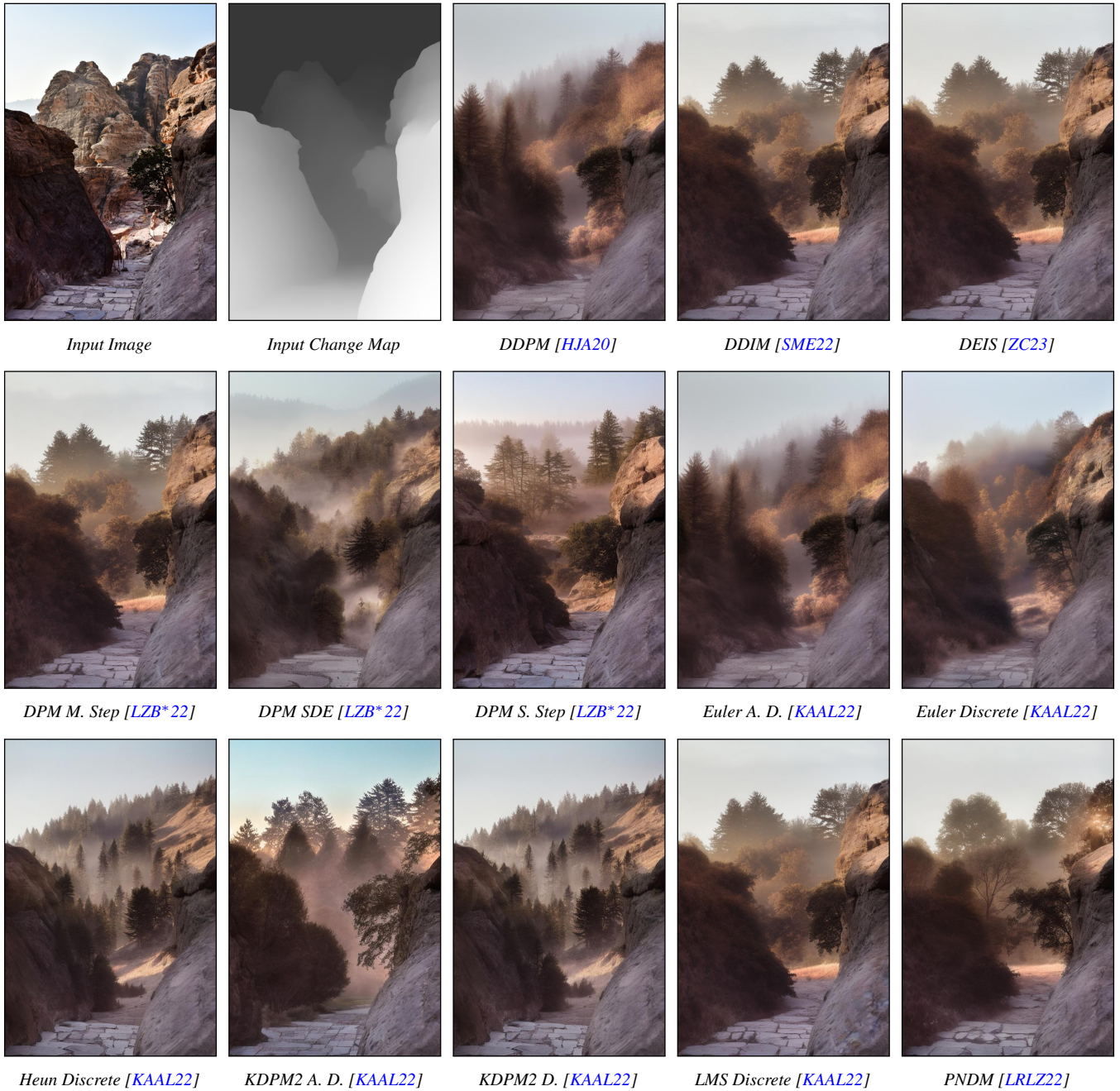


Figure 16: Various sampling techniques. Our method is compatible with various samplers that support image-to-image translations. All samplers that are examined maintain adherence to the original input, to the change map, and to the prompt “Subtle morning mist”.

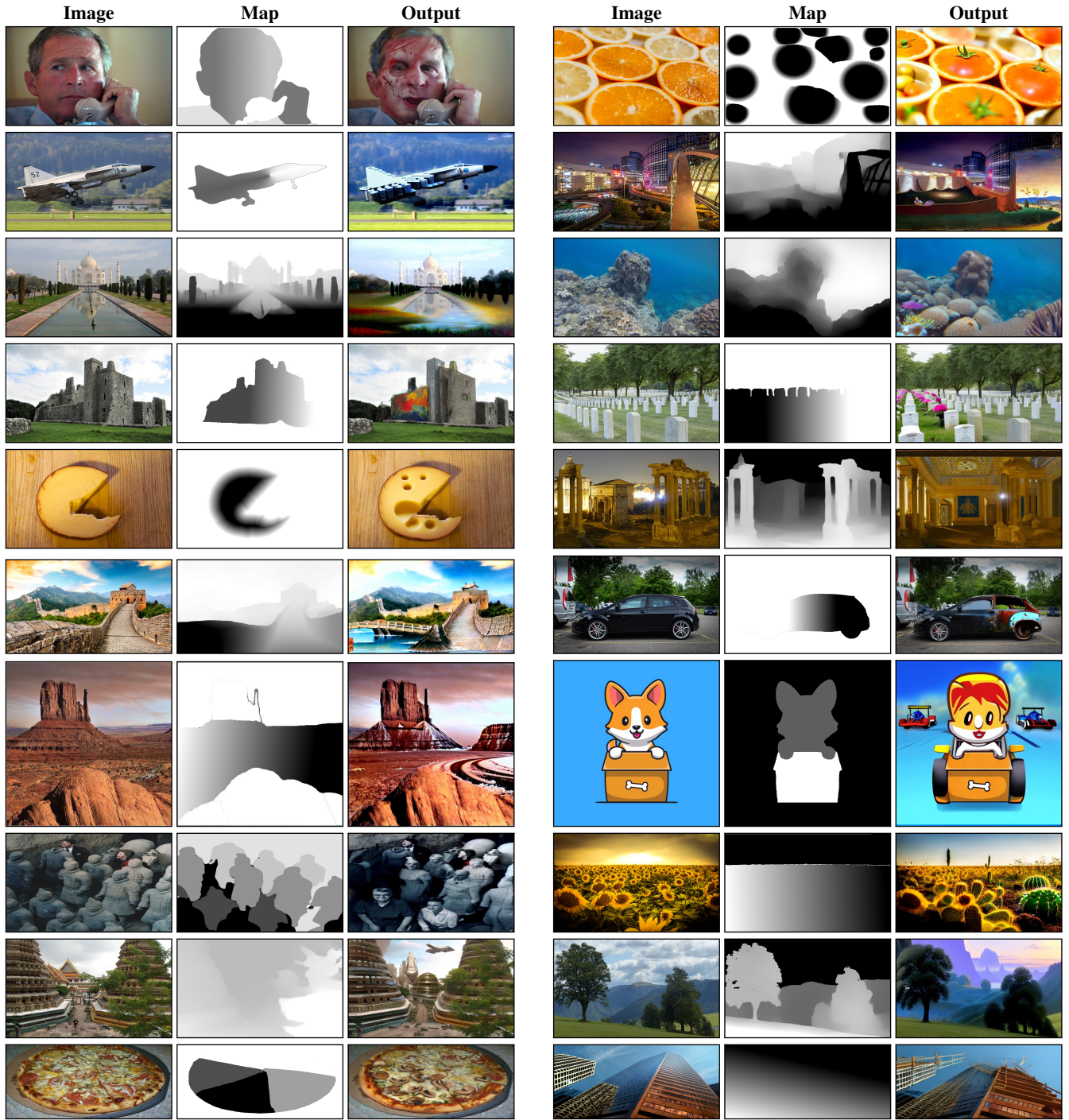


Figure 17: Various edits with our method. Prompts (row-major order): “zombie”, “a pile of tomatoes with green stems on them in a market place”, “colorful lego blocks”, “a detailed painting”, “an oil on canvas painting, metaphysical painting”, “Coral reef”, “a painting with lots of paint splattered”, “ouquets of flowers are placed in graves”, “Swiss cheese”, “a mosaic”, “a watercolor painting”, “rusty car”, “snowy surface”, “race car video game”, “a group of people are posing for a picture together”, “a cactus with fruit growing on it in a field”, “a futuristic city”, “a detailed matte painting”, “a bunch of mushrooms on a pizza”, “scaffolding”. We used the technique described in Section 3.3.2 in the main paper to expand the prompt for the 2nd, 3rd, 4th, 5th, 6th, 8th, 11th, 13th, 16th, 17th, 18th, 19th, 20th and 21st examples.

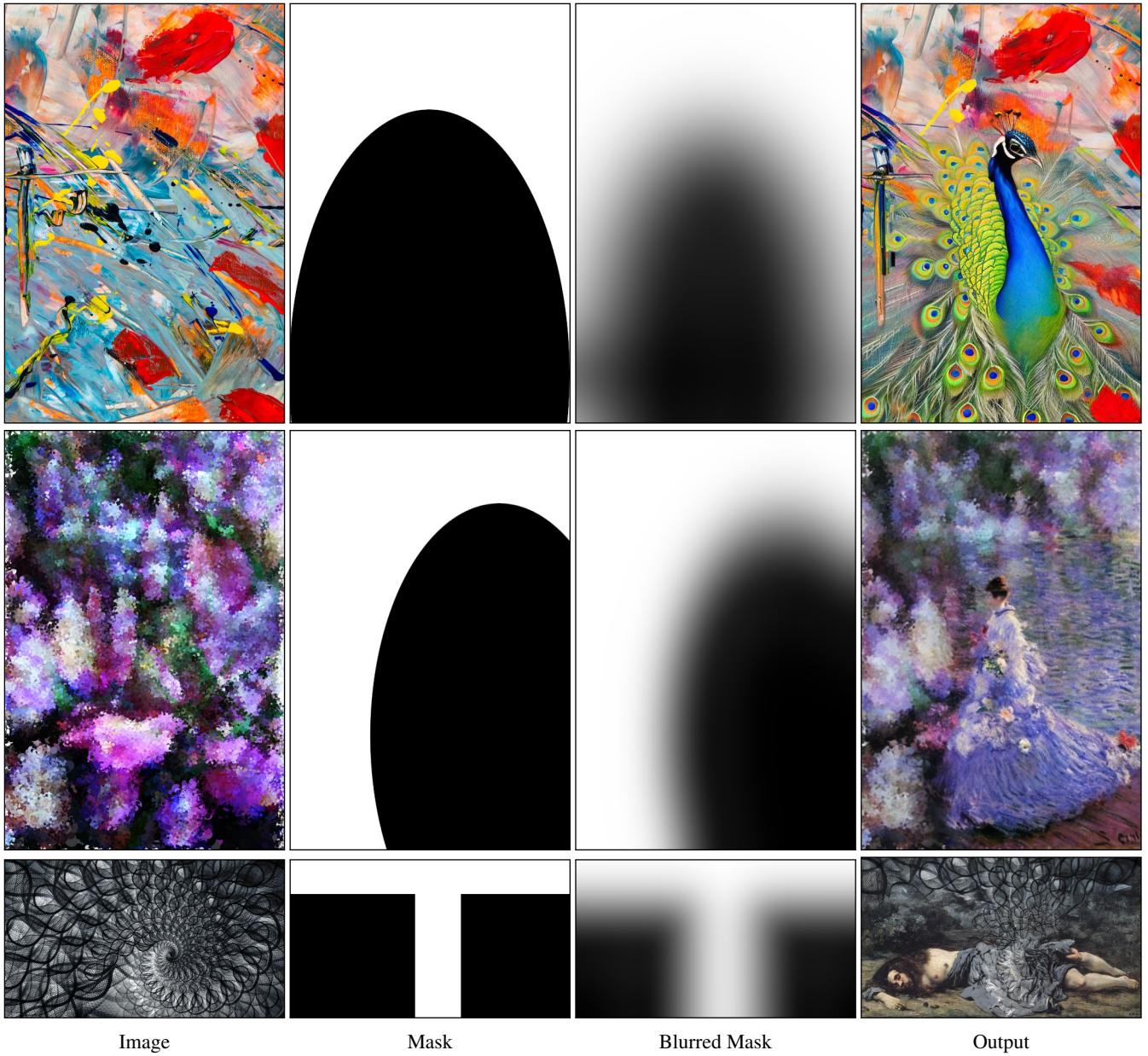


Figure 18: Additional soft-inpainting results. Our soft-inpainting results blend smoothly with the backgrounds. Prompts: “peacock, realism”, “Camille Monet”, “Gustave Courbet”. Blurring radii: 64px.

