

Hands-on-Lab: Image Generation in Action

Estimated time needed: 45 minutes

Welcome to the hands-on lab image generation in action!

What is image generation in generative AI?

Images are vital to visual details because they comprehensively represent concepts, data, and information that users can easily grasp. For effective communication, images are powerful elements that add depth, context, and visual appeal to the information.

Generative AI models have advanced significantly in recent years, enabling the creation of new and realistic images using artificial intelligence techniques. Using generative AI-powered models and tools, you can generate and modify high-quality images that are difficult to distinguish from real photographs.

This lab focuses on the image-generation capabilities of generative AI. In this lab, you will look at the three image generation tools: Stable Diffusion, Hugging Face, and Simplified. You will explore these tools and use them to create images from textual descriptions, generate captions for images, and create thumbnails from the textual description.

Learning Objectives

After completing this lab, you will be able to:

- Explore the image generation capabilities of generative AI
- Create realistic images from text using Stable Diffusion XL
- Generate an image caption using Hugging Face's image-to-text
- Create small thumbnail images representing larger images using Simplified

Exercise 1: Create realistic images from text using Stable Diffusion Online

In this exercise, you will create realistic images using Stable Diffusion Online based on the textual prompts you provide to the generative AI tool. Stable Diffusion Online is powered by Stable Diffusion XL, a state-of-the-art tool offered by Stability AI. It takes English text as an input, called the "text prompt" and generates images based on the text description.

Step 1: Navigate to Stable Diffusion Online

1. Click [Stable Diffusion Online](#) to launch it.

If you cannot access Stable Diffusion Online by clicking on the link, copy the following URL and paste it into a browser.

1. <https://stablediffusionweb.com/>

Copied!

2. Once launched, you will view the webpage that provides an interface for generating images for free. Click on **Get Started Now** to navigate to the AI image generator section.

Stable Diffusion Online

Stable Diffusion is a latent text-to-image diffusion model capable of generating photo-realistic images given any text input, cultivates autonomous freedom to produce incredible imagery, empowers billions of people to create stunning art within seconds.

Stable Diffusion is a deep learning model that generates images from text descriptions. The model is considered to be a part of the ongoing AI spring, which refers to the rapid development of artificial intelligence technologies.

Get started with Stable Diffusion AI Image generator, for [FREE](#).

[Get Started Now](#)

► [AI Art Generator PRO](#)

3. You will be navigated to the Stable Diffusion AI Image Generator section.

Stable Diffusion AI Image Generator

Stable Diffusion powered AI image generator that creates images from textual descriptions

Prompt

Enter your negative prompt here

Styles

cinematic-default ▾

Advanced Options ▶

Generate

Step 2: Generate the desired image

1. Consider a context for the realistic image.
2. Write the prompt or statement in the **Prompt** field.
For example,

1. 1

1. Create a captivating and scientifically accurate image of the solar system. Include depictions of the sun and the planets.

Copied!

Note: Include specific details about the required image for better results.

Stable Diffusion AI Image Generator

Stable Diffusion powered AI image generator that creates images from textual descriptions

Prompt

Create a captivating and scientifically accurate image of the solar system. Include depictions of the sun and the planets.

Styles

cinematic-default

Advanced Options

Generate

3. From the **Styles** drop-down list, you can select the desired style in which you want to generate the image. Let's select the **sai-cinematic** style.

Stable Diffusion AI Image Generator

Stable Diffusion powered AI image generator that creates images from textual descriptions

The screenshot shows the user interface of the Stable Diffusion AI Image Generator. On the left, under the 'Prompt' section, there is a text input field containing the instruction: 'Create a captivating and scientifically accurate image of the solar system. Include depictions of the sun and the planets.' Below this is a 'Advanced Options' button. In the center, there is a 'Styles' dropdown menu with the following options listed: 'cinematic-default' (selected), '✓ cinematic-default', 'sai-3d-model', 'sai-analog film', 'sai-anime', and 'sai-cinematic'. The 'sai-cinematic' option is highlighted with a red rectangle. At the bottom of the interface, there is a large 'Generate' button.

4. Next, let's access the **Advanced** options to specify other attributes for the desired image. Click the left-pointing arrow for the Advanced options.

Stable Diffusion AI Image Generator

Stable Diffusion powered AI image generator that creates images from textual descriptions

Prompt

Create a captivating and scientifically accurate image of the solar system. Include depictions of the sun and the planets.

Styles

sai-cinematic

Advanced Options

Generate

5. Let's enter the negative prompt, which is used to specify any elements or attributes that you do not want in the generated image. Enter the following prompt in the **Negative Prompt** field.

1. 1

1. Inaccurate Proportions, Unrealistic Colors, Fantasy Elements

Copied!

Stable Diffusion AI Image Generator

Stable Diffusion powered AI image generator that creates images from textual descriptions

Prompt

Create a captivating and scientifically accurate image of the solar system. Include depictions of the sun and the planets.

