
Generative AI for Video Generation (1)

COMP4431 Lab 4

Lab 04 Agenda

- Video Generation Models
 - Introductions of Video Generation Models
 - Products of Video Generation Models
 - Access to Video Generation Models
- Basic Video Generation Techniques
 - Text-to-video Generation
 - Image-to-video Generation

Introductions of Video Generation Models

■ Concept

- Video generation models deal with the creation of synthetic videos.

- New video clips are generated by describing the content of the desired video footage (text-to-video)
- or applying the style and composition via text or image prompt to a source video (video-to-video).

■ History

- Early works are limited to simple scenes (low-resolution, single-object, and short-duration).
- Benefiting from the success achieved by the diffusion model in the generative area, current works are generating more complex videos.

Duke Text2Filter
Swimming in swimming pool



TGANs-C
a cook puts noodles into some boiling water.



NUWA
running on the sea



S. stability ai An exploding cheese house



Pika LED chinese dragon hovering over chinese town



Morph Studio male wizards in wizard hats and wizard cloaks casting magic spell at McDonalds



runway The late afternoon sun peeking through the window of a New York City loft.



Introductions of Video Generation Models

■ Video Generation Models in 2024

□ OpenAI's Sora

- On February 15th, 2024, OpenAI introduces Sora, a vision foundation model that generates videos from text prompts.
- Sora can generate videos up to a minute long while maintaining visual quality and adherence.

□ Meta's Meta Movie Gen

- On October 4th, 2024, Meta introduces Meta Movie Gen, another vision foundation model for video generation.
- Meta Movie Gen generates high-quality HD videos of up to 16 seconds duration that follow the text prompt.

□ Is Sora/ Meta Movie Gen available right now to me?

■ A: No, it is **not yet widely available**.



<https://openai.com/index/sora/>



<https://ai.meta.com/research/movie-gen/>

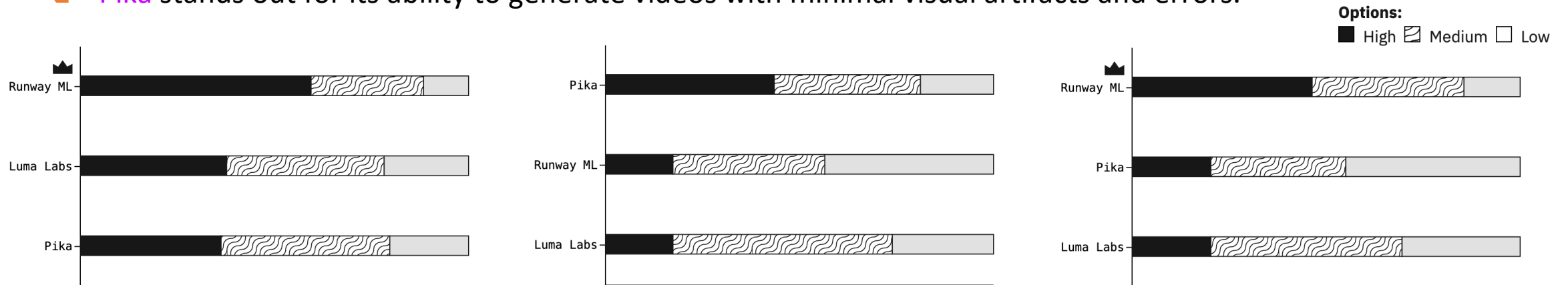
Products of Video Generation Models

■ Available Generative Artificial Intelligence Products for Video Generation

Company	Generation Type	Max Length	Extend?	(zoom, pan)	(amount)	Other Features	Format
Runway	Text-to-video, image-to-video, video-to-video	4 sec	Yes	Yes	Yes	Motion brush, upscale	Website
Pika	Text-to-video, image-to-video	3 sec	Yes	Yes	Yes	Modify region, expand canvas, upscale	Website
Genmo	Text-to-video, image-to-video	6 sec	No	Yes	Yes	FX presets	Website
Kaiber	Text-to-video, image-to-video, video-to-video	16 sec	No	No	No	Sync to music	Website
Stability	Image-to-video	4 sec	No	No	Yes		Local model, SDK
Zeroscope	Text-to-video	3 sec	No	No	No		Local model
ModelScope	Text-to-video	3 sec	No	No	No		Local model
AnimateDiff	Text-to-video, image-to-video, video-to-video	3 sec	No	No	No		Local model
Morph	Text-to-video	3 sec	No	No	No		Discord bot
Hotshot	Text-to-video	2 sec	No	No	No		Website
Moonvalley	Text-to-video, image-to-video	3 sec	No	Yes	No		Discord bot
Deforum	Text-to-video	14 sec	No	Yes	No	FX presets	Discord bot
Leonardo	Image-to-video	4 sec	No	No	Yes		Website
Assistive	Text-to-video, image-to-video	4 sec	No	No	Yes		Website
Neural Frames	Text-to-video, image-to-video, video-to-video	Unlimited	No	No	No	Sync to music	Website
MagicHour	Text-to-video, image-to-video, video-to-video	Unlimited	No	No	No	Face swap, sync to music	Website
Vispunk	Text-to-video	3 sec	No	Yes	No		Website
Decohere	Text-to-video, image-to-video	4 sec	No	No	Yes		Website
Domo AI	Image-to-video, video-to-video	3 sec	No	No	Yes		Discord bot
FullJourney	Text-to-video, image-to-video	8 sec	No	Yes	No	Lipsyncing, face swap	Discord bot