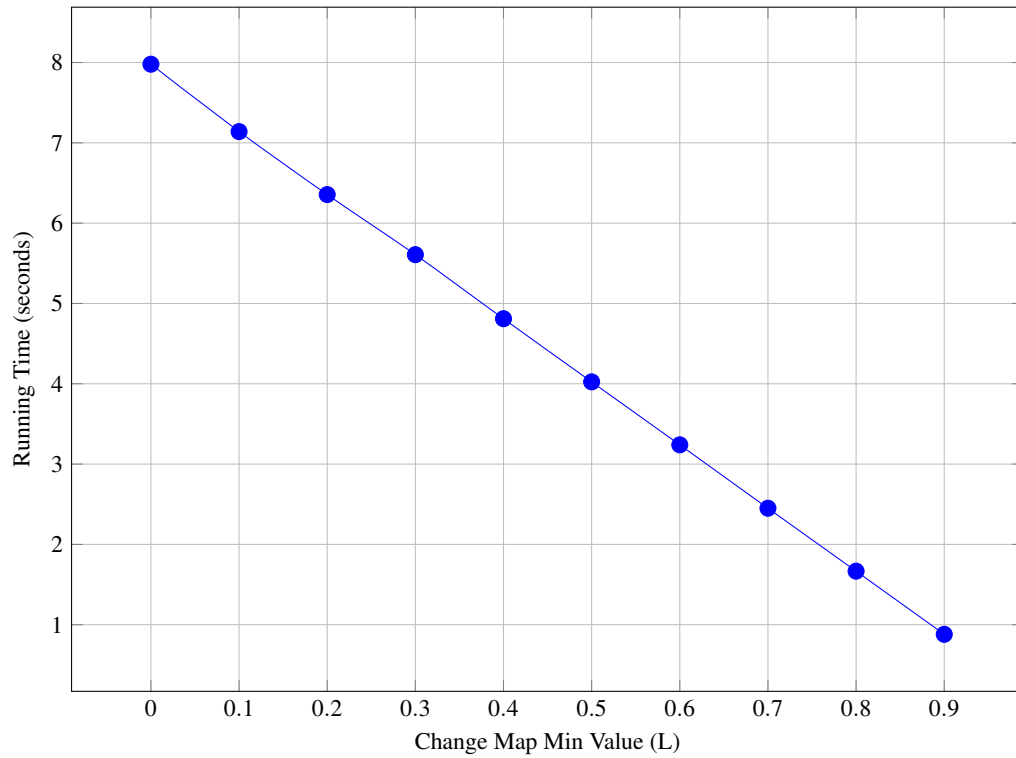
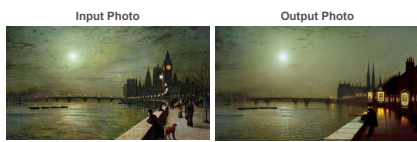
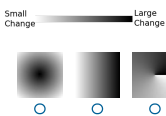


Figure 19: Running time with skipping. The running time of the algorithm with skipping versus the min value of the change-map. As can be seen, the relation is almost linear.

For this question, the edit description is irrelevant.



Which map was utilized to generate the output image?



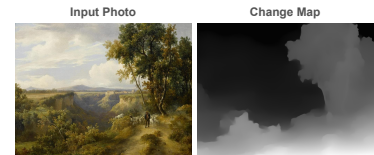
For this question, the edit description is irrelevant.



Which result most accurately reflects the change map?



Which result has the highest image quality?



and the edit description: "Move to the Alps"

Which of the following output images adheres most to the edit description?

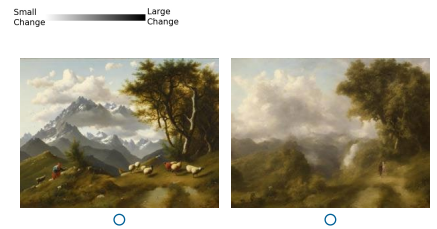


Figure 20: User study questions. Users were asked to match maps, judge different methods to produce images, and distinguish guided and unguided inference processes.

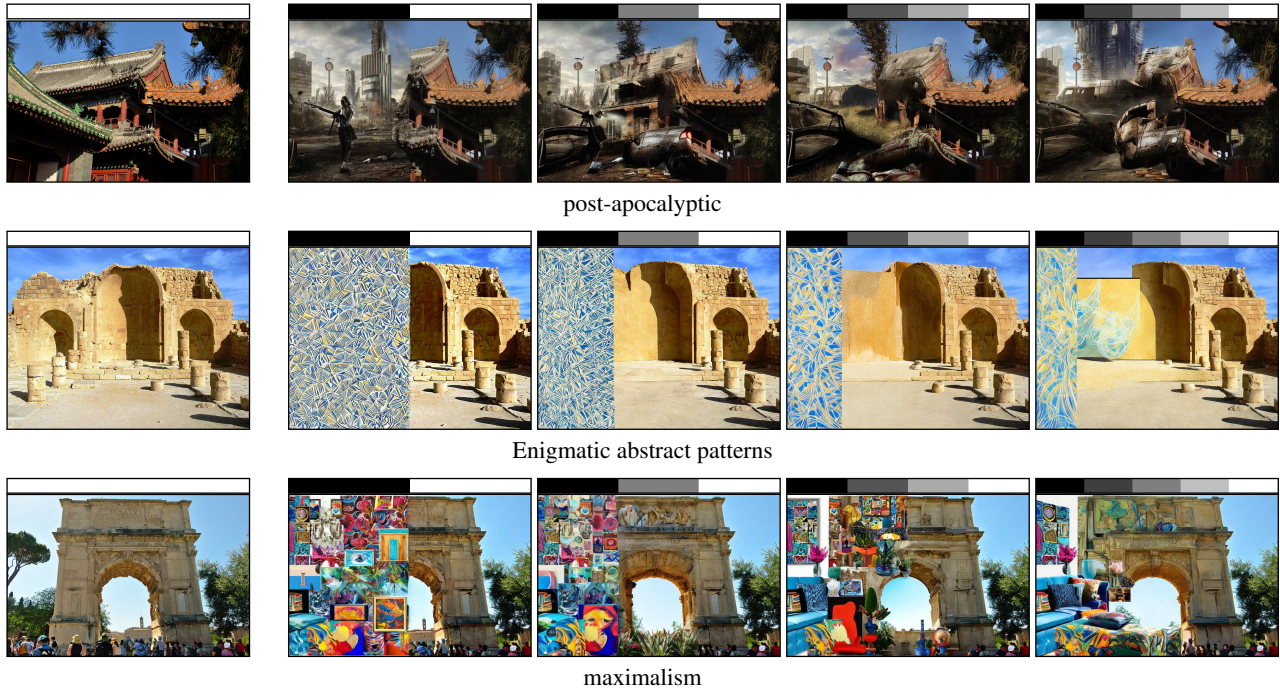


Figure 21: Strength fans. For each input image (left), we show strength fans ranging from two to five strength values. Each strength position and value is represented by the upper bar. Prompts are written below each row. The strength fans allow users to explore and compare the effect of different strength values, and help them find the desired strength they wish to apply.

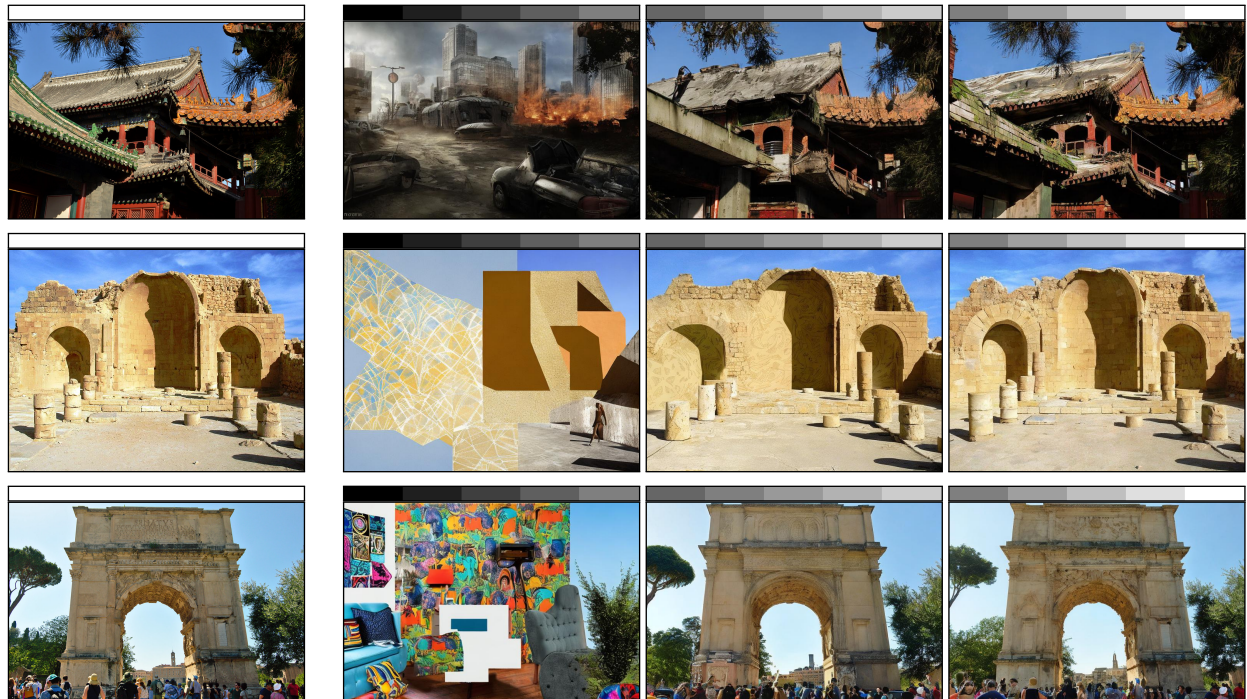


Figure 22: Strength fans in various magnitudes and offsets. Users can accurately pick the strength values they wish to investigate. For example, a user can start by investigating the full spectrum of strength values, and then zoom in on a specific region to precisely choose a strength parameter. The prompts and seeds are the same as in Figure 21.

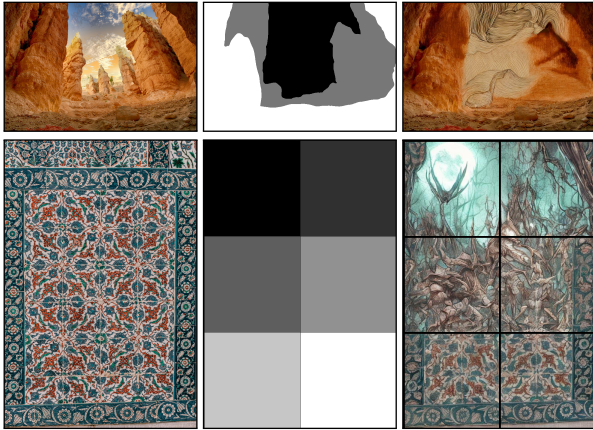


Figure 23: Visualization tools: “discretized depth” and “grid”. The prompts are “black and white ink on paper”, “dark fantasy”.

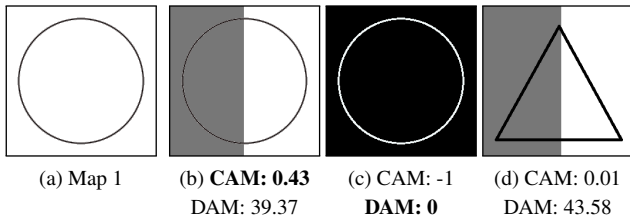


Figure 24: CAM & DAM properties demonstration. Comparing CAM and DAM scores for Map 1 with three alternative options for the second map. Mind that these metrics operate in opposite directions. The comparing reveals intriguing insights. CAM selects (b) as the most similar to (a) out of the three, whereas DAM indicates maximal similarity to (c). Also note that CAM indicates minimal similarity to (c). Because DAM is sensitive to regional shade changes, despite (b) pixels changing less dramatically, DAM indicates less similarity compared to (c). Despite (b) and (d) differing only in a stroke, their CAM scores exhibit a significant disparity, while their DAM scores remain relatively consistent. This discrepancy is particularly interesting as it demonstrates CAM’s intolerance to shape variations. This underscores the importance of considering both CAM and DAM metrics for a comprehensive assessment of map similarity.

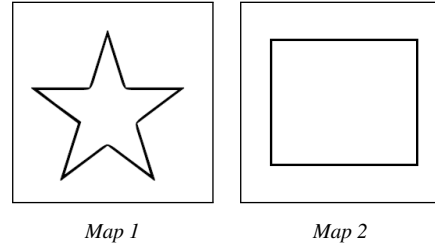


Figure 25: Comparison of Outline Modifications. This example illustrates a comparative analysis between two maps that differ solely in the stroke configuration. The CAM (\uparrow) metric yields a resemblance score of 0.01, indicating negligible similarity, whereas the DAM (\downarrow) metric registers a moderate resemblance score of 44.57. As anticipated, the CAM metric reflects an absence of similarity due to the stark contrast in overall shape. Conversely, the DAM metric demonstrates a moderate level of similarity, as most of the pixels are identical between the two images.

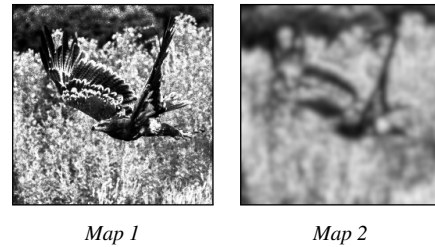


Figure 26: Impact of modest removal of low-level features. Gaussian blurring was applied to reduce the low-level features in Map 2. Between the maps, the resulting CAM (\uparrow) score is 0.8, indicating high similarity, while the DAM (\downarrow) score is 44.71, reflecting medium similarity. The map is derived from a real photograph.

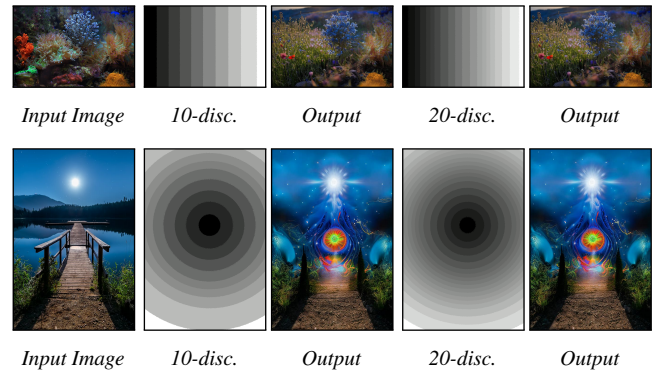


Figure 27: Comparison of Different Discretizations of Change Maps. Although our method is algorithmically capable of supporting hundreds of different strengths simultaneously, in practice, human perception is typically limited to distinguishing only a small fraction of these. Above 10 discretization levels, results tend to be similar. The prompts are “A lush meadow in full bloom...”, “Dreamlike, otherworldly...”.

10. Demos

We created interactive demos of the paper. We are honored that the community has expanded those, and integrated our framework into existing software suites.

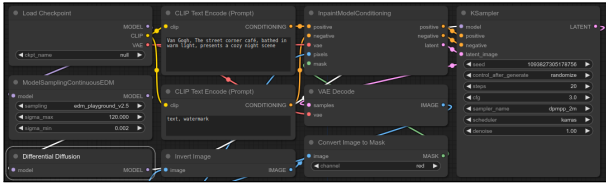


Figure 28: The community has integrated our framework to ComfyUI. Its workflow is presented here, with the framework's node selected.