Styles

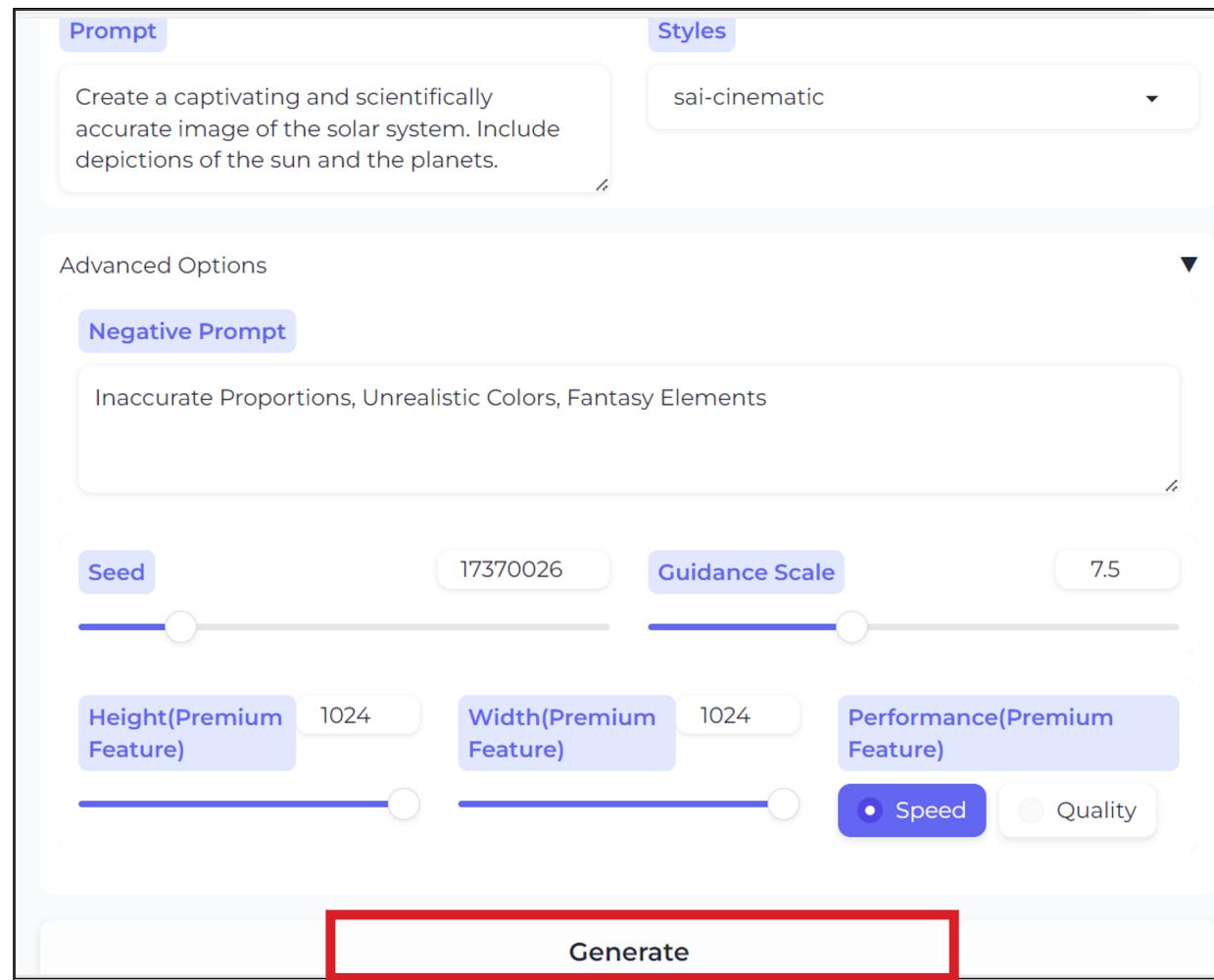
sai-cinematic

Advanced Options

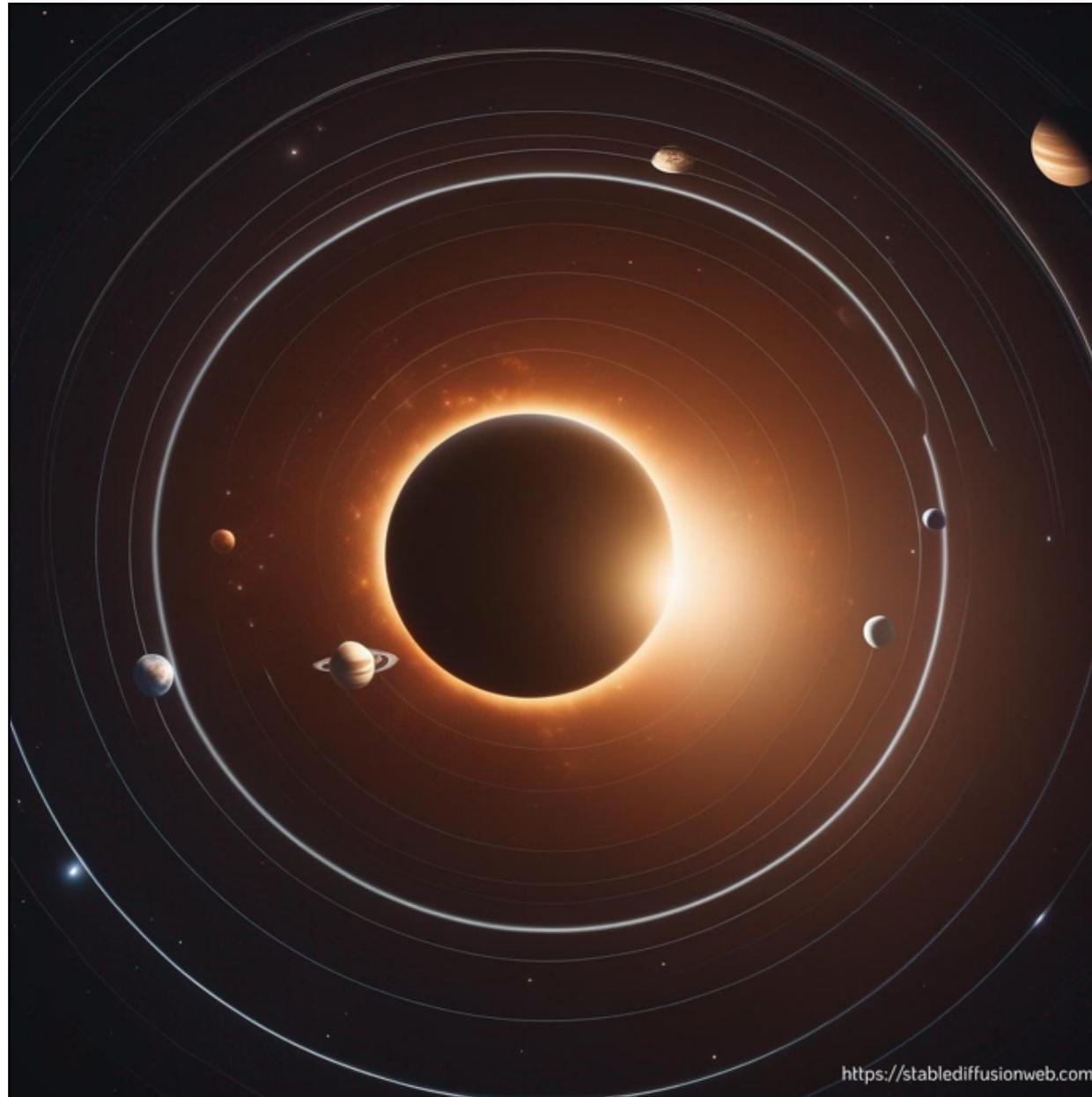
Negative Prompt

Inaccurate Proportions, Unrealistic Colors, Fantasy Elements

6. Click **Generate** to send the message to the model.



7. After a few seconds, the image will be generated. The image that was generated at our end as the output is displayed here.