Products of Video Generation Models

- **Video generation leaderboard** evaluates AI models on their ability to generate high-quality videos from textual descriptions.
 - **Runway ML** excels in producing videos that closely match the given prompts and appear highly realistic.
 - **Pika** stands out for its ability to generate videos with minimal visual artifacts and errors.



- **Consistency with prompt**
 - assesses the model's ability to generate responses that are relevant, on-topic, and faithful to the context and requirements specified in the prompt.

- **Artifacts and errors**
 - includes identifying issues like flickering, color distortions, or other anomalies that detract from the visual quality and realism of the video output.

- **Is realistic**
 - measures whether the video looks like genuine footage, with realistic movements, lighting, and details, as opposed to appearing artificial.

Access to Video Generation Models (1)

- Runway ML
 - Navigate to <https://app.runwayml.com/>
 - Log in using your account and password
 - You can create account with different email address.
 - You can login with Google account.
 - You can login with Apple account.
 - Select a plan
 - Free Plan (\$0): Can't remove watermarks or upscale resolution in Gen-1 and Gen-2; 125 one-time credits (125 credits = 25s of Gen-2, 8s of Gen-1).
 - Standard Plan (\$12): Remove watermarks from video generations. 625 credits/month (625 credits = 125s of Gen-2, 44s of Gen-1)
 - Pro Plan (\$28): 2250 credits/month (2250 credits = 450s of Gen-2; 160s of Gen-1)
 - Prompt to create a video

Create an account

Already have an account? [Log in](#)

Next

OR



Sign up with Google



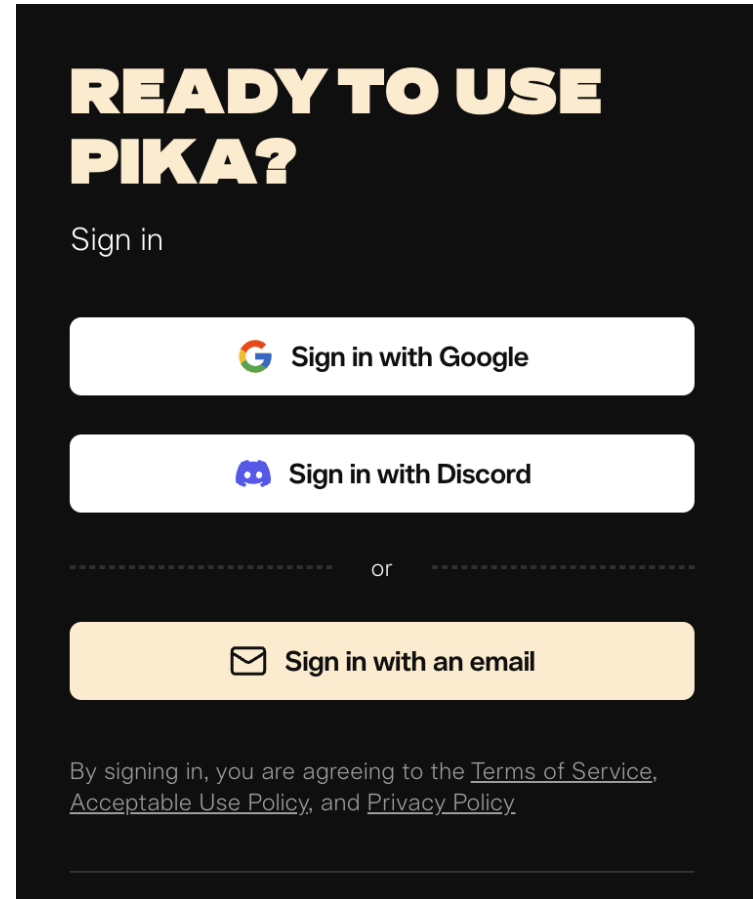
Sign up with Apple

Use Single Sign-On (SSO)

By clicking "Next", "Sign up with Google" or "Sign up with Apple" you agree to our [Terms of Use](#) and acknowledge that you have read and understand our [Privacy Policy](#).