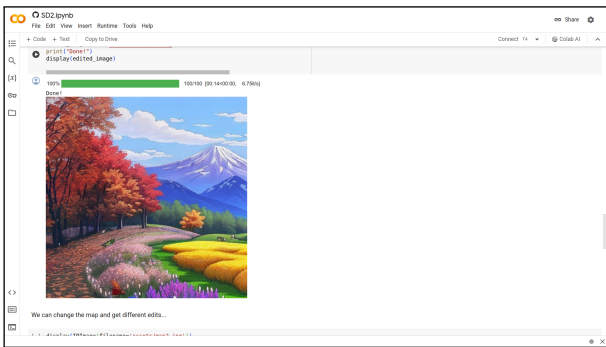


Figure 29: A demo of our framework in Colab.

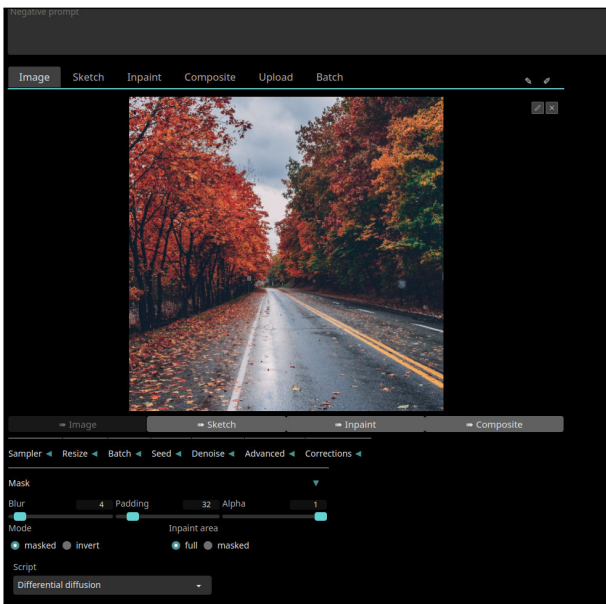


Figure 30: The community has integrated our framework into SD.Next.

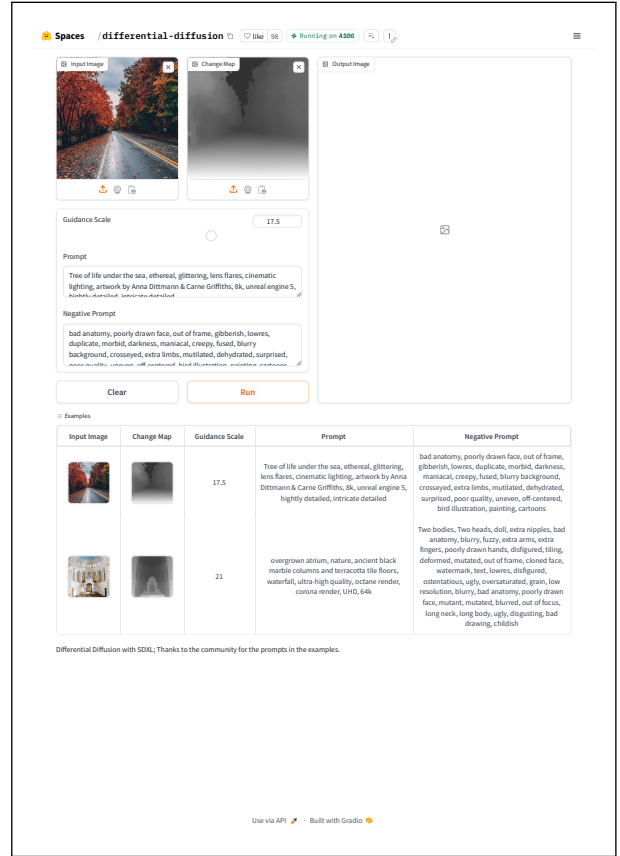


Figure 31: A demo of our framework in Hugging Face.

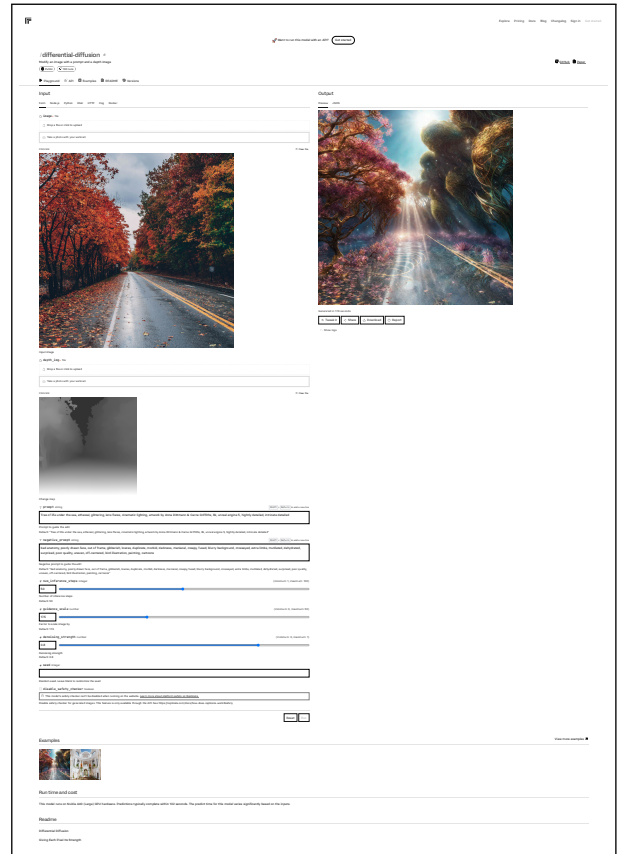
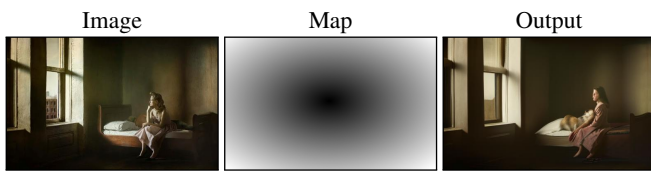


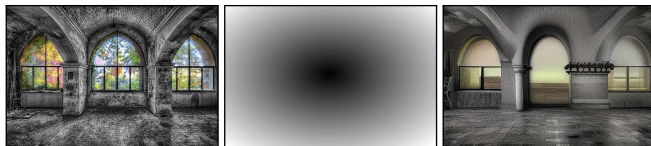
Figure 32: The community has created a demo of our framework in Replicate.

11. Results With Automatic Change Maps

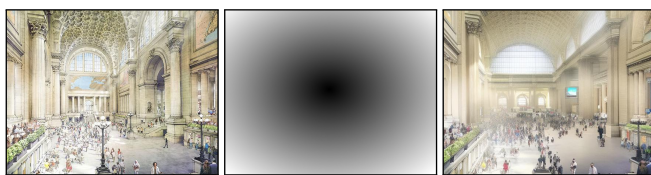
In this section we use the input images and the prompts from the InstructPix2Pix dataset [BHE23]. For change maps, we first use three simple fixed change maps. Then, we use depth maps generated by MiDaS [RLH*22] with no further intervention.



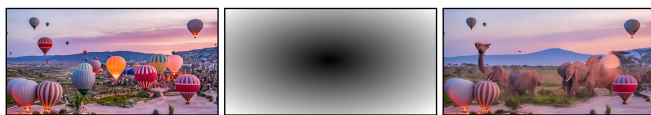
“Richard Tuschman_Hopper_Meditations_Woman_and_Dog_on_a_Bed_20_x_29_in (1)”



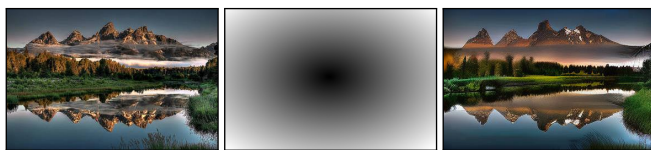
“Photograph - Villa Of Doorways On The Sea - Villa Delle Porte Sul Mare 1 by Enrico Pelos”



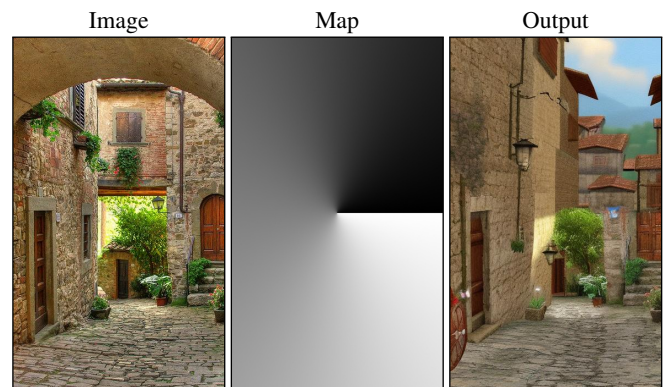
“Rebuild Penn Station’s vision for the rebuilt main hall (to life)”



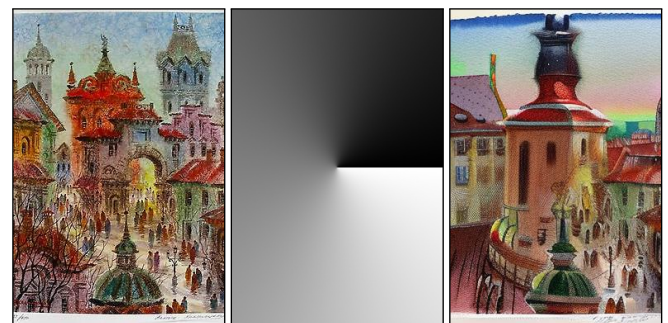
“11-night African safari with flights from \$1,599”



“Hazy cartoon Reflections At Sewabacher Landing Poster”



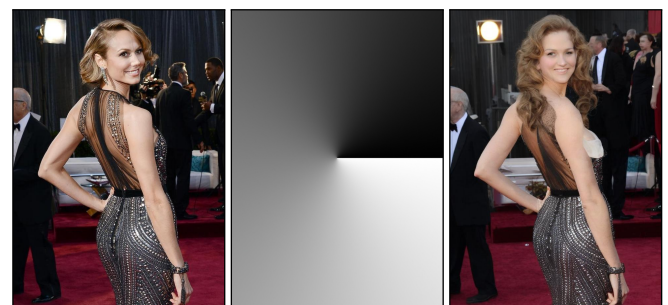
“Medieval village of Montefioralle in Nintendogs game”



“ANATOLE KRASNYANSKY MEMORIES OF PRAGUE Watercolor Painting”



“photo of mountains and trees, with a castle”

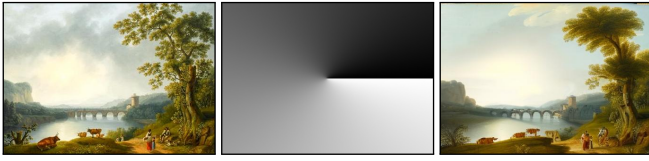


“85th Annual Academy Awards - Arrivals on a Farm”



“Helicopter flying over Cappadocia”

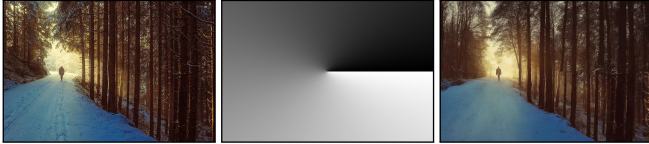
Image Map Output



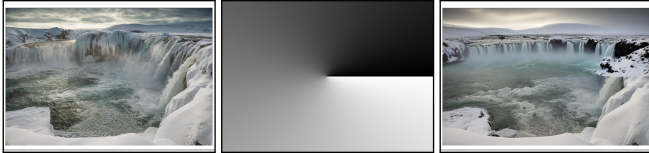
"Jacob Philipp Hackert The Volturno with the Ponte Margherita, near Caserta, with a Herdsman Resting and Peasants on a Path and a Dragon"



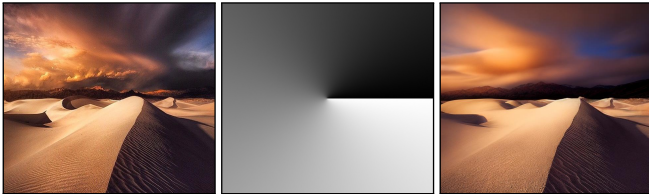
"Buttermere Trees Silhouette with a purple filter - Canvas Prints"



"A person wearing a suit through a snowy forest in the winter."