<https://stablediffusionweb.com>

Note: You can continue interacting with Stable Diffusion to generate more and more images and iterations. You can click Generate image to regenerate the images for the description. You can also modify the description, and Stable Diffusion will show the revised set of images on the display field.

Note: The generated images may not be very good in quality or precision (in relation to the shared prompt) as this is a free image generator. To generate faster and better-quality images, you may consider upgrading to premium (pricing) plans.

Step 3: Try yourself

Now, you can try to describe and generate realistic images of your choice using Stable Diffusion.

Exercise 2: Generate image caption using Hugging Face

In the digital age, image captions have gained even more importance due to their role in improving accessibility for people with visual impairments. Screen readers and other assistive technologies can read these captions aloud, allowing individuals who cannot see the image to understand its content. Generative AI models simplify the task of predicting the captions for the given image.

In this exercise, you will generate the caption for an image using GPT-powered Hugging face.

What is Hugging Face?

Hugging Face is well-known for creating and maintaining the "transformers" library, which provides a wide range of pre-trained NLP models, including the famous GPT (Generative Pre-trained Transformer) series models.

The Hugging Face "transformers" library allows researchers, developers, and data scientists to easily access and use state-of-the-art NLP models for various tasks, such as text generation, sentiment analysis, language translation, and more. The library provides interfaces and pre-trained models for both PyTorch and TensorFlow, making it convenient for users to work with their preferred deep-learning framework.

Let's generate captivating captions for the images using the Hugging Face's image-to-text model.

Step 1: Select the image for caption generation

Select the image for which you want to generate the caption.

Step 2: Launch Hugging Face

1. Click [Hugging Face](#) to launch it.

Note: If you cannot access Hugging Face by clicking on the link, copy the following URL and paste it into a browser.

1. 1

1. <https://huggingface.co/tasks/image-to-text>

Copied!

2. Once launched, you can directly use the platform for caption generation.

Note: Sign up or login is not mandatory to use the image-to-text model of Hugging Face.

Hugging Face Models Datasets Spaces Docs Solutions Pricing Log In Sign Up

< Tasks

Image-to-Text

Image to text models output a text from a given image. Image captioning or optical character recognition can be considered as the most common applications of image to text.

Inputs

Output

Detailed description
a herd of giraffes and zebras grazing in a field

Compatible libraries
js Transformers.js

Image-to-Text demo
using Salesforce/blip-image-captioning-base
Image-to-Text

Drag image file here or click to browse from your device

This model can be loaded on the Inference API on-demand.

JSON Output Maximize

Models for Image-to-Text Browse Models (265)

Salesforce/blip-image-captioning-large
Image-to-Text Updated Aug 1 1.17M 357
Note A robust image captioning model.

nlpconnect/vit-gpt2-image-captioning

About Image-to-Text

Use Cases

3. You can work on the model workspace on the right side of the home page, as shown below.

The screenshot shows the Hugging Face Model Hub interface. At the top, it displays "Compatible libraries" and a "Transformers.js" button. Below this, the "Image-to-Text demo" section is shown, using the "Salesforce/blip-image-captioning-base" model. It features a dashed box for dragging images and a "Maximize" button. A note indicates the model can be loaded on demand. The "JSON Output" button is also visible. In the bottom section, there's a "Models for Image-to-Text" heading and a "Browse Models (265)" button. A specific model, "Salesforce/blip-image-captioning-large", is highlighted, showing its details: "Image-to-Text", "Updated Aug 1", "1.17M", and "357" likes.

Compatible libraries

JS Transformers.js

⚡ Image-to-Text demo

using [Salesforce/blip-image-captioning-base](#)

Image-to-Text

Drag image file here or click to browse from your device

This model can be loaded on the inference API on-demand.

JSON Output Maximize

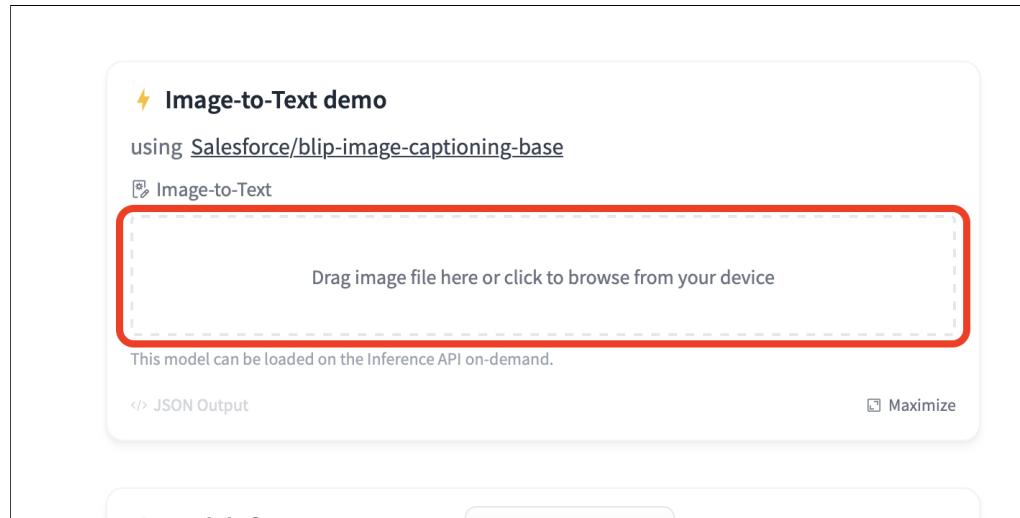
Models for Image-to-Text [Browse Models \(265\)](#)