Access to Video Generation Models (2)


- Pika
 - Navigate to <https://pika.art/my-library>
 - Log in using your account and password
 - You can create account with different email address.
 - You can login with Google account.
 - You can login with Discord account.
 - Select a plan
 - Free Plan (\$0): 150 monthly video credits
 - Standard Plan (\$10): 700 monthly video credits
 - Pro Plan (\$35): 2000 monthly video credits
 - Unlimited Plan (\$95): Unlimited monthly video credits
 - Prompt to create a video

A screenshot of the Pika login interface. At the top, it says "READY TO USE PIKA?" in bold yellow text. Below that, it says "Sign in". There are three main login buttons: "Sign in with Google" (with the Google logo), "Sign in with Discord" (with the Discord logo), and "Sign in with an email" (with an envelope icon). The "Sign in with an email" button is highlighted in orange. Below the buttons, there is a line of text: "By signing in, you are agreeing to the Terms of Service, Acceptable Use Policy, and Privacy Policy." with links to each of these documents.


READY TO USE PIKA?

Sign in

 Sign in with Google

 Sign in with Discord

or

 Sign in with an email

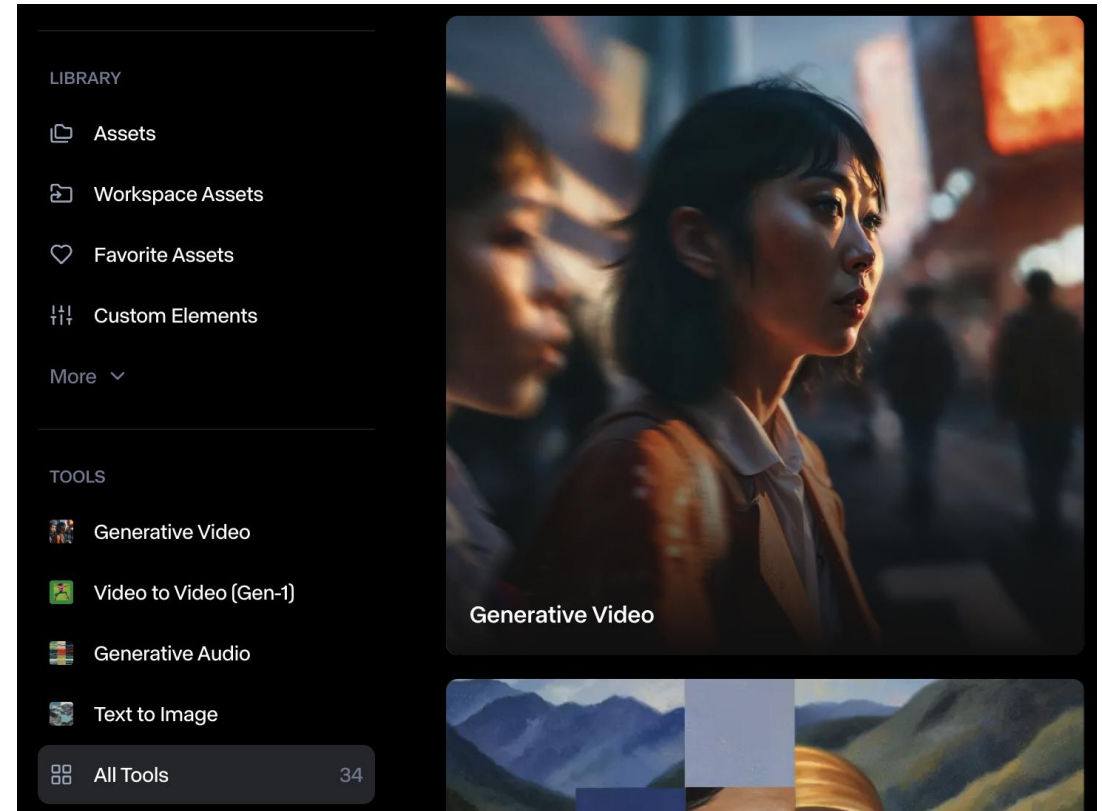
By signing in, you are agreeing to the [Terms of Service](#), [Acceptable Use Policy](#), and [Privacy Policy](#).