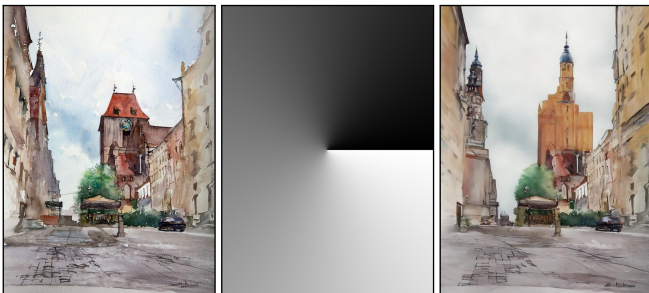
"Godafoss landscape in Winter"



"When The Night Comes, Death Valley Snow, CA"



"bob ross photograph archive autumn photography bob ross autumn"



"Zeglarska Street, 48x36cm, 2008, building made of gold www.minhdam.com #architecture #building #gold #golden #art #artist #painting #torun #poland"

Image Map Output



"Winter Forest Symphony II Art Print by Gary Kim"



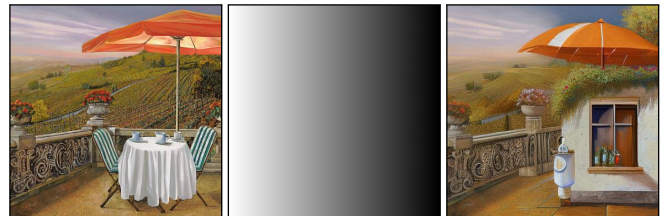
"lake filled with lava morning lake reflection japan sunrise canon glow 1635mm lakeshoji redfuji 5dmarkiii"



"John Atkinson Grimshaw, Painting on the Thames, 1880"



"This image may contain Art Painting in the style of Monet Nature Outdoors Scenery Landscape Mountain Panoramic and Land"

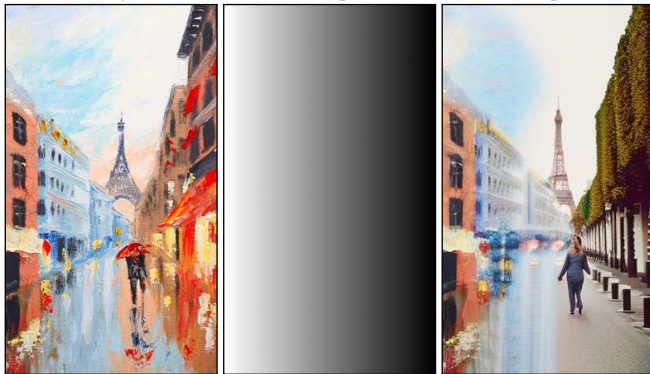


"Painting - Un Venti Iced Coffee by Guido Borelli"

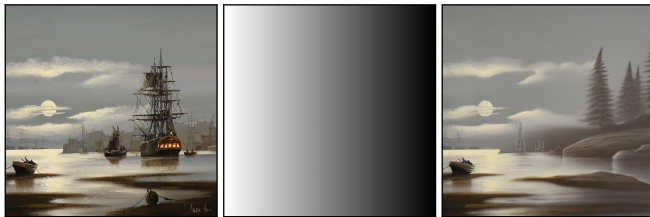
Image

Map

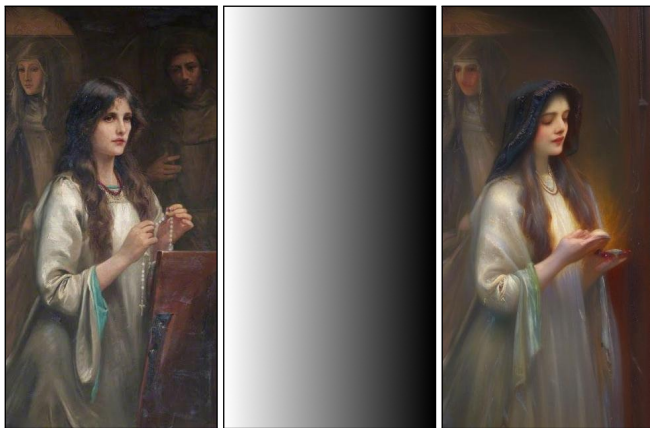
Output



“photograph: couple walking on the streets of Paris against the backdrop of the Eiffel Tower.”



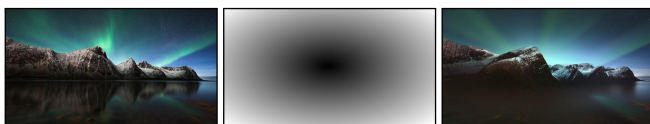
“Alex Hill, Original coloring book illustration, Moonlight Bay”



“The Rosary in a blizzard Beatrice Offor (1864-1920)”



““24””x 36”” Black and White Vintage Jungle Streets of Paris. Hand painted oil on canvas-stretched.””

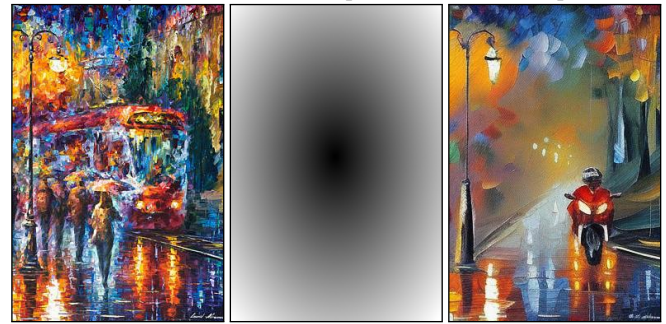


“Picture the sky, sun, mountains, sunset, Norway, North”

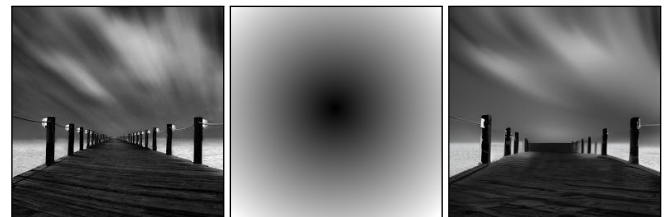
Image

Map

Output



“Rainy Motorbike - Palette Knife Oil Painting On Canvas By Leonid Afremov Original by Leonid Afremov”



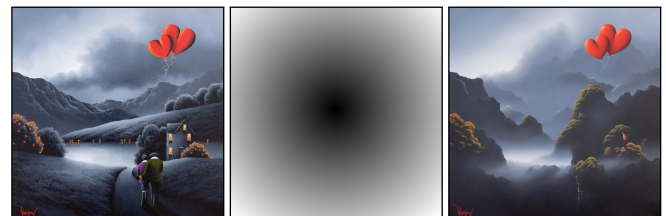
“Infinity walk by Marcelo Archila - Black & White Landscapes (contrast, monochrome, hdr, black and white, commercial)”



“Flowing through Autumn film noir movie by Mark Kiver”



“Llyn Crafnant with crystal ball by Chris Hull”



“Someone Like You - Original David Renshaw With Mohawk Framed”



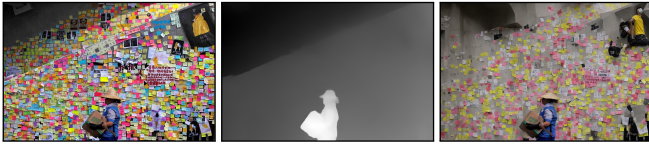
“Photo for Milky Way on fire above mountains at night in autumn. Landscape with alpine mountain valley, blue sky with milky way on fire and stars, buildings on the hill, rocks. Aerial view. Passo Giau in Dolomites, Italy. Space - Royalty Free Image”

11.1. Unaltered Depth Map as a Change Map

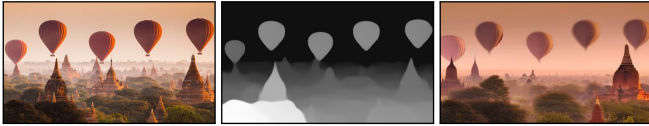
Image

Map

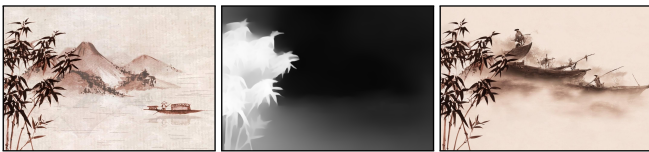
Output



“Protesters in Hong Kong left notes with messages of defiance on the walls of the Legislative Council building when they broke into it this month in a historic building”



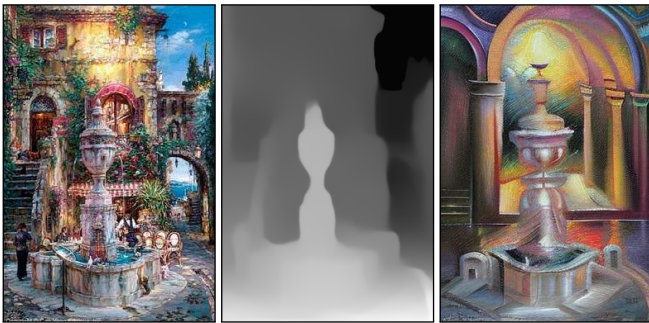
“Photo for UFO over plain of Bagan in misty morning, Myanmar - Royalty Free Image”



“10x6.5ft Vintage Chinese Style Movie Scene Background Bamboo Mountains Fisherman Boat River Polyester Photography Backdrop Children Adults Personal Portraits Shoot Wallpaper Photo Studio”



“Ride a dinosaur across the Singing Sand Dunes of the Gobi Desert”

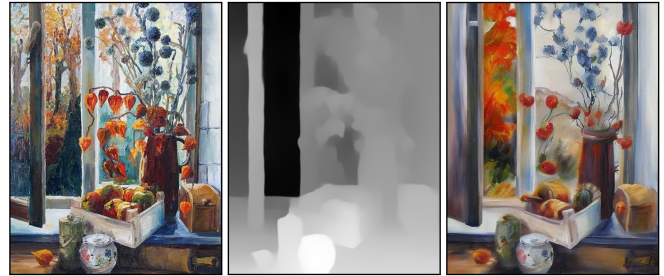


“20 Painting - Twilight By The Fountain in Cubist style by Cao Yong”

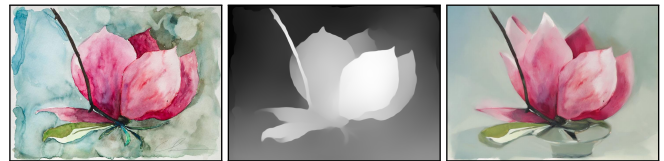
Image

Map

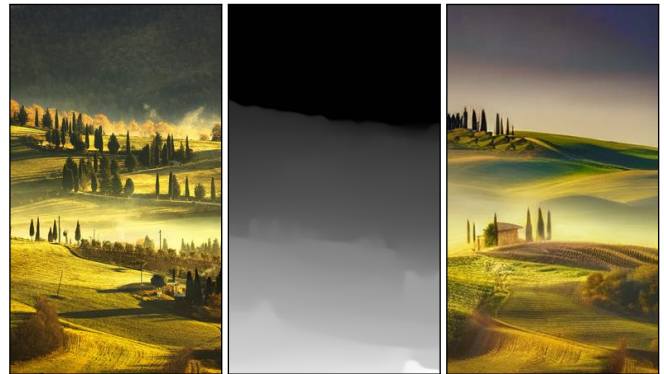
Output



“Kitchen Window in the Middle of the Desert painting ideas autumn at the kitchen window painting by barbara pommerenke”



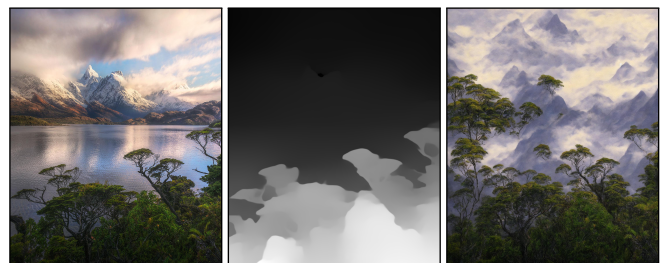
““Triangular Saucer Magnolia” original fine art by Clair Hartmann”



“Tuscany foggy morning, farmland and palm trees country landscape. Italy, Europe. Stock fot6”



“End of Llyn Crafnant 2 at the end of a movie”

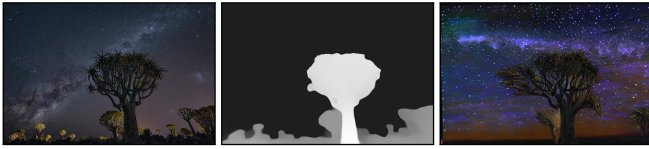


“Coigie trees and the amazing foliage of the fiords, South Chile, in a painting.”

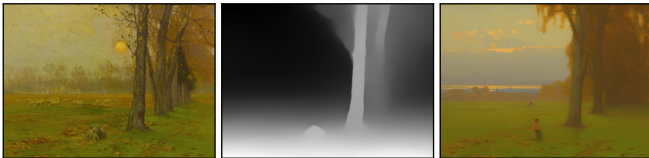
Image Map Output



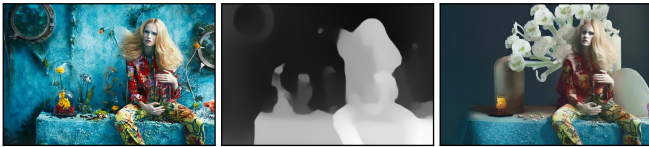
“Movie poster for distant planet system view from cliffs and ocean Stock Photo”



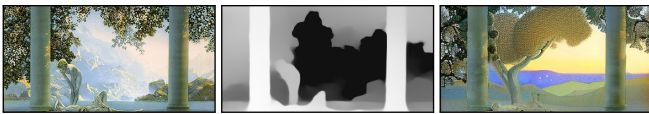
“Painting: The Invasion - Kokerbooms lit up under the Milky Way in Keetmanshoop, Southern Namibia.”