Salesforce/blip-image-captioning-large

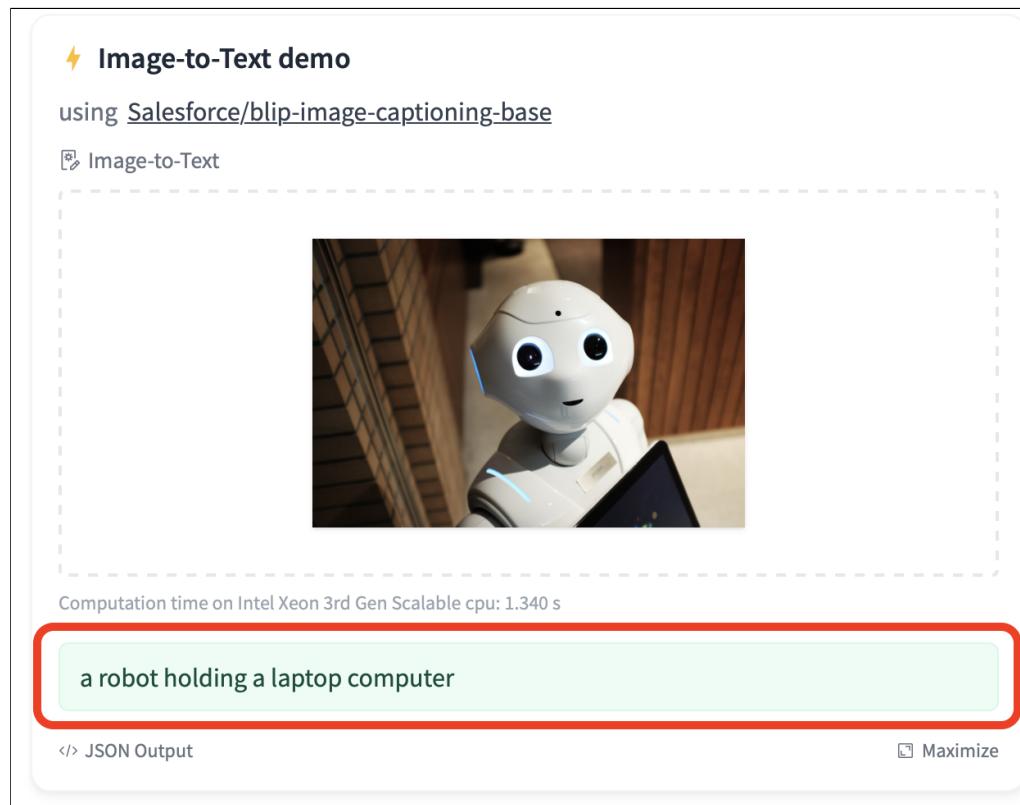
Image-to-Text • Updated Aug 1 • 1.17M • 357

Step 3: Generate the caption for the image

1. Insert the image for which you want to generate the caption in the prompt field. You can drag the image into the field or use the field option to select the image from your system or device.



2. The system then generates the caption within seconds, as shown in the screenshot below.



3. Now, you must try generating captions for three images using Hugging Face's image-to-text generation.

Exercise 3: Create small thumbnail images representing larger images using Simplified [Optional]

Note: This exercise is optional as using the Simplified tool may require credits or paid subscription.

Thumbnails are small images that provide a preview of larger content, such as videos, images, or documents. This visual preview improves navigation, enhances the user experience, aids content organization, and contributes to marketing efforts across various digital platforms.

In this exercise, you will create a small thumbnail image representing a larger image using the Simplified tool.

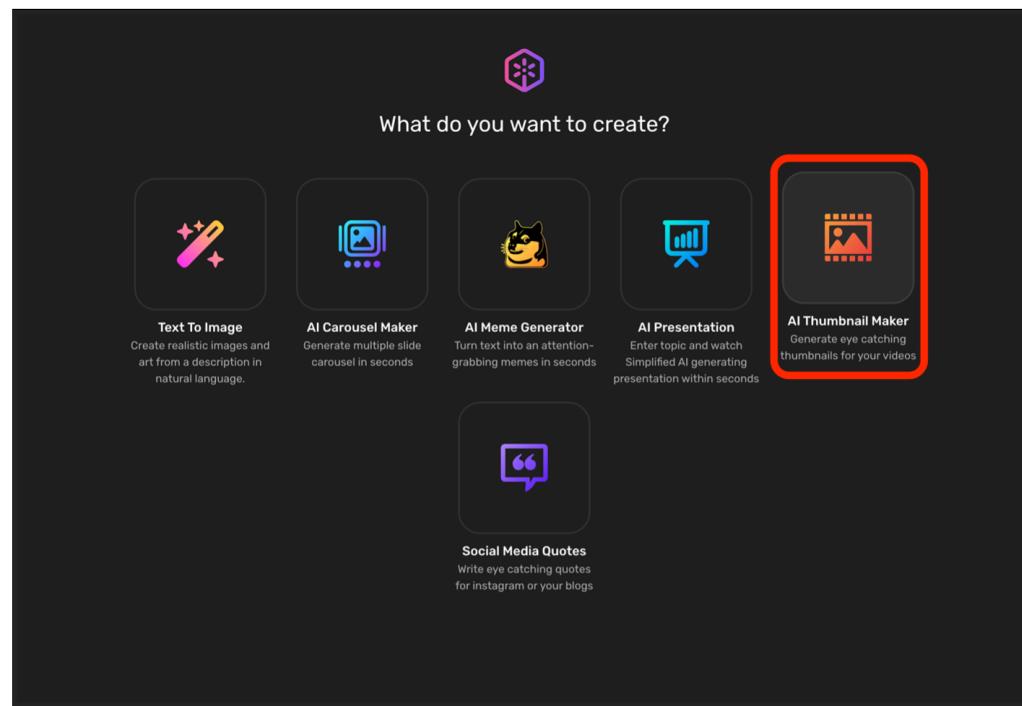
What is Simplified?

Simplified is a powerful web app that uses Generative AI models to create art, images, text, and more. It helps create text-to-image, carousels, memes, presentations, thumbnails, and quotes.

Let's generate a thumbnail representing the images or videos using Simplified AI Thumbnail Maker.