Lab 04 Agenda

- Video Diffusion Models
 - Concepts of Video Diffusion Models
 - Access to Video Diffusion Models
 - Leaderboard of Video Generation Mode
- Basic Video Generation Techniques
 - Text-to-video Generation
 - Image-to-video Generation

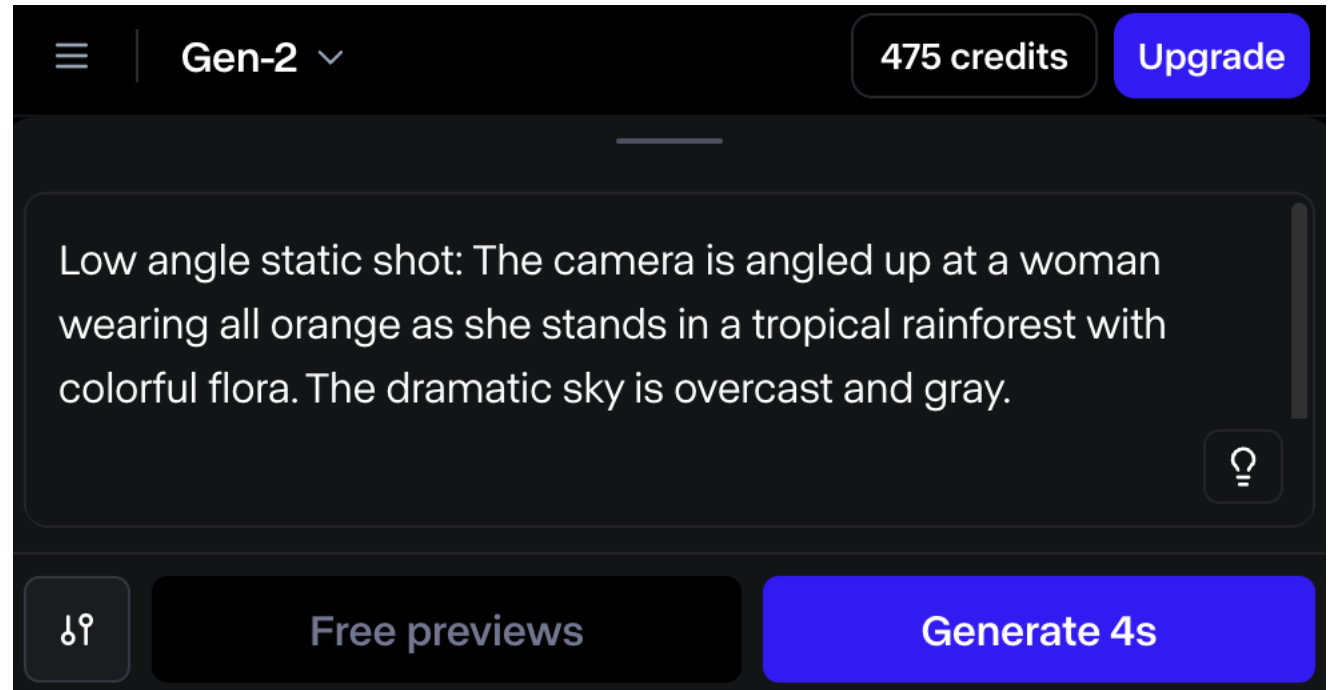
Text-to-video Generation

- A step-by-step guide to Text-to-Video Generation:
 - Select "Generative Video" from your dashboard.
 - Choose "Gen-2" from the model dropdown menu.
 - Enter a text prompt describing the video you want to create.
 - Click on the "Style" icon to add a style to your prompt.
 - Select an aspect ratio for your video.
 - Adjust the General Motion slider, which ranges from 1 to 10:
 - 1 represents little motion, while 10 indicates extreme motion.
 - Click "Generate" to create a video based on your settings and prompt.