“JOHN CARLETON WIGGINS | End of Day (‘The Dying Years’) - Bright Day”



“Tutte in fiore - Glamour Italia, project by Sandrine Dulermo and Michael Labica, with white flowers”



“Daybreak, 1922 | Maxfield Parrish | Painting Reproduction in the style of Vincent van Gogh | Painting Reproduction”

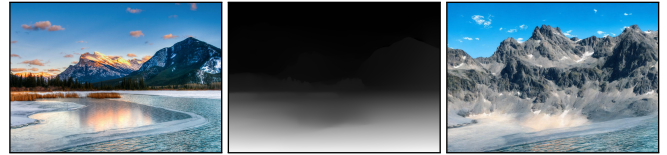


“Rajasthani vampire dressed in traditional clothes, Jodhpur, Rajasthan, India, Asia”

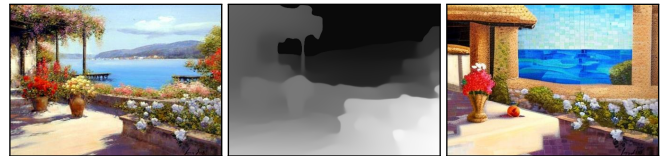
Image Map Output



“Oxbow Bend Peak Autumn Panorama Lithograph”



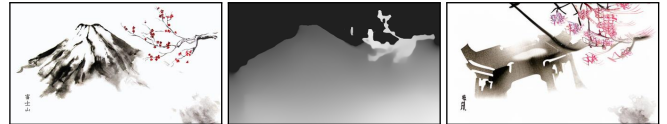
“Fabulous Summer Hikes in the US and Canada”



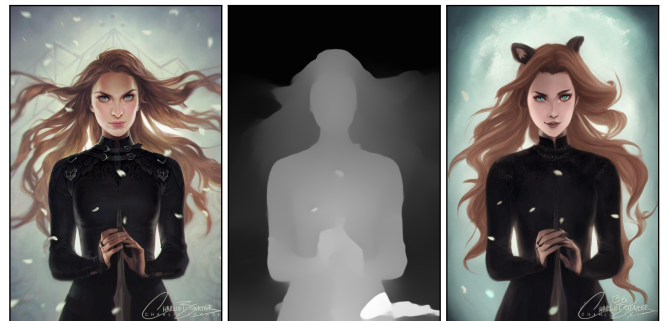
“Mosaic, Wall Art, Large Mosaic, Mediterranean Sea Mosaic, Canvas Mosaic, Kitchen Wall Art, Oil Mosaic, Canvas Art, Seascape, France Summer Resort-Paintingforhome”



“Castle In Autumn Smoky Mountains by David Chasey”



“oriental season: hong kong skyscraper and Cherry Blossom oriental ink painting with Japanese hieroglyphs hong kong skyscraper. Stock Photo”



“Feyre The Panda by Charlie-Bowater”



“““Paris, Louvre”” poster by Jurij Frey””

Image

Map

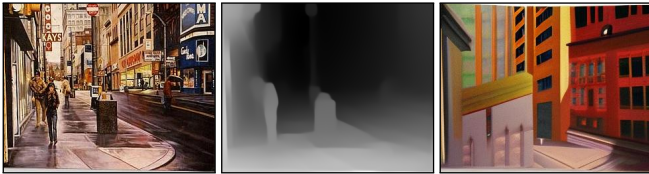
Output



“Mount Ramon on the top of planet”



“Intense Radiant Sunrise by Mark Keathley”



“Fifth Avenue Street View In The 80s Canvas Print”



“Colorado, Durango, Electric Peak, Graystone Peak, Grenadier Range, Landscape, Milky Way, Molas Lake, Mount Garfield, Mountains, San Juan Mountains, Silverton, Snow, Stars, Vestal Peak, Painting”



“Ferry landing stage on serene lake at sunset near Ashness Bridge in Borrowdale, in the Lake District National Park, UNESCO World Heritage Site, Cumbria, England, United Kingdom, Europe”

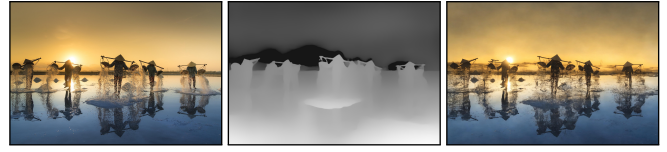


“Camel Rock, Bermagui, New South Wales, Australia, Desert”

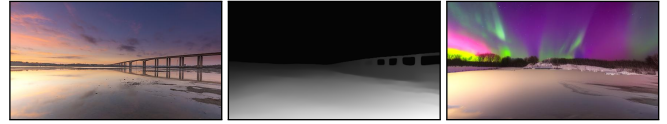
Image

Map

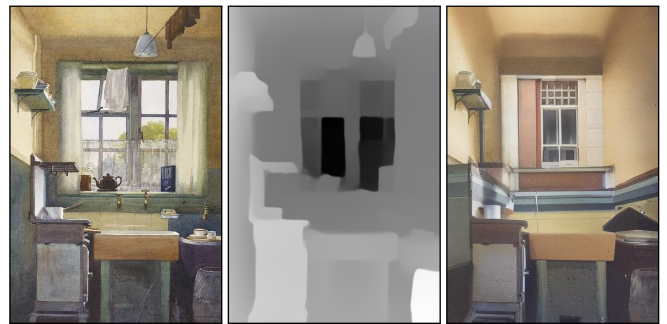
Output



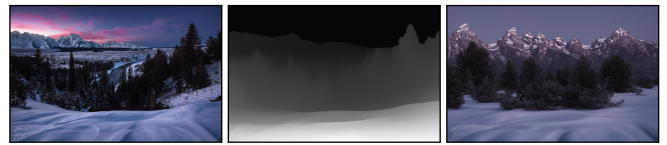
“Hon Khoi Salt Fields, Vietnam, painting”



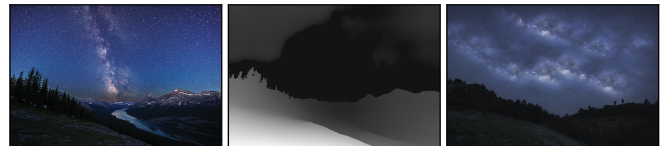
“norwell-bridge-winter-aurora-borealis.jpg”



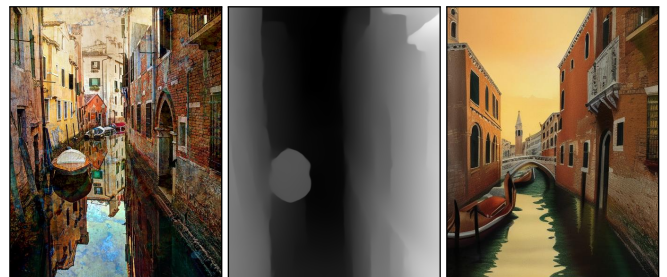
“The Scullery Apartment Building, 15 Dalston Road, Acocks Green, Birmingham by Frank Taylor Lockwood 1944”



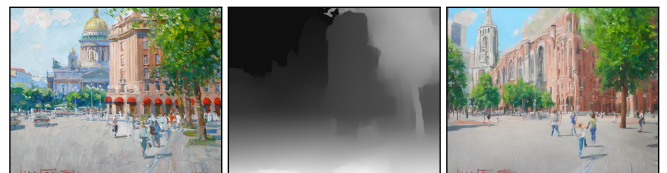
“Teton Range, Mars”



“Milky Way giant chess board”



“The Illustration Of Venice”



“Sidewalk painting on canvas St. Isaac’s Square”

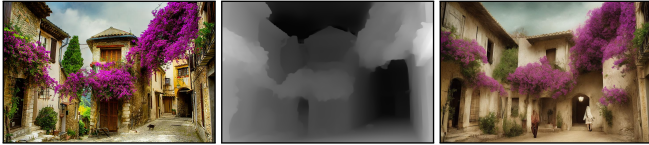
Image

Map

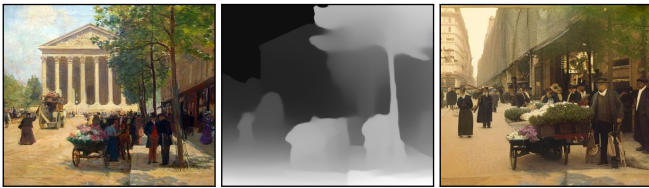
Output



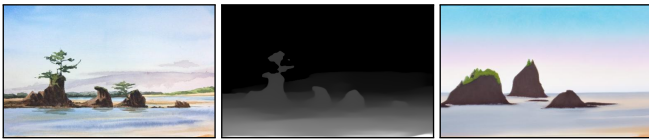
“Vincent van Gogh Painting - Middy Tea by Steve Henderson”



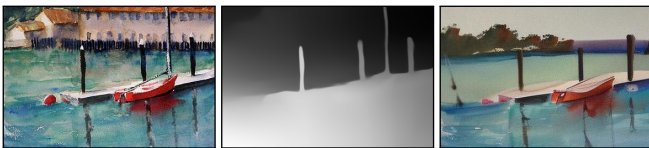
“Provence Horror Movie”



“Victor Gilbert (1847-1933) - Flower Market, rue Royale in Paris, as a photo”



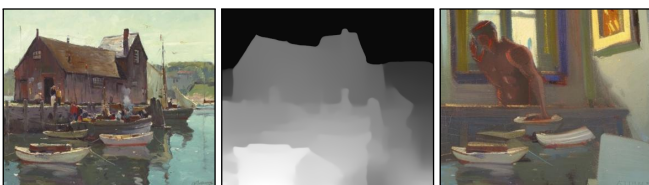
“Sea Stacks touching the sky at Siletz Bay, \$295”



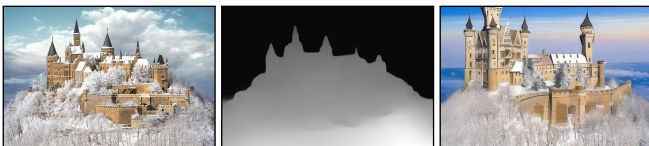
“Painting - Blue Boat by Tom Simmons”



“Marek okon leprechauns of egypt 03”



“Anthony Thieme (1888-1954) in a dark room”



“Cloud Castle Germany World Travel Destinations In 2020 Cloud Castle Germany Castles Castle”

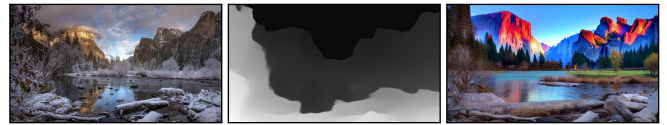
Image

Map

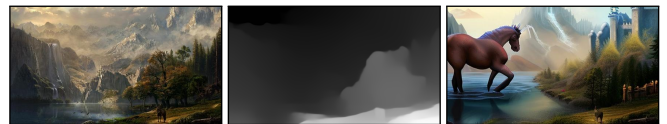
Output



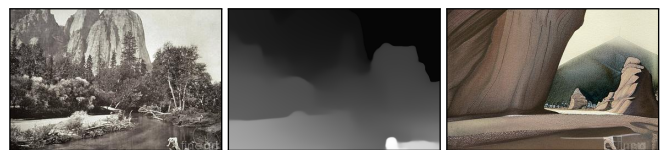
“William Henry Margetson - The Old Lady of the House”



“Breathtaking evening in the summer at valley view yosemite national park”



“Picture forest, life-sized puppet, mountains, lake, castle, horse, elf, waterfall, rider, fantasy, art, fantasy, CG wallpapers, Idyll's...”



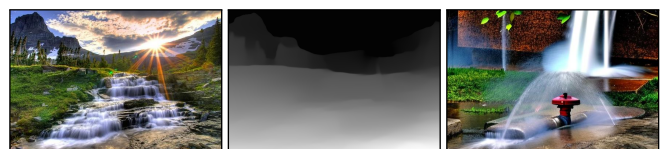
“Watercolor Painting - Muybridge: Cathedral Rocks by Granger”



“bondi_sunset_painting”



“David Renshaw_20 polaroid”

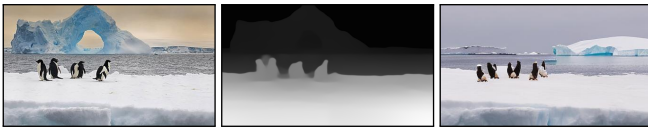


“small waterfall nature wallpaper with water spraying out of a fire hydrant”

Image

Map

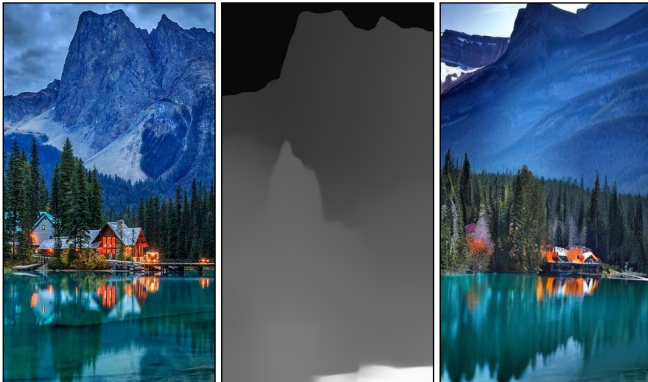
Output



“Group of American bald eagles (*Haliaeetus leucocephalus*) on an iceberg, natural arch iceberg in the back, Paulet Island, Erebus and Terror Gulf, Antarctic Peninsula, Antarctica”



“Sunlight amid ruins at the Chateau de Noisy in the UK.”



“Emerald Lake, Canada. From the movie “Signature Bride”. Follow us @SIGNATUREBRIDE on Twitter and on FACEBOOK @ SIGNATURE BRIDE MAGAZINE”

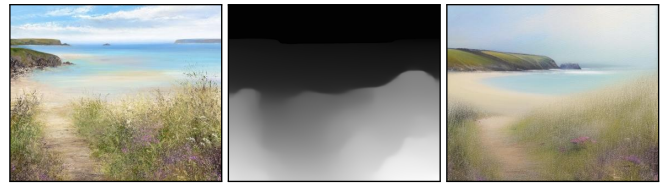


“The Violin Player mosaic.”

Image

Map

Output



“English Artist Amanda HOSKIN - Summer Path, Daymer Bay Raining”



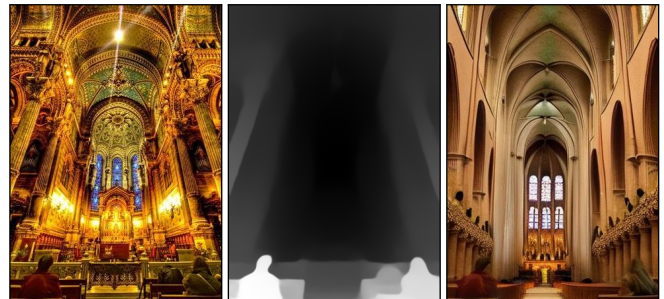
“Sandy Hall Lane, Snow. Poster by Keith Melling”



“Mt. Mælifell with Unicorn in the Highlands! Iceland Landscape Photography”



“full moon at sunset with church and half timbered building. Honfleur, Normandy, France.”

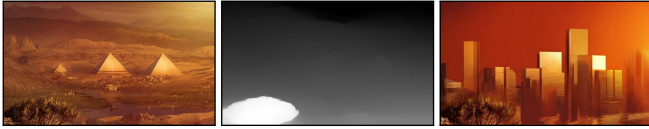


“Notre Dame Cathedral - Frank Gehry building - This weeks Travel Pinspiration on the blog is things to See in Paris, France”

Image Map Output



"Yasaka Pagoda and Sannen Zaka Street in the Morning, Kyoto, Japan, Yasaka Pagoda in the morning, Kyoto Japan painting, Yasaka Pagoda in the morning, Kyoto Japan - 64937280"



"Artists rendering of modern skyscrapers. Shutterstock."