Step 1: Launch Simplified AI Thumbnail Maker

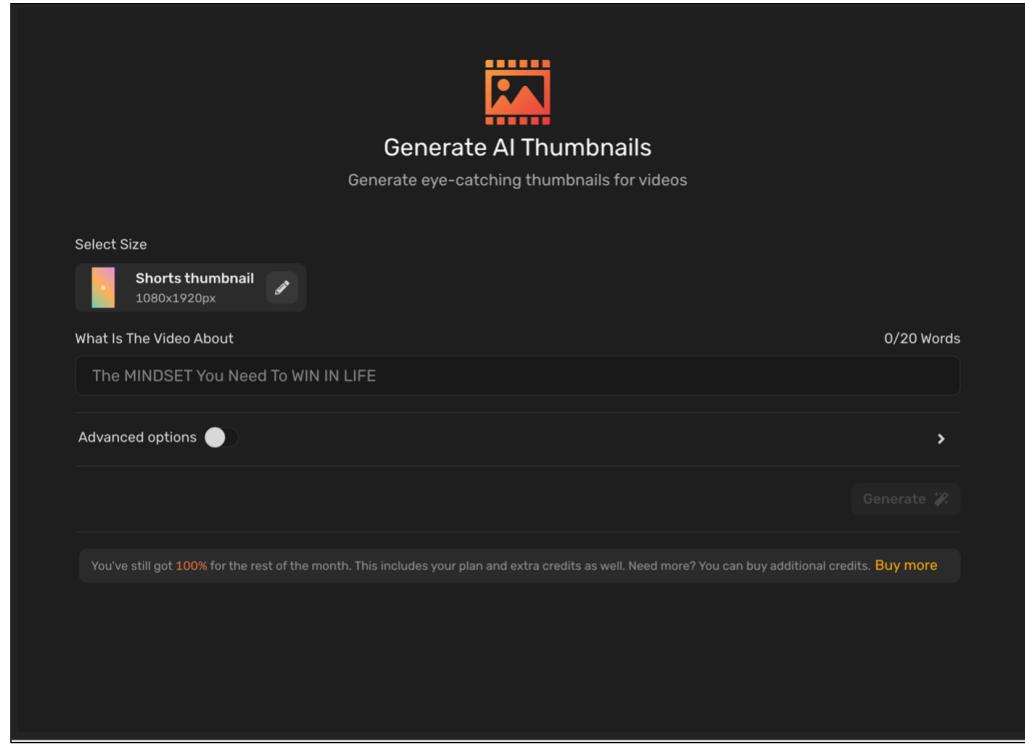
1. Click [Simplified](#) to launch its generative AI apps. Select the **AI Thumbnail Maker** from the available options, as shown in the screenshot below.



Alternatively, you can directly launch the AI Thumbnail Maker by clicking [Simplified AI Thumbnail Maker](#).

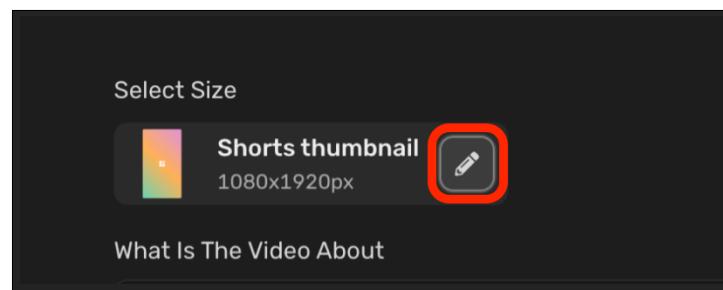
2. Once launched, you will view the **AI Thumbnail Maker workspace**, as shown in the screenshot below.

Note: You need to sign up or login to use Simplified's AI Thumbnail Maker. You get unlimited free credits for a month after signing up, followed by limited free credits to generate images using Simplified AI.