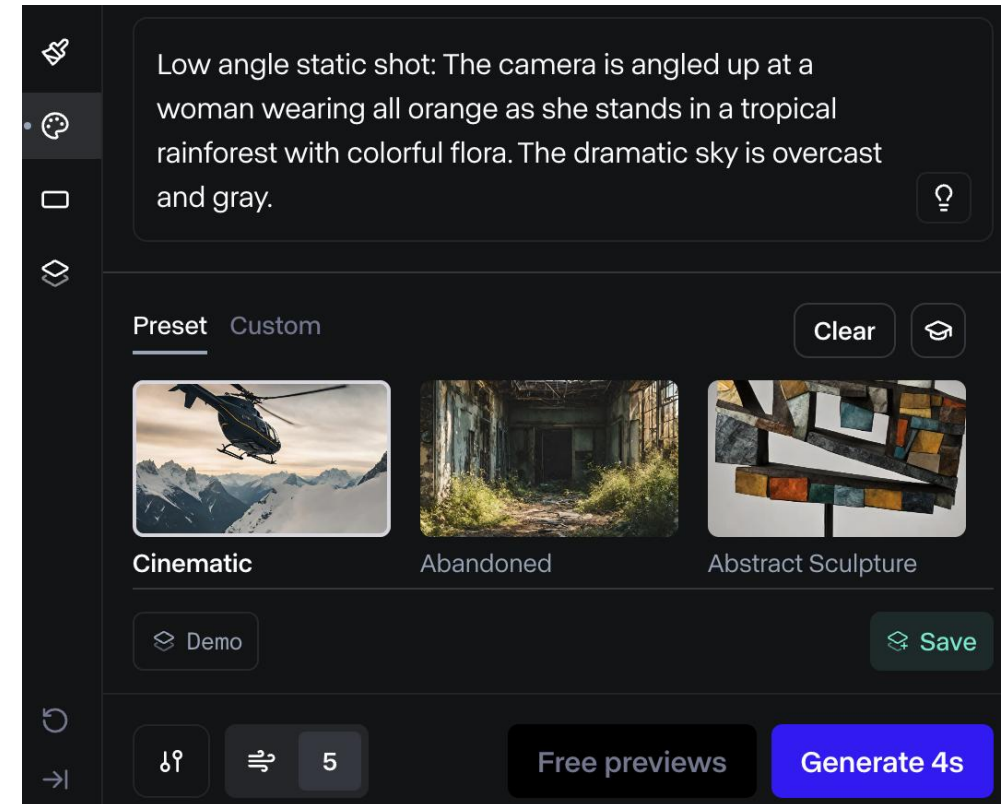
Text-to-video Generation

- A step-by-step guide to Text-to-Video Generation:
 - Select "Generative Video" from your dashboard.
 - Choose "Gen-2" from the model dropdown menu.
 - Enter a text prompt describing the video you want to create.
 - Low angle static shot: The camera is angled up at a woman wearing all orange as she stands in a tropical rainforest with colorful flora. The dramatic sky is overcast and gray.
 - Click on the "Style" icon to add a style to your prompt.
 - Select an aspect ratio for your video.
 - Adjust the General Motion slider, which ranges from 1 to 10:
 - 1 represents little motion, while 10 indicates extreme motion.
 - Click "Generate" to create a video based on your settings and prompt.



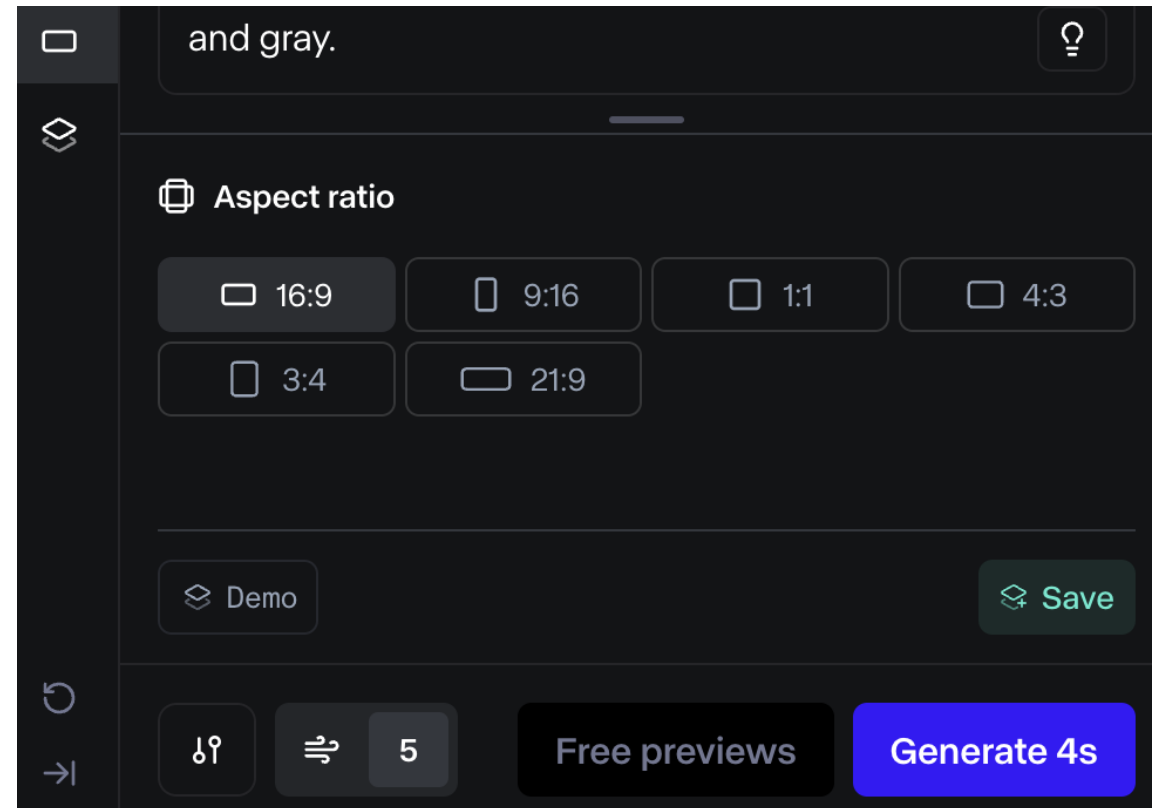
Text-to-video Generation

- A step-by-step guide to Text-to-Video Generation:
 - Select "Generative Video" from your dashboard.
 - Choose "Gen-2" from the model dropdown menu.
 - Enter a text prompt describing the video you want to create.
 - Click on the "Style" icon to add a style to your prompt.
 - Select an aspect ratio for your video.
 - Adjust the General Motion slider, which ranges from 1 to 10:
 - 1 represents little motion, while 10 indicates extreme motion.
 - Click "Generate" to create a video based on your settings and prompt.