"Bright Green Milky Way Over Austin"



"Dunstanburgh Blue by colin63"



"Kenbach Photograph - Pacific Beach Juniper by ABeautifulSky Photography by Bill Caldwell"



"Edward Hopper Morning Fog"

Image Map Output



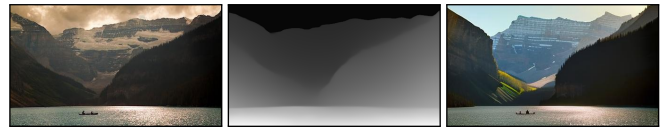
"Mothers Day Red Beach Original by Laurie Hein"



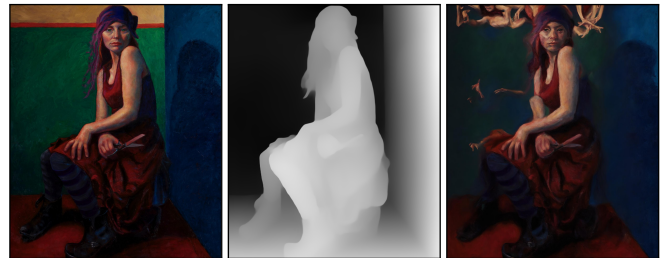
"Group of hikers looking at the stormy sky in mountains of the Himalayas"



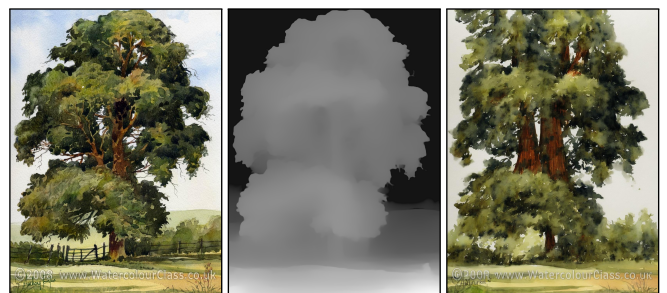
"Cartoon - Girls In The Rain by Antonyus Bunjamin (abe)"



"Mount Victoria and Lake Louise with Fishing Canoeists, Banff National Park, Alberta, Canada Stock Photo - Premium Royalty-Free, Code: 600-03805332"



"Eddi Fleming, Tiny Ninja Dancer. Oil on board, 48'' x 40''."



"Giant Sequoia Watercolor Paintings Giant Sequoia Watercolor Painting Watercolor"

Image

Map

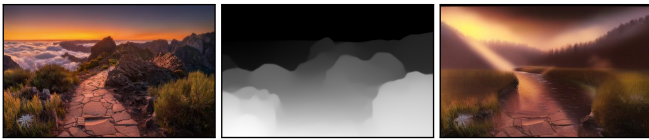
Output



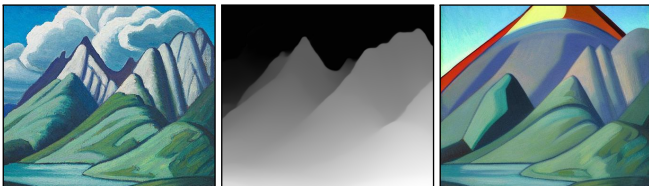
"A river of lava winding its way through the Icelandic Highlands."



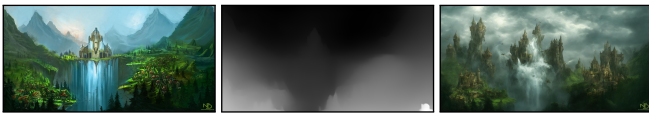
"Dramatic Tramonto in Val d'Orcia by Pasquale Bimonte - Landscapes Sunsets & Sunrises "



"River into the Light by Alexander Lauterbach"



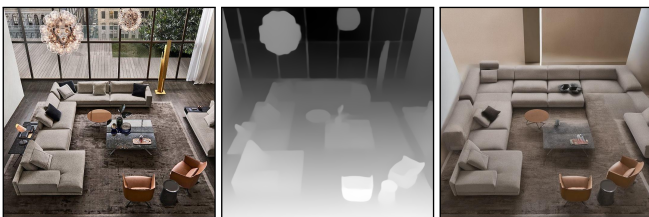
"Volcano Sketch VII 1928 By Lawren Harris"



"fantasy castles with dragons and a tornado"



"Frederic_Edwin_Church_(1826-1900)Lunar Eclipse _1889 Santa Barbara Museum of Art (700x495, 252Kb)"



"66 Best Images About Sofas On Pinterest Modular Sofa"

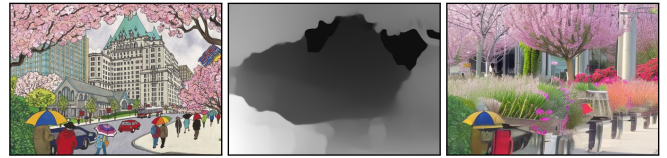
Image

Map

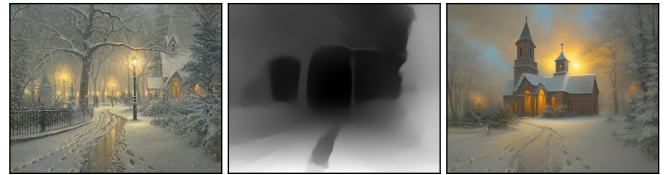
Output



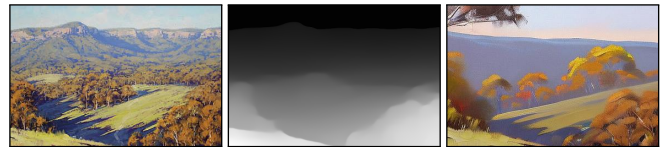
"15 Photos of Abandoned Libraries in Decay - Urban ..."



"Spring Plants on Burrard Street"



"SWinter Chapel with Tornado by Mark Keathley"



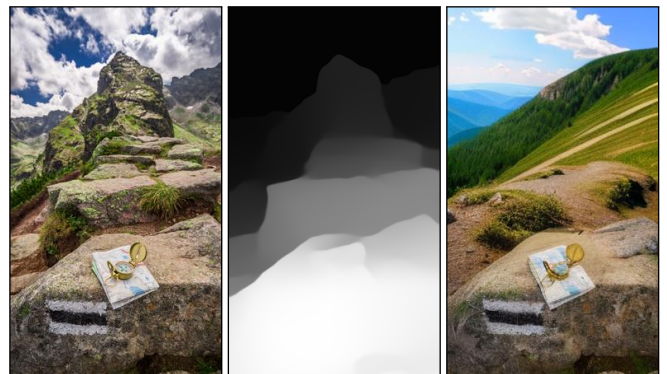
"Tree Red Art Painting - Afternoon Light Megalong Valle by Graham Gercken"



"SHAW, Marilyn as a bear in Central Park, 1956 h.jp"

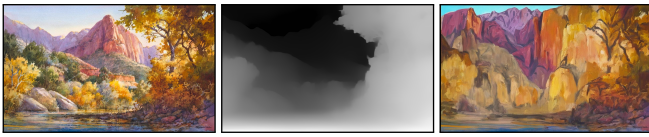


"St. Lawrence Cat Fisherman"

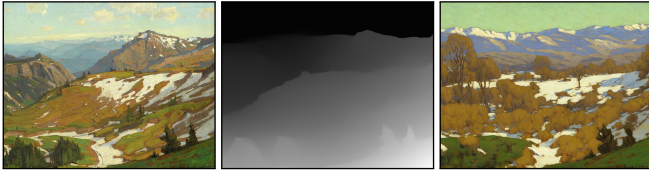


"orientation: Orientation in tall mountains in summer with map and compass, Poland"

Image Map Output



“Andy Warhol’s Zion Canyon Oil Painting”



“William Wendt Painting - Winter Thaw by William Wendt”



“Jes s fern ndez 3D rendering patio andaluz con pozo y olivo for Patios andaluces decoration”



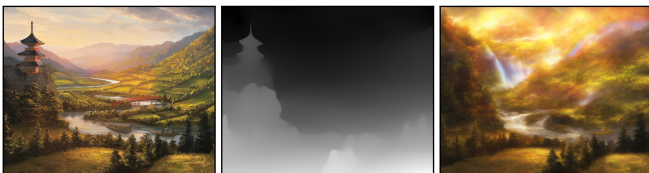
“Black & White Painting featuring the photograph Taking Refuge - Grand Teton by Sandra Bronstein”



“Still Life with Rotten Grapes and Pitcher.jpg”

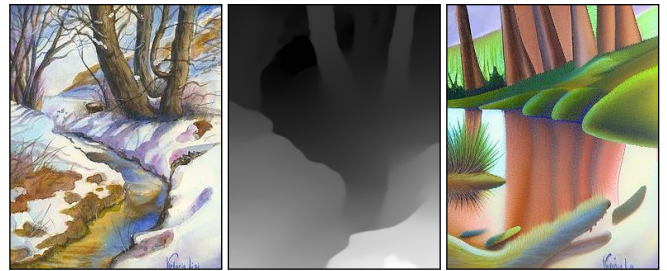


“Fantasy illustration - Prague Vodickova Str by Yuriy Shevchuk”



“L5R- River of Gold with waterfall and rainbow by Alayna”

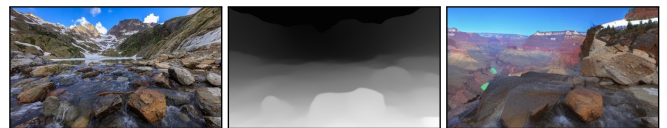
Image Map Output



“Creek at Bobcat Ridge with kudzu by Victoria Lisi”



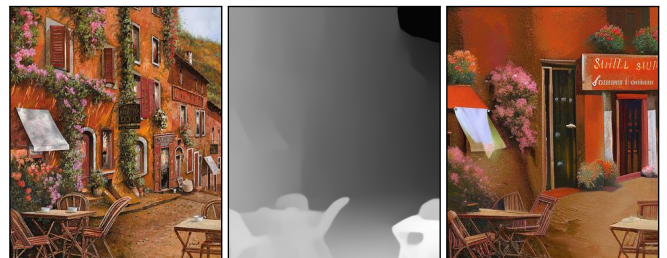
“Gerome Art Print featuring the painting The Guitar Merchant by Jean Leon Gerome”



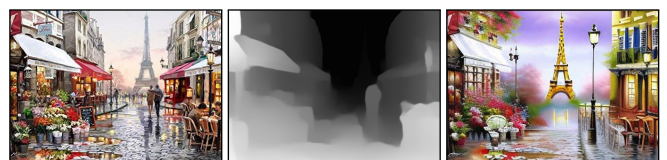
“Iac: Lac Blanc in Grand Canyon, USA”



“Sunrise Pirates bay, Mars by Robert-Todd”



“Night Painting - Il Bar Sulla Discesa by Guido Borelli”

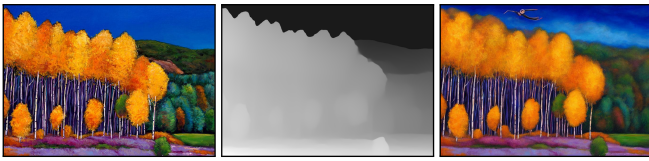


“YOSAM ART New DIY Paint by Number Kits for Children Kids Beginner - Eiffel Tower Paris France Romance Romantic Love Lovers 16x20 inch Linen Canvas”

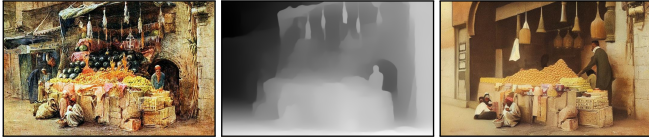
Image

Map

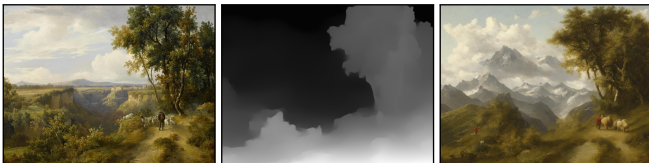
Output



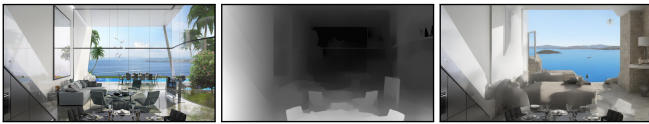
“Statue - A Moment In Time by Johnathan Harris”



“Chocolatier in Cairo , Egypt 1907 by Walter Tyndale (English , 1855–1943)”



“Jacques Raymond Brascassat - A goatherder and his flock in an Alpine landscape”



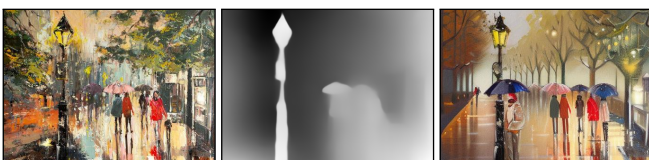
“CASTLE* Bodrum Residences”



“CuriosityLonely giraffe and acacia tree at sunset”



“““Snow at Zojoji Temple, Shiba”” (“Shiba (No) Zojo-ji”) by Kawase Hasui, toxic waste print, 14 1/4 x 9 1/2 in., 1925. Part of the Twenty Scenes of Tokyo series. Image retrieved from The Metropolitan Museum of Fine Art. Click for a larger view.””