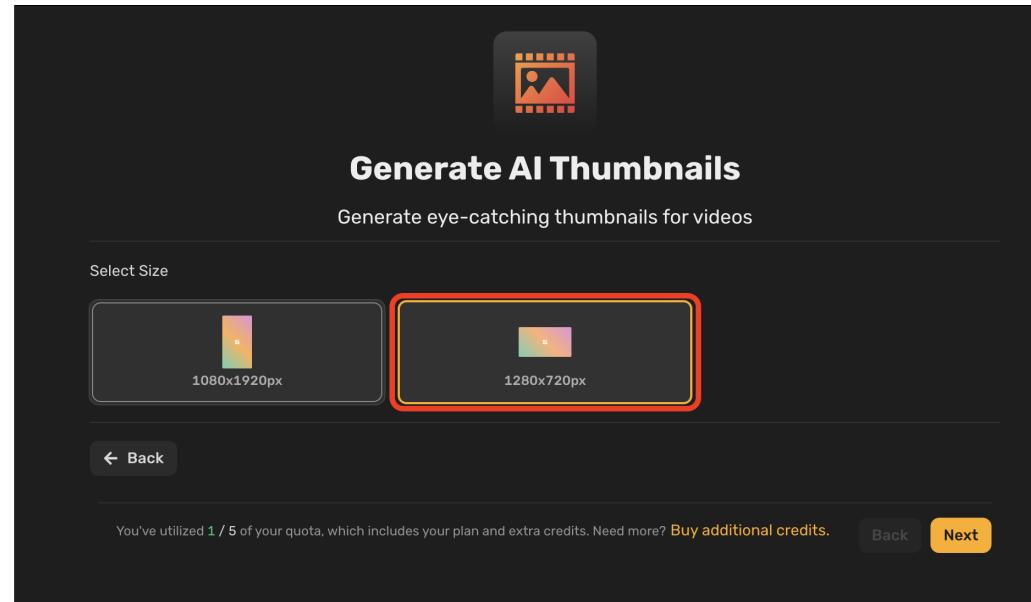


Step 2: Exercise the thumbnail generation

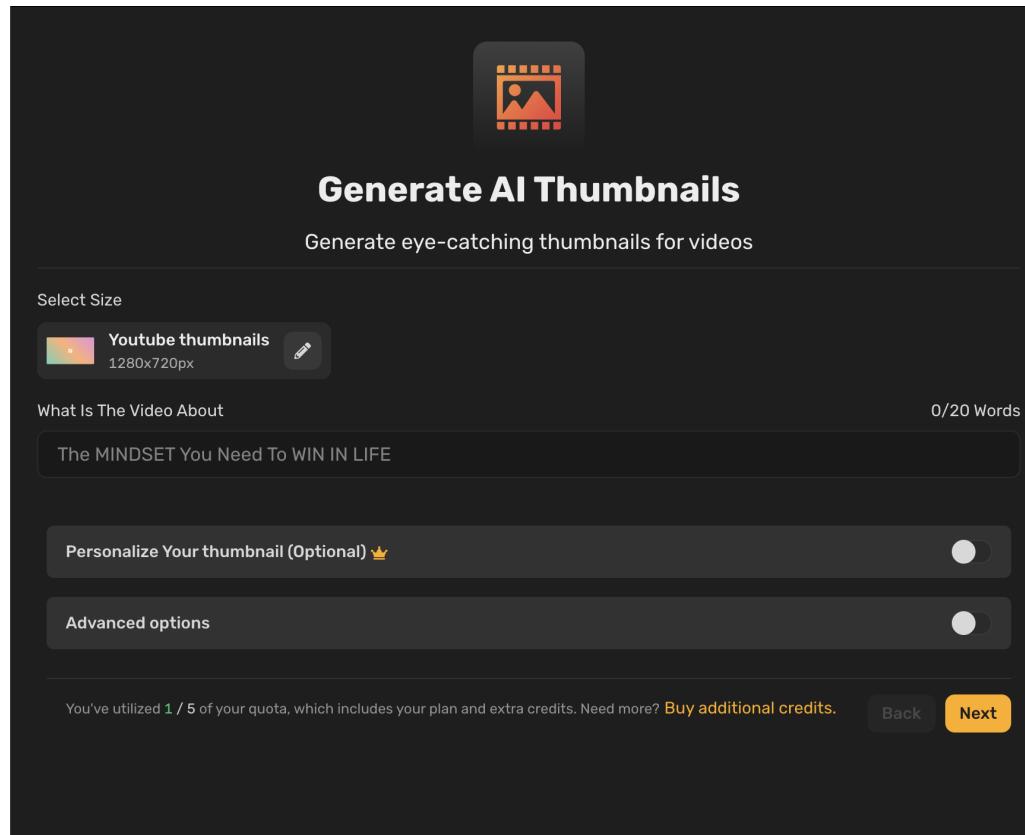
1. Click the pencil icon to select the thumbnail size.



2. You will see two options for the thumbnail size: Shorts thumbnail (1080x1920px) and Youtube thumbnails (1280x720px).



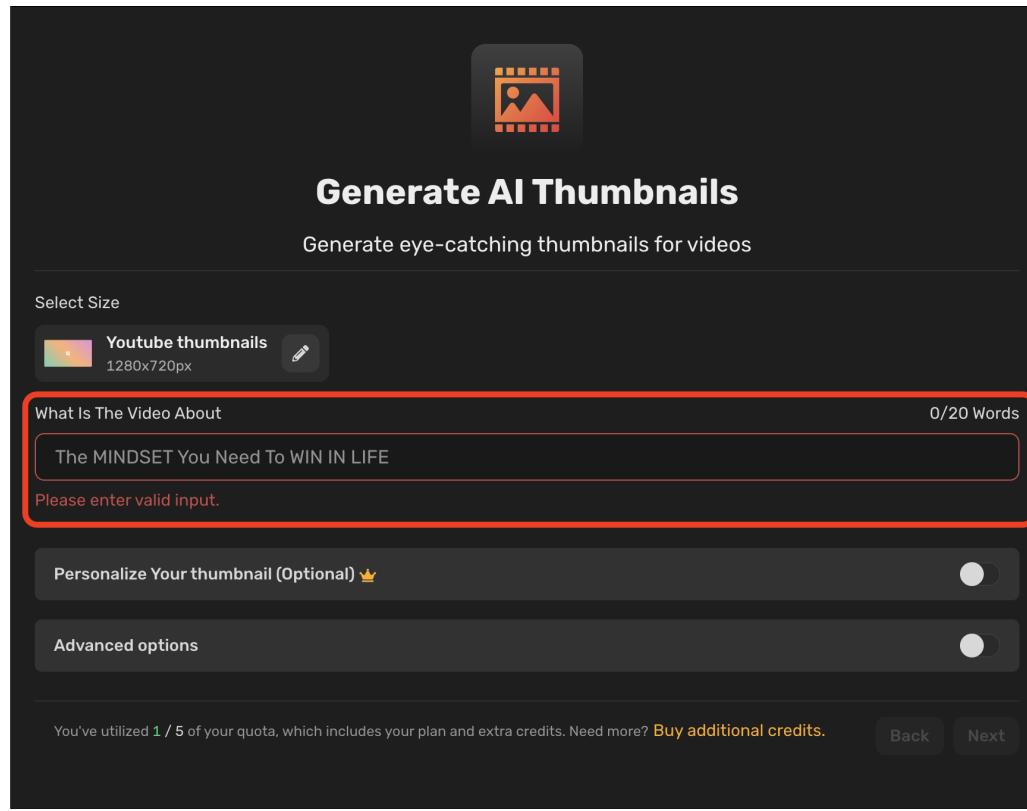
3. Select the **Youtube thumbnails** for this exercise. You will see the updated selection below **Select Size**, as shown in the screenshot below.