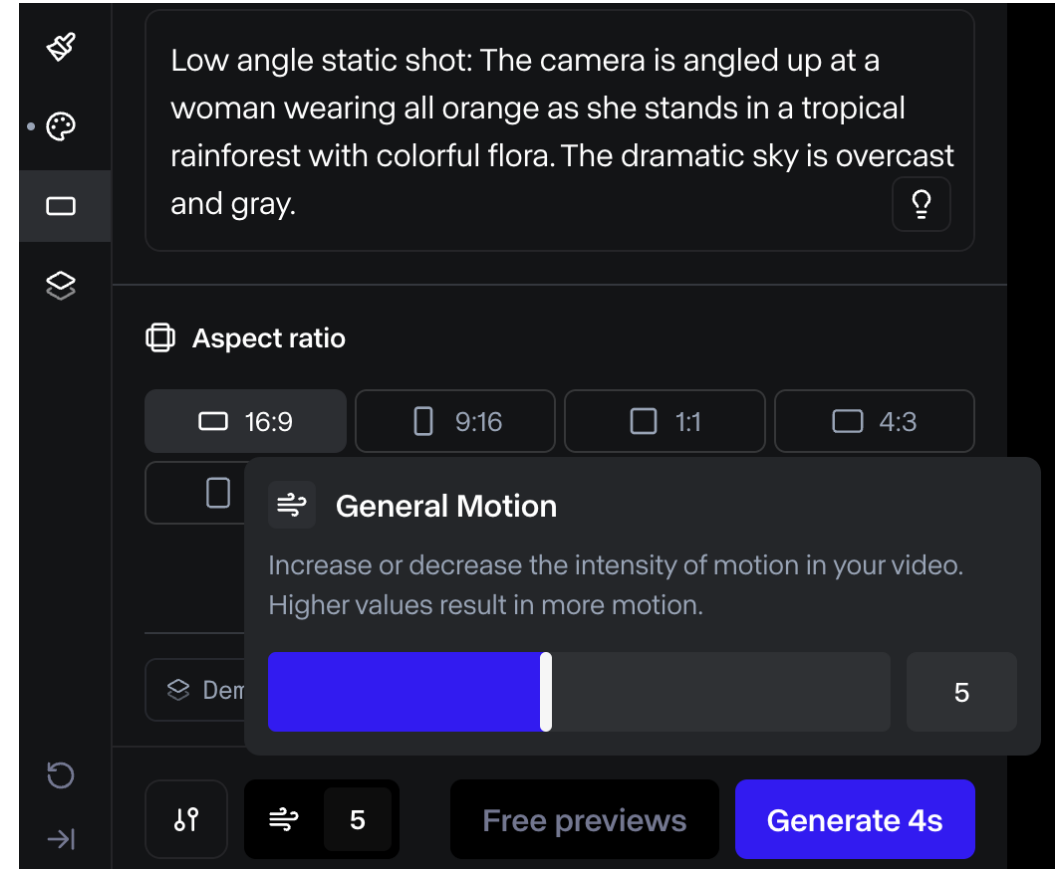
Text-to-video Generation

- A step-by-step guide to Text-to-Video Generation:
 - Select "Generative Video" from your dashboard.
 - Choose "Gen-2" from the model dropdown menu.
 - Enter a text prompt describing the video you want to create.
 - Click on the "Style" icon to add a style to your prompt.
 - **Select an aspect ratio for your video.**
 - Adjust the General Motion slider, which ranges from 1 to 10:
 - 1 represents little motion, while 10 indicates extreme motion.
 - Click "Generate" to create a video based on your settings and prompt.



Text-to-video Generation

- A step-by-step guide to Text-to-Video Generation:
 - Select "Generative Video" from your dashboard.
 - Choose "Gen-2" from the model dropdown menu.
 - Enter a text prompt describing the video you want to create.
 - Click on the "Style" icon to add a style to your prompt.
 - Select an aspect ratio for your video.
 - Adjust the General Motion slider, which ranges from 1 to 10:
 - 1 represents little motion, while 10 indicates extreme motion.
 - Click "Generate" to create a video based on your settings and prompt.