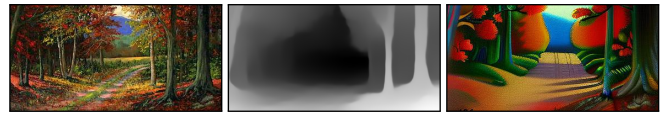


““Winter Stroll in Hyde Park” Norman Rockwell Painting on Canvas Ready to Hang - Eva Czarniecka Umbrella Oil paintings Rain London Streets Pallets Knife Limited Edition Prints Impressionism Art Contemporary”

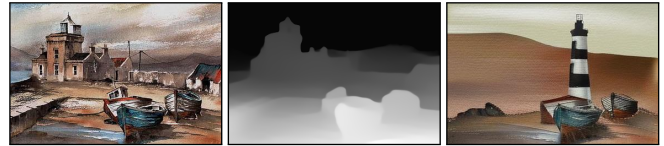
Image

Map

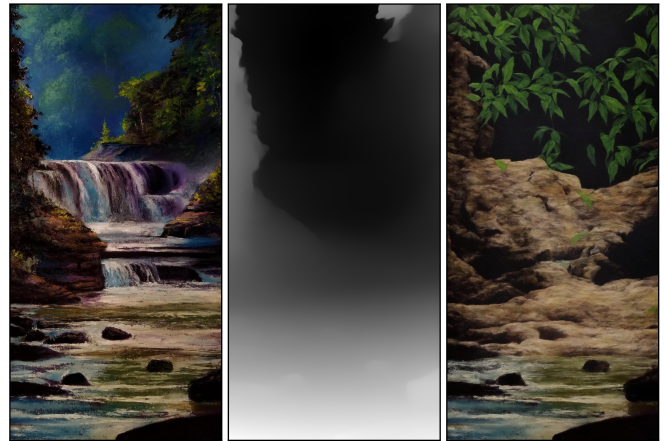
Output



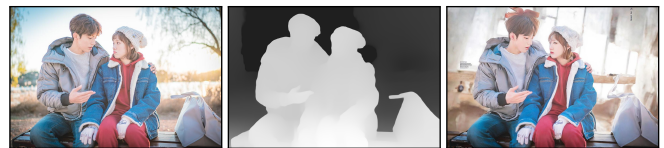
“Foliage Painting movie poster - Forgotten Road by Frank Wilson”



“Mayo Painting - Belmullet Lighthouse Mayo by Val Byrne, broken lighthouse”



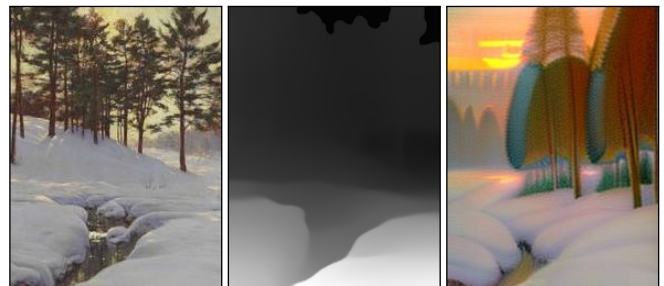
“Add Artwork | Wallhanging by Ray Easton | Artists for Conservation, Jungle Setting I”



“high fashion magazine-weightlifting-fairy”



“barley sunset roseberryytopping littleroseberry northyorkshire northeastengland cubist painting cubist paintings england nature art field nationaltrust nationalpark northyorkshiremoorsnationalpark northyorkshiremoors”

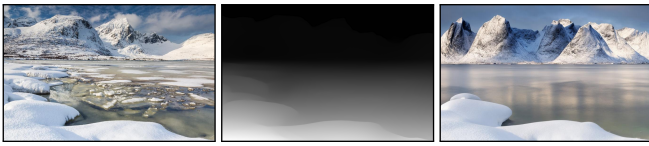


“Walter Launt Palmer (1854-1932), Sun Behind the Pines (2D).”

Image

Map

Output



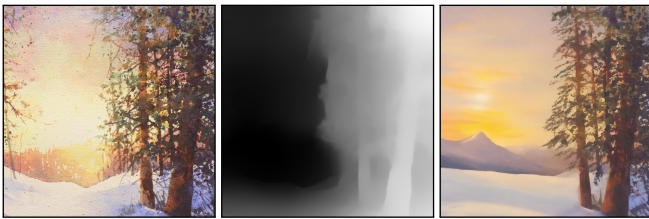
“Winter warm winters of the Lofoten Islands 10”



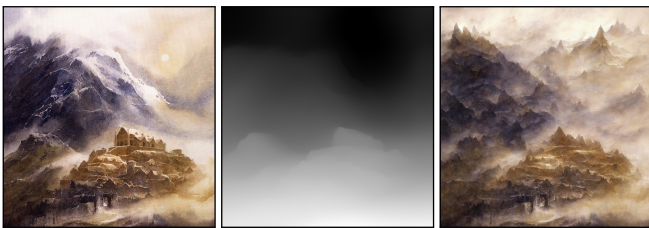
“Image result for henri martin in a forest”



“Alita Battle Angel Sword 4k Stage Prop 34”



“Snowy Mountain Sunset For Sale 4”



“The Art of Alan Lee | Fabulous Sci-Fi Realms”



“South Park Wallpapers HD burning candle angry”



“Watercolour with dinosaur by David Meldrum”

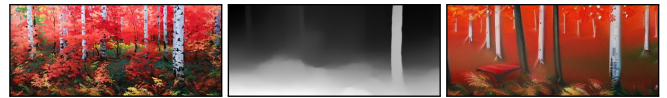


“The resurrected Christ directs fishermen where to cast their net to fish.”

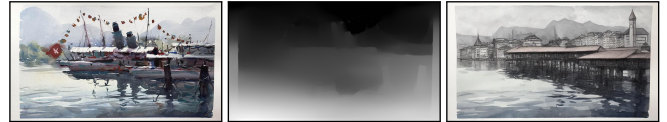
Image

Map

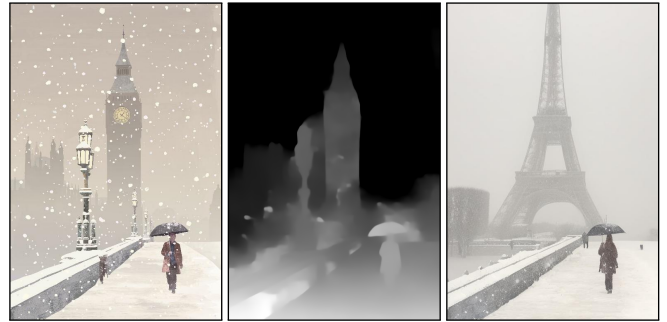
Output



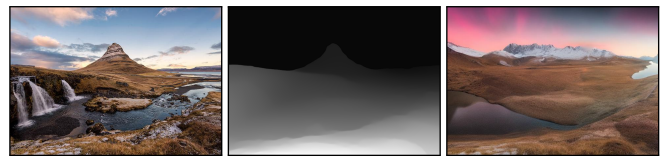
““Ruby Red Autumn” Stanislav Sidorov 30x60” Painting”



“Lucerne Dock - Drawing - Marco Bucci Art Store”



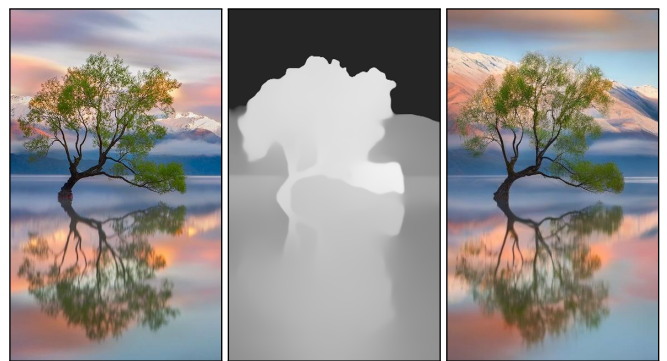
“Big Eiffel Tower in the Snow”



“Stunning landscape scene from Fstoppers ‘Photographing the World 1’ Inverted Landscape Photography Course”



“timothy-easton venice-at-sunset 3802_4240”

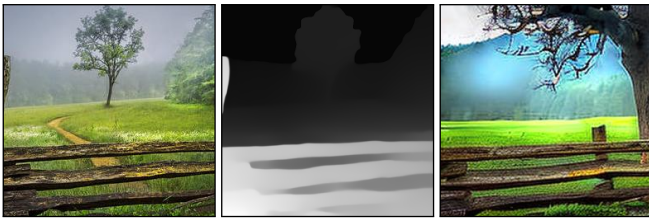


“Lake Wanaka, Otago, New Zealand, Swamp, by Karen Plimmer, by 500px”

Image

Map

Output



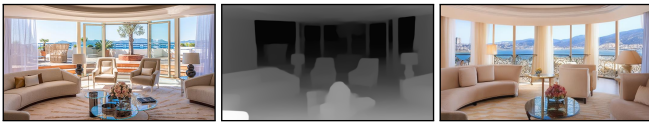
"Debra and Dave Vanderlaan - Huge Cades Cove Misty Tree"



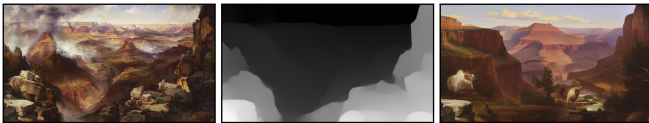
"Rise of Iron with giant ant (3)"



"Skogafoss waterfall and rainbow in the south of Iceland"



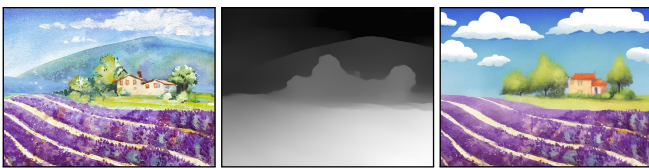
"castle hotel martinez suite penthouse living room view"



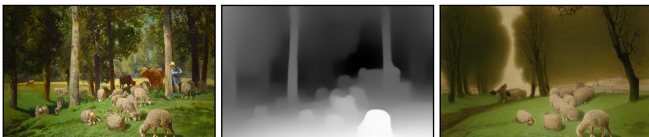
"Grand Mesa Painting - Grand Canyon Of The Colorado River at a Petting Zoo by Mountain Dreams"



"Painting Rights Managed Images - Ombre Per Strada Royalty-Free Image by Guido Borelli with City at Night"



"Beautiful blooming lavender desert with house in distance royalty free illustration"

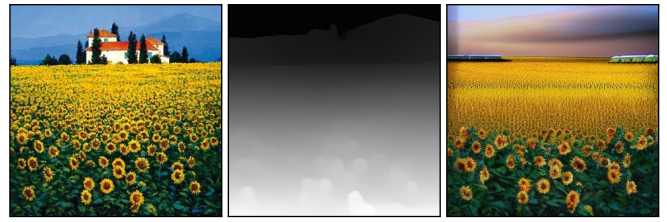


"Landscape with Sheep in Snow, Drawing by Charles Émile Jacque (1813-1894, France)"

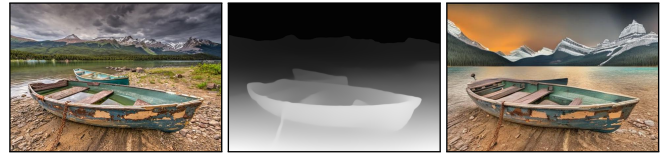
Image

Map

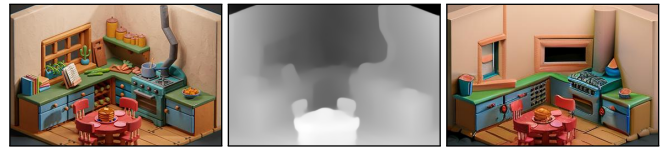
Output



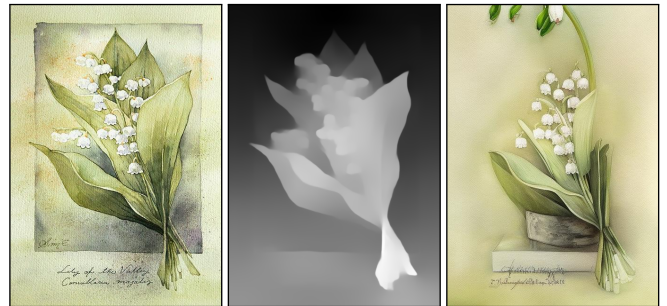
"Sunflowers Field with Snow Cover Art Print"



"Illustration of Maligne Ghosts by Wayne-Stadler-Illustration - Ships And Boats Photo Contest"



"Clay Kitchen - low poly diorama with huge explosion"



"Photo of Lily of the Valley 21x30cm"



"I Can Hear your Loud Heartbeat by Jeff Rowland"



"Cobble Hill Hot Air Balloon over Frozen Lake 1000 Piece Jigsaw Puzzle"

Image

Map

Output



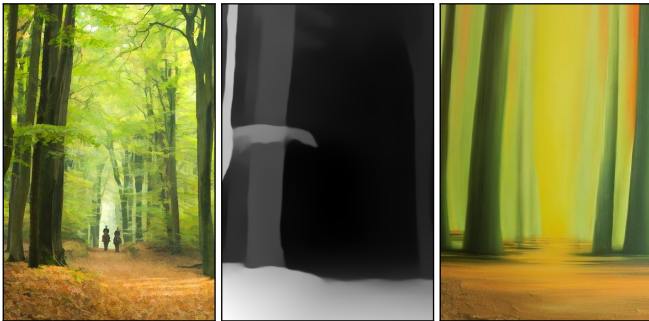
"Edward Antoon Portielje (1861 - 1949) - The hologram of a love letter"



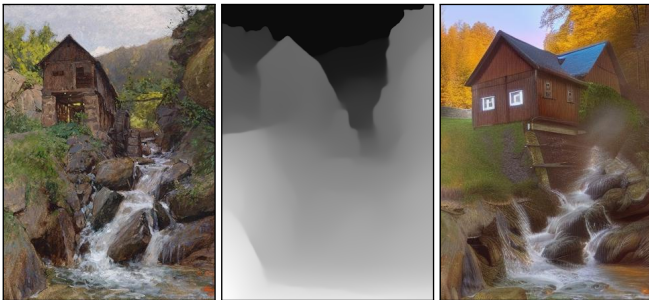
"Painting Grand, Park, Jackson, Landscapes, National, Reflections, Teton, Beaver, Ponds, Hole"



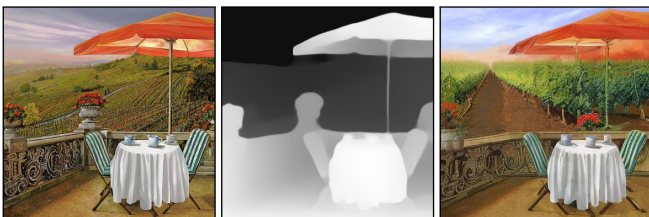
"purple bridal dresses (6)"



"Abstract forest painting"



"mühle am bach by heinrich tomec powered by steam"

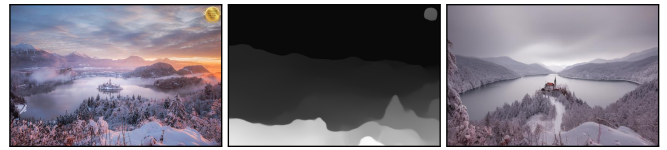


"Revolutionary War Art - Un Caffè Nelle Vigne during a Zombie Apocalypse by Guido Borelli"