4. You must provide a description for the thumbnail in the field **What Is The Video About** in 20 words.
For example,

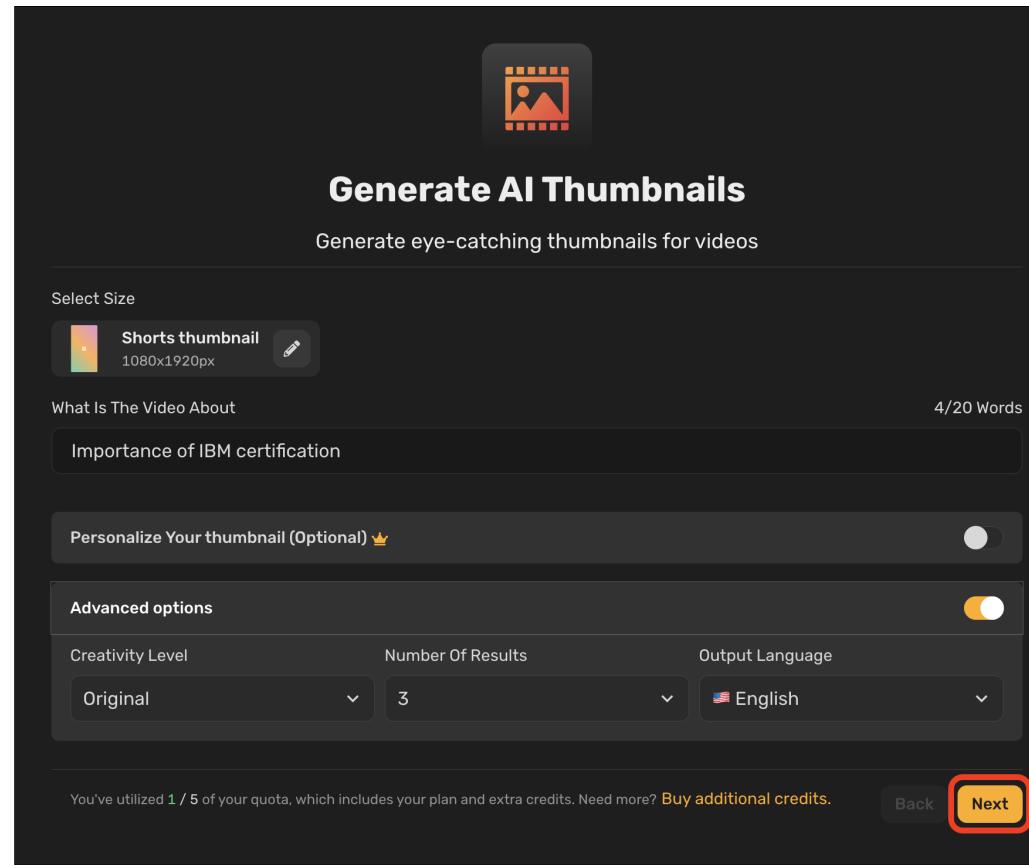
1. 1
1. Importance of IBM certification

Copied!

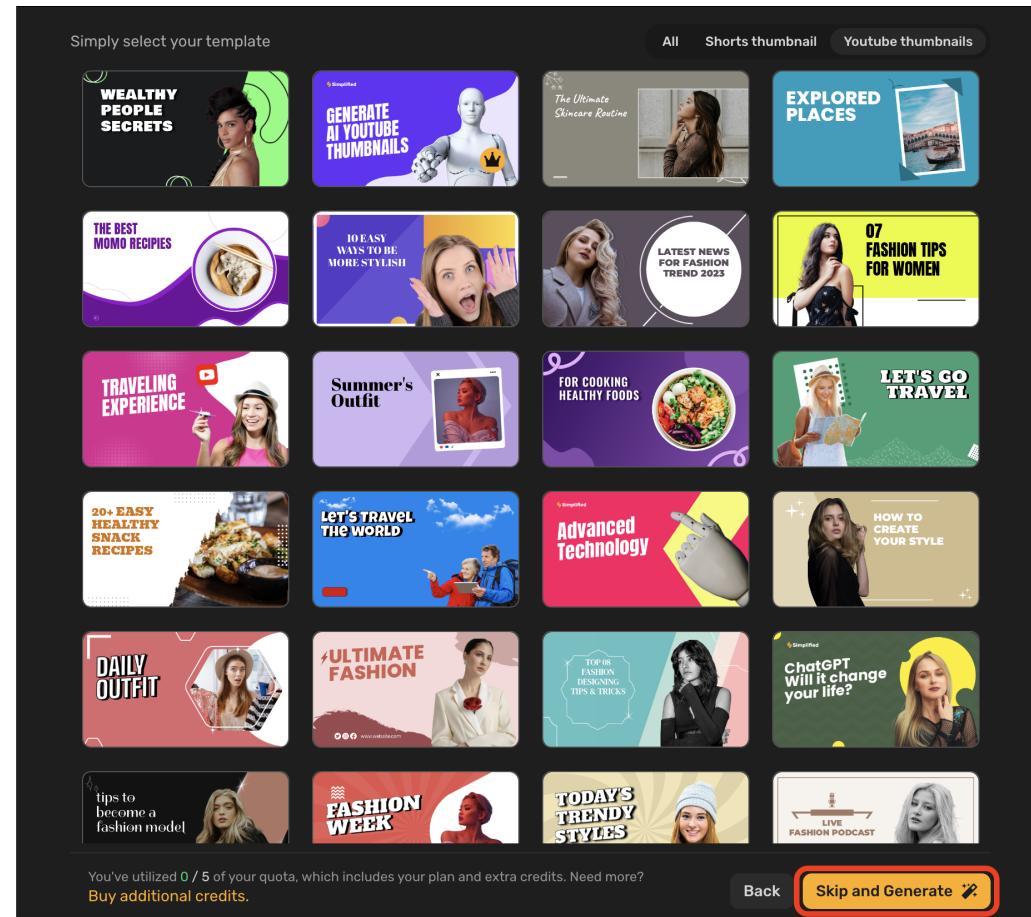


Note: You can also set the advanced options, like the number of results, creativity level, and output language.