Text-to-video Generation

- What makes a good text prompt

- A well-structured prompt typically follows this format: [camera movement]: [establishing scene]. [additional details].
 - Example: Top-down-drone-shot: a colossal iceberg in Antarctica collapses into the frigid sea. White and blue hues contrast under the piercing sunlight, creating a mesmerizing visual spectacle.

- How to generate a good text prompt

- Navigate to the Runway Prompt Builder at <https://word.studio/tool/runway-prompt-builder/>
- Set camera details: Low angle, High angle, Overhead, FPV (First Person View), etc. (refer to the following link)
- Define the subject and location of your scene.
- Describe the action or movement: Dynamic motion, Slow motion, Fast motion, etc. (refer to the following link)
- Select a style for your video.
- Choose the desired prompt length.
- Click "Generate Prompt" to create your optimized text-to-video prompt.

The screenshot shows the Runway Prompt Builder interface. It has a light purple background. At the top, there's a section titled "Camera Details: Specify the camera angle, shot type, or camera movement" with a text input field containing "Top-down-drone-shot". Below this, there are two columns: "Subject" with a text input field containing "icebergs", and "Location" with a text input field containing "Antarctica". Underneath these is a section titled "Action or movement" with a text input field containing "Iceberg falls". At the bottom, there are two dropdown menus: "Style" with "Cinematic (live action)" selected, and "Prompt Length" with "20-30 words" selected. A large dark grey button labeled "Generate Prompt" is at the very bottom.

Lab 04 Agenda

- Video Diffusion Models
 - Introductions of Video Generation Models
 - Products of Video Generation Models
 - Access to Video Generation Models
- Basic Video Generation Techniques
 - Text-to-video Generation
 - Image-to-video Generation

Image-to-video Generation

- A Step-by-Step Guide to Image-to-Video Generation:
 - Refer to Lab 2 Generative Artificial Intelligence for Image Generation to create your starting image.
 - Navigate to <https://genai.polyu.edu.hk/> or <https://fal.ai/models>
 - Select a diffusion model
 - Prompting the diffusion mode
 - A beautiful and powerful mysterious sorceress, smile, sitting on a rock, lightning magic, hat, detailed leather clothing with gemstones, dress, castle background, digital art, hyper-realistic, fantasy, dark art, art-station, highly detailed, sharp focus, sci-fi, dystopian, iridescent gold.
 - Click on "Generative Video" and upload your generated image.
 - Provide additional text prompts alongside your image to better guide the video result.
 - Describe new elements to appear throughout the video. For example: "New lightning appears."
 - Describe any movements you want to occur during the video. For example: "The sorceress smiles and waves her hands."
 - Click the "Generate" button. The model will interpret your image and prompts to produce the best possible video results.