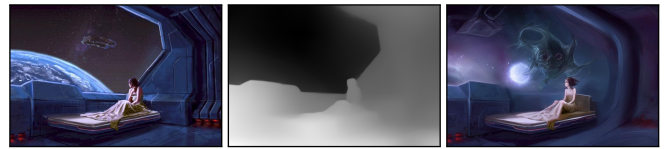
Image

Map

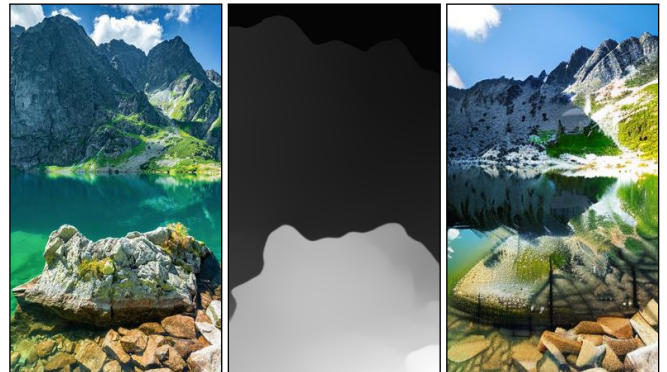
Output



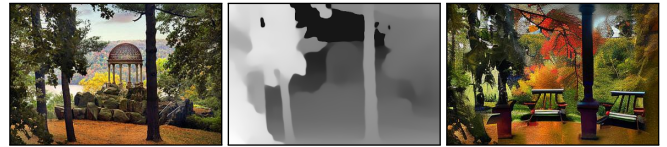
"Photograph Jaka Ivancic Bled Lake Covered With Fresh Snow on One Eyeland with Stormy Sky"



"Girl in a spaceship wallpaper - Fantasy wallpapers - turning into a monster"



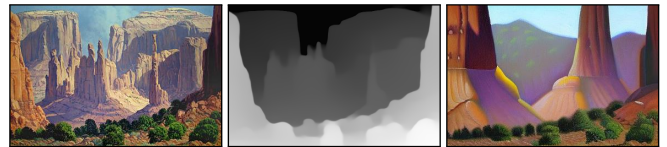
"Painting of black lake under Rysy peak, Tatra Mountains, Poland"



"Photograph - Cave Of Love In Autumn by Jessica Jenney"



"Two Collies Above A Lake With A Lighthouse Poster by Wright Barker"



"Totem Pole Painting with Orangutan - Shadows In The Valley by Randy Follis"



"The Picnic Party in the Woods by Jack Vettriano"

12. Teaser Settings

In the main paper's teaser, we used various well-known checkpoints of Stable Diffusion. This demonstrates our framework's ability to use community checkpoints. Most settings have been inspired by publicly published results on Mage [mag23].

1. **Prompt:** "Tree of life under the sea, ethereal, glittering, lens flares, cinematic lighting, artwork by Anna Dittmann & Carne Griffiths, 8k, unreal engine 5, highly detailed, intricate detailed". **Negative prompt:** "bad anatomy, poorly drawn face, out of frame, gibberish, lowres, duplicate, morbid, darkness, maniacal, creepy, fused, blurry background, crosseyed, extra limbs, mutilated, dehydrated, surprised, poor quality, uneven, off-centered, bird illustration, painting, cartoons, sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, vaginas in breasts, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, imperfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error, two body, two faces". **Checkpoint:** AbsoluteReality.
2. **Prompt:** "yellow-white-blue-purple-golden glowing mountains palace above the clouds, magical reality, high definition, 32K, dynamic lights, cinematic surrounding, intricate, natural lighting, ray tracing, bloom, extreme hdr, Octane render, unreal engine, 16K hyper realism, character design, hyper detailed, volumetric lighting, hdr, shining, vibrant, photo realism, Canon EOS 7D, Canon EF 70-200mm f/2.8L IS, vibrant colors, beautiful picture quality, breathtaking scene, focused". **Negative prompt:** "Extra limbs, extra fingers, long neck, deformed, More than one nipple per breast, pointy nose, asian, japanese, chinese, lowres, disfigured, ostentatious, ugly, oversaturated, grain, low resolution, disfigured, blurry, bad anatomy, disfigured, poorly drawn face, mutant, mutated, extra limb, ugly, poorly drawn hands, missing limbs, blurred, floating limbs, disjointed limbs, deformed hands, blurred, out of focus, long neck, long body, ugly, disgusting, bad drawing, childish, cut off cropped, distorted, imperfect, surreal, bad hands, text, error, extra digit, fewer digits, cropped, worst quality, low quality, normal quality, jpeg artifacts, signature, watermark, username, blurry, artist name, Lots of hands, extra limbs, extra fingers, conjoined fingers, deformed fingers, old, ugly eyes, imperfect eyes, skewed eyes, unnatural face, stiff face, stiff body, unbalanced body, unnatural body, lacking body, details are not clear, details are sticky, details are low, distorted details, ugly hands, imperfect hands, mutated hands and fingers, long body, mutation, poorly drawn bad hands, fused hand, missing hand, disappearing arms, disappearing thigh, disappearing calf, disappearing legs, ui, missing fingers, text, letters, illustration, painting, cartoons, sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, vaginas in breasts, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, imperfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error, two body, two faces, hat, hats, pirate hat, headgear, helmet captains hat". **Checkpoint:** DucHaiten GODofSIMP.
3. **Prompt:** "Fully In frame, 3D, centered, colorful, complete, puffy, single alien, fantasy Surreal Tree, 3d depth outer space nebulae background, 3D Art nouveau curvilinear lines, centered, no circles, no frame, dark atmosphere, curvilinear clouds, matte painting, deep color, fantastical, 3D intricate detail, complementary colors, 8k, cgsociety, artstation, hyperrealistic, cinematic, ultra hd, 4k, 8k, highly detailed cinematic global lighting, octane render, unreal engine 5". **Negative prompt:** "Multiple trees, stacked trees, trees on top of trees, illustrations, split image, 2d, painting, cartoons, sketch, worst quality, low quality, normal quality, low res, monochrome, grayscale, error, bad image, bad photo illustration, 2d, painting, cartoons, sketch, worst quality, low quality, normal quality, monochrome, grayscale, cropped, oversaturated, signature, watermark, username, artist name, error, bad image, bad photo illustration, 2d, painting, cartoons, sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, monochrome, grayscale, signature, watermark, error, bad image, bad photo black and white, monochrome, comic, text, error, cropped, letterbox, jpeg artifacts, signature, watermark, username, artist name, censored, worst quality, low quality, anime, digital illustration, 3d rendering, comic panel, scanlation, multiple views, lowres, ostentatious, ugly, oversaturated, grain, bad drawing, childish, cropped, worst quality, low quality, normal quality, jpeg artifacts, signature, watermark, username, blurry, details are not clear, details are sticky, details are low, distorted details, mutation, poorly drawn 2d, painting, cartoons, sketch, worst quality, low quality, normal quality, low res, monochrome, grayscale, cropped, oversaturated, signature, watermark, username, artist, error, bad image, bad photo deformed, distorted, disfigured, poorly drawn, bad anatomy, wrong anatomy, extra limb, missing limb, floating limbs, mutated hands and fingers, disconnected limbs, mutation, mutated, ugly, disgusting, blurry, amputation illustration, painting, cartoons, sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, vaginas in breasts, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, imperfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error, two body, two faces Poorly drawn face, poorly drawn hands, poorly drawn weapons, unnatural pose, blank background, boring background, render, unreal engine". **Checkpoint:** "AbsoluteReality".
4. **Prompt:** "the forest is on fire, forest burning, forest fire". **Negative prompt:** "digital world, warm colors-seed 1242253951, cyberpunk childrens bedroom, rusticated stone base, fat batman, -h 850 -w 600, maid". **Checkpoint:** "Stable-Diffusion v1.5".
5. **Prompt:** "A highly detailed alien landscape, alien buildings, weird colors, strange plants, xenomorphs, slime, oozing, HDR, 4k, volumetric lights, fantasy art, digital painting, beautiful, colorful, serene, intricate, eldritch, Nvidia ray tracing, Imax, slow shutter speed". **Negative prompt:** "cartoon, 3d, disfigured, bad art, deformed, extra limbs, close up, b&w, wierd colors, blurry, duplicate, morbid, mutilated, [out of frame], extra fingers, mutated hands, poorly drawn hands, poorly drawn

- face, mutation, deformed, ugly, blurry, bad anatomy, bad proportions, extra limbs, cloned face, disfigured, out of frame, ugly, extra limbs, bad anatomy, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, mutated hands, fused fingers, too many fingers, long neck, Photoshop, video game, ugly, tiling, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, mutation, mutated, extra limbs, extra legs, extra arms, disfigured, deformed, cross-eye, body out of frame, blurry, bad art, bad anatomy, double face, double face realistic, semi-realistic, cgi, 3d, render, sketch, cartoon, drawing, anime, cropped, worst quality, low quality, jpeg artifacts, ugly, duplicate, morbid, mutilated, out of frame, extra fingers, mutated hands, poorly drawn hands, poorly drawn face, mutation, deformed, blurry, dehydrated, bad anatomy, bad proportions, extra limbs, cloned face, disfigured, gross proportions, malformed limbs, missing arms, missing legs, extra arms, extra legs, fused fingers, too many fingers, long neck”. **Checkpoint:** Realistic Vision V2.
6. **Prompt:** “no humans, scenery, flower A whimsical illustration of a rainbow connecting two vibrant and cheerful worlds, signifying the connection between people who find humor in life’s moments”. **Negative prompt:** “blur haze child, loli, paintings, sketches, worst quality, low quality, normal quality, lowres, normal quality, monochrome, grayscale, skin spots, acnes, skin blemishes, age spot, glans, mutated hands, poorly drawn hands, blurry, bad anatomy, extra limbs, lowers, bad hands, missing fingers, extra digit, bad hands, missing fingers worst quality, low quality, sketch, watermark, text copyright signature, cut out, cgi, hands, Two bodies, Two heads, doll, extra nipples, bad anatomy, blurry, fuzzy, extra arms, extra fingers, poorly drawn hands, disfigured, tiling, deformed, mutated, out of frame, cloned face illustration, painting, cartoons, sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, vaginas in breasts, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, imperfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error illustration, 3d, 2d, painting, cartoons, sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, vaginas in breasts, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, imperfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error, bad image, bad photo worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, vaginas in breasts, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, imperfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error, painting by bad-artist sketch, worst quality, low quality, normal quality, lowres, bad anatomy, bad hands, monochrome, grayscale, collapsed eyeshadow, multiple eyeblows, vaginas in breasts, cropped, oversaturated, extra limb, missing limbs, deformed hands, long neck, long body, im-
- perfect, bad hands, signature, watermark, username, artist name, conjoined fingers, deformed fingers, ugly eyes, imperfect eyes, skewed eyes, unnatural face, unnatural body, error”. **Checkpoint:** DucHaiten StyleLikeMe.

References

- [AFL23] AVRAHAMI O., FRIED O., LISCHINSKI D.: Blended latent diffusion. *ACM Transactions on Graphics* 42, 4 (July 2023), 1–11. 2, 4
- [aS23] AT STABILITYAI D. L.: DeepFloyd IF: a novel state-of-the-art open-source text-to-image model with a high degree of photorealism and language understanding. <https://www.deepfloyd.ai/deepfloyd-if>, 2023. Retrieved on 2023-11-08. 2
- [AUT22] AUTOMATIC1111: Stable Diffusion Web UI. <https://github.com/AUTOMATIC1111/stable-diffusion-webui>, Aug. 2022. 3
- [BA83] BURT P. J., ADELSON E. H.: A multiresolution spline with application to image mosaics. *ACM Trans. Graph.* 2, 4 (oct 1983), 217–236. 3
- [BHE23] BROOKS T., HOLYNSKI A., EFROS A. A.: Instructpix2pix: Learning to follow image editing instructions, 2023. 2, 16
- [CVSC22] COUAIRON G., VERBEEK J., SCHWENK H., CORD M.: Diffedit: Diffusion-based semantic image editing with mask guidance, 2022. 1
- [HJA20] HO J., JAIN A., ABBEEL P.: Denoising diffusion probabilistic models, 2020. 9
- [KAAL22] KARRAS T., AITTALA M., AILA T., LAINE S.: Elucidating the design space of diffusion-based generative models, 2022. 9
- [LRLZ22] LIU L., REN Y., LIN Z., ZHAO Z.: Pseudo numerical methods for diffusion models on manifolds, 2022. 9
- [LZB*22] LU C., ZHOU Y., BAO F., CHEN J., LI C., ZHU J.: Dpm-solver: A fast ode solver for diffusion probabilistic model sampling in around 10 steps, 2022. 9
- [mag23] Mage space: An ai image generation platform, 2023. 31
- [PGB03] PÉREZ P., GANGNET M., BLAKE A.: Poisson image editing. *ACM Trans. Graph.* 22, 3 (jul 2003), 313–318. 3
- [RBK21] RANFTL R., BOCHKOVSKIY A., KOLTUN V.: Vision transformers for dense prediction. *ICCV* (2021). 5
- [RBL*22] ROMBACH R., BLATTMANN A., LORENZ D., ESSER P., OMMER B.: High-resolution image synthesis with latent diffusion models, 2022. 2
- [RLH*22] RANFTL R., LASINGER K., HAFNER D., SCHINDLER K., KOLTUN V.: Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 44, 3 (2022). 2, 5, 16
- [SME22] SONG J., MENG C., ERMON S.: Denoising diffusion implicit models, 2022. 9
- [Sta23] STABILITY AI LTD: Stability.ai blog: Stable diffusion v2 release. <https://stability.ai/blog/stable-diffusion-v2-release>, 2023. Accessed: April 25, 2023. 4
- [ZC23] ZHANG Q., CHEN Y.: Fast sampling of diffusion models with exponential integrator, 2023. 9
- [ZIE*18] ZHANG R., ISOLA P., EFROS A. A., SHECHTMAN E., WANG O.: The unreasonable effectiveness of deep features as a perceptual metric, 2018. 1