

Creative fusing of images

Ken Kahn - [home page](#)

19 September 2025



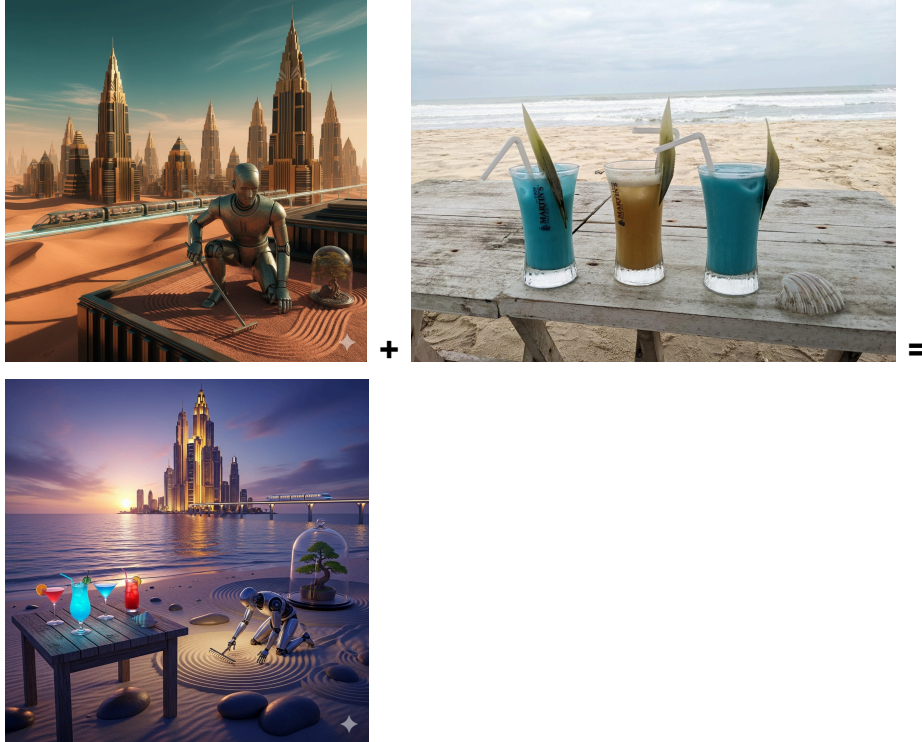
I was curious how well generative AI systems could create images that blend elements from two different images.

Prompting-by-example app development

People have been exploring programming by example systems for over 50 years. See, for example, the book [Your Wish is My Command](#). (I have a chapter in it.) I experimented with applying some of those ideas to using chatbots to create apps that use generative AI.

[I started with uploading two images and making a series of requests to Gemini 2.5 Pro:](#)

1. Describe each image in detail
2. Now combine the two descriptions to make a single coherent image description
3. Make an image from that description [*a collage of the images was produced*]
4. I meant combine the image descriptions to come up with a new image description that captures as many elements of the original two images into a single scene.
5. Now render that



Happy with this result, I then prompted:

Can you make an app that uses the Gemini API to work like this? Two images uploaded and one produced.

The app was buggy and Gemini struggled to fix it. And Gemini became confused about how to access the Gemini API from a Gemini app. I gave it an example of this but bugs persisted.

Instead of giving up, [I started a new session with Gemini where I uploaded the buggy code and the most recent error message](#). It took a while to clarify that I wanted a corrected Gemini app but Gemini produced a working app.

I then asked for series of improvements:

1. When the fused image has been generated then display the result as an "equation":
Picture 1 + Picture 2 $\approx$ Fused image
2. And if a user clicks on any of the images the description is displayed. Add a note at the bottom explaining this feature.
3. The prompt to fuse the images should emphasize that the aspects of both images should be fused but the images themselves should not just be assembled. [*Motivated because it sometimes produced collages.*]
4. Let's add a button to save the equation (along with the ability to see the descriptions) to a downloadable HTML file.
5. The descriptions sometimes contain markdown - please render that nicely

6. The saved HTML should use the width of the tab to display the equation. [*I should've said dimensions, not just the width.*]
7. Let's add a button next to "save result" that is "Try again". This will generate another fusion description where the prompt asks for something different from any of the previously generated descriptions.
8. The saved HTML should make the equation bigger. It should make good use of the height of the tab as well.
9. When I used the try again button I got an image that showed the main elements of each image together in the image but not really blended or fused creatively.

Evolution of the “fuse prompt”

The first version of the prompt that Gemini generated to combine two image descriptions was:

You are an imaginative artist creating a prompt for an AI image generator. Your task is to combine two sets of concepts into a single, new, and coherent scene. Be creative and merge the elements seamlessly.

Concepts 1: "\${summary1}"

Concepts 2: "\${summary2}"

Now, create the final, detailed, and visually rich prompt for the new scene. Do not describe your process, just output the final prompt.

This eventually evolved into this code generated by Gemini:

```
let prompt = `You are an imaginative artist creating a prompt
for an AI image generator. Your task is to combine two sets of concepts
into a single, new, and coherent scene. Both concepts must be represented
with equal importance. The goal is a true fusion of styles, subjects, and
atmospheres, not a simple collage or side-by-side placement of the
elements from the source images. The final image should look like a
single, unified piece of art.\n\nConcepts 1: "${summary1}"\n\nConcepts 2:
"${summary2}"\n\n`;

if (history.length > 0) {
  prompt += `You have already tried the following
prompt(s):\n`;
  history.forEach((p, index) => {
    prompt += `Attempt ${index + 1}: "${p.substring(0,
250)}"..."\n`;
  });
}
```

```
prompt += `\\nCRITICAL INSTRUCTION FOR THIS NEW ATTEMPT:  
The previous result may have failed to balance the two concepts, with one  
dominating the other. For this new version, you MUST give EQUAL or even  
GREATER creative weight to the underrepresented concept from the previous  
attempt. Do not simply repeat one idea. Generate a completely new,  
imaginative, and BALANCED fusion.\\n`;  
}  
  
prompt += `\\nNow, create the final, detailed, and visually  
rich prompt for the new scene. Do not describe your process, just output  
the final prompt.`;
```

Results

Here is a series of image fusions that started with a photo of a unicorn fruit plate (a joint effort of myself and my grandchildren) and an image generated by ChatGPT. The app first generated this “equation”:



Clicking on “Try again” produced these fusions:







Notice how the generated images often include hands and candle sticks. Here is the entire image that it processed:



Combining works of art



+



≈







Accessing the image descriptions

The app produces web pages where you can click on the input images and the output image and read the descriptions used. Here is [an example](#). And here is [the app](#).

About me



I've been researching AI, creativity, and education for fifty years. I started when working on my phd from the MIT AI lab. Most recently I wrote a book about how anyone can use chatbots to co-create apps, adventures, illustrated stories, and discussions.

[*The Learner's Apprentice: AI and the Amplification of Human Creativity*](#)

The book presents many examples of tinkering with chatbots as well as co-creating apps and games with chatbots.

You can follow me on [LinkedIn](#), [Facebook](#), [Threads](#), or [BlueSky](#).