5. Now, click **next** and you will get various template options to choose from.

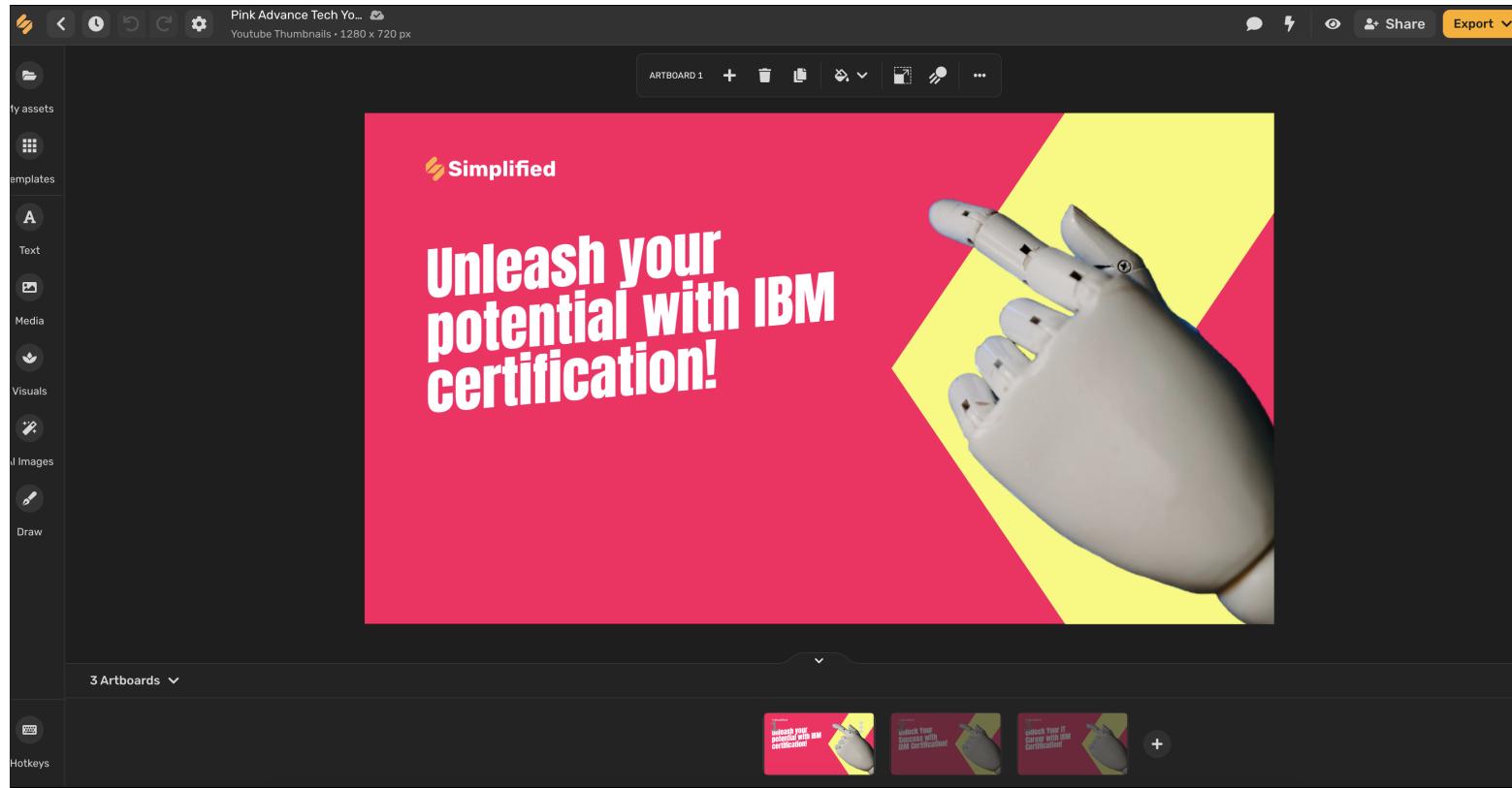


6. Select the desired template from available template options and click **Generate** to get the thumbnail results.

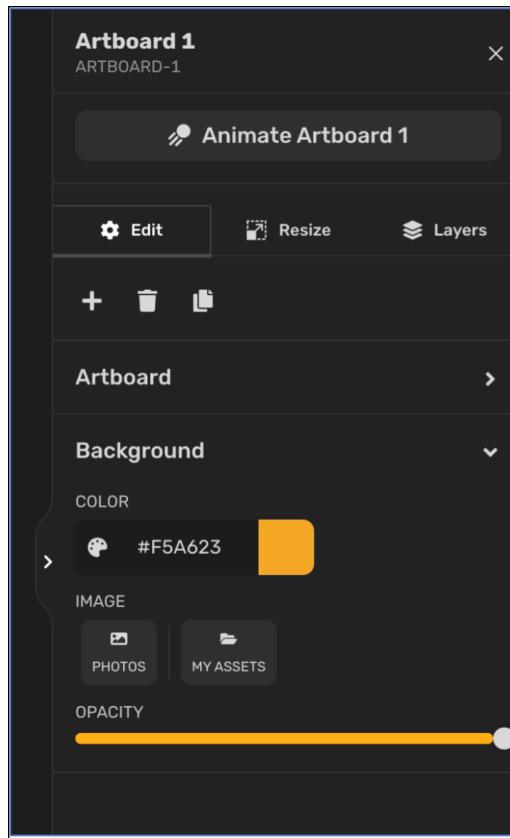


Note: You can skip the template selection, which will generate a random result.

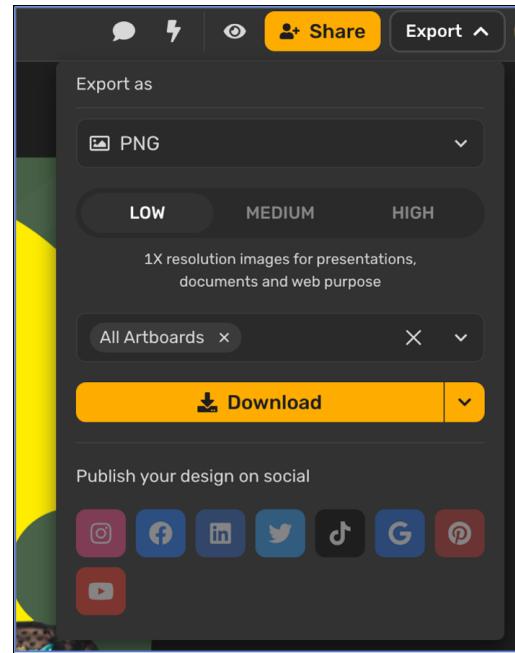
7. The page will refresh to a new window and show the result options, as shown in the screenshot below.



8. You can select the desired thumbnail from the options and edit it further in the **Artboard**, as shown in the screenshot below. You can edit, resize, or add layers as required.



9. You can download the desired thumbnail using the **Export** option available in the top-right corner of the screen, as shown in the screenshot below.



10. Now, you must try generating three thumbnails using Simplified AI Thumbnail Maker.

Summary:

Congratulations! You just completed the hands-on lab Image Generation in Action.

In this lab, you explored image generation capabilities of Generative AI. You learned how to create realistic images from text using Stable Diffusion XL. You learned to generate captions for images using GPT-2-powered Hugging Face's image-to-text. You also learned how to create small thumbnail images using Simplified AI Thumbnail Maker.

© IBM Corporation 2023. All rights reserved.



Skills Network