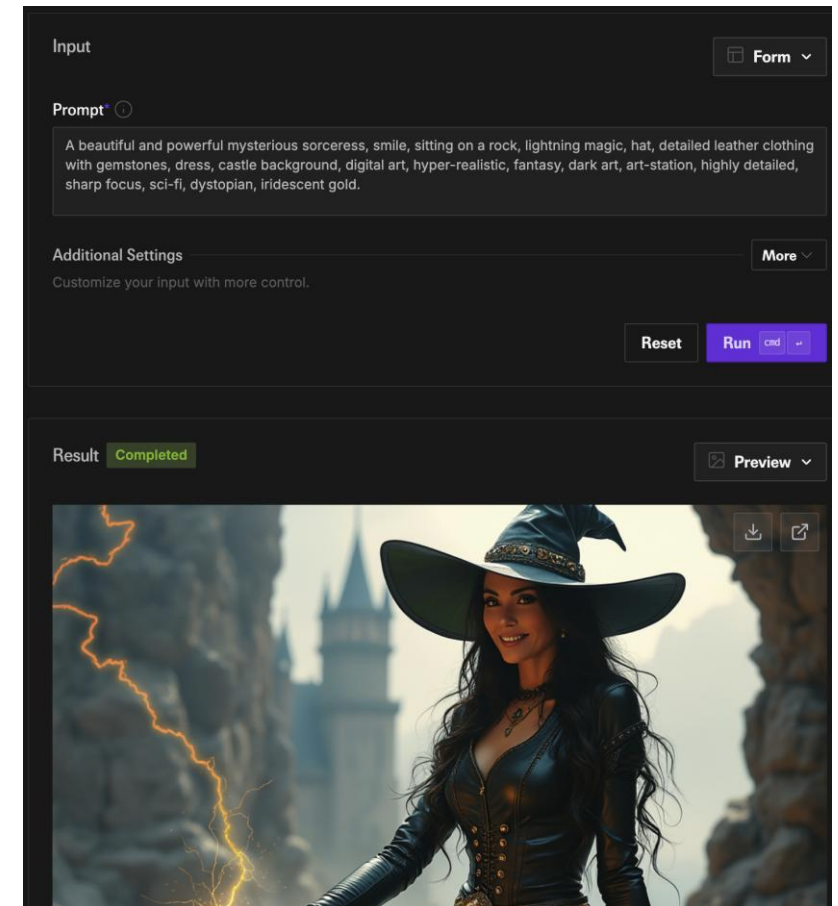


Image-to-video Generation

- A Step-by-Step Guide to Image-to-Video Generation:
 - Refer to Lab 2 Generative Artificial Intelligence for Image Generation to create your starting image.
 - Navigate to <https://genai.polyu.edu.hk/> or <https://fal.ai/models>
 - Select a diffusion model
 - Prompting the diffusion mode
 - A beautiful and powerful mysterious sorceress, smile, sitting on a rock, lightning magic, hat, detailed leather clothing with gemstones, dress, castle background, digital art, hyper-realistic, fantasy, dark art, art-station, highly detailed, sharp focus, sci-fi, dystopian, iridescent gold.
 - Click on "Generative Video" and upload your generated image.
 - Provide additional text prompts alongside your image to better guide the video result.
 - Describe new elements to appear throughout the video. For example: "New lightning appears."
 - Describe any movements you want to occur during the video. For example: "The sorceress smiles and waves her hands."
 - Click the "Generate" button. The model will interpret your image and prompts to produce the best possible video results.

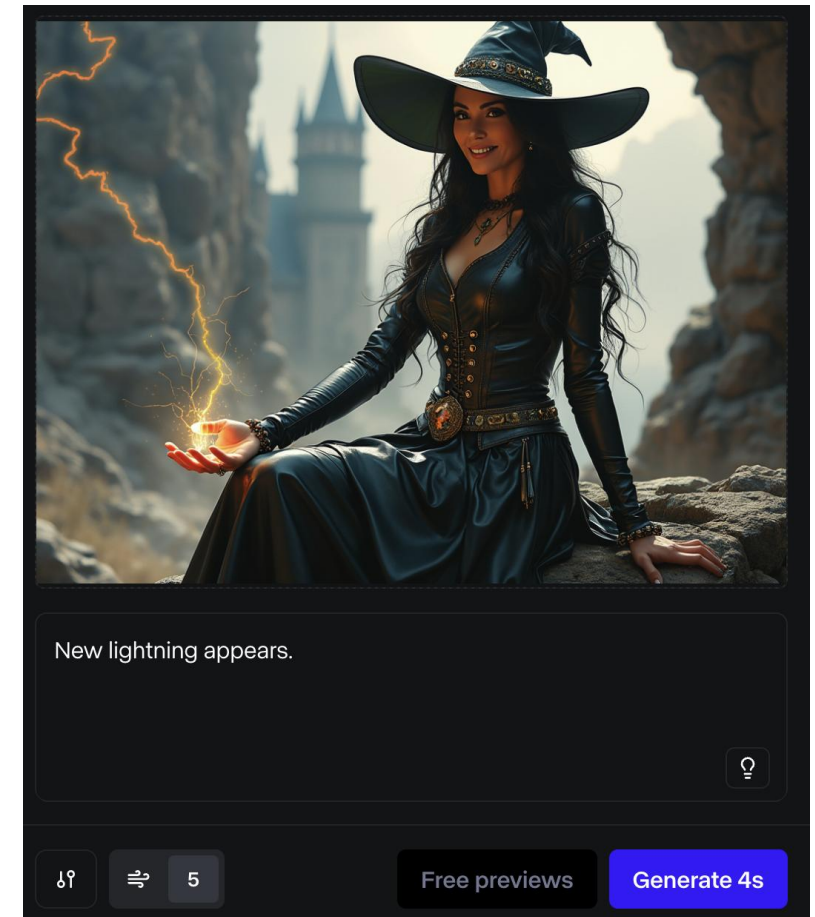


Image-to-video Generation

- A Step-by-Step Guide to Image-to-Video Generation:
 - Refer to Lab 2 Generative Artificial Intelligence for Image Generation to create your starting image.
 - Navigate to <https://genai.polyu.edu.hk/> or <https://fal.ai/models>
 - Select a diffusion model
 - Prompting the diffusion mode
 - A beautiful and powerful mysterious sorceress, smile, sitting on a rock, lightning magic, hat, detailed leather clothing with gemstones, dress, castle background, digital art, hyper-realistic, fantasy, dark art, art-station, highly detailed, sharp focus, sci-fi, dystopian, iridescent gold.
 - Click on "Generative Video" and upload your generated image.
 - Provide additional text prompts alongside your image to better guide the video result.
 - Describe new elements to appear throughout the video. For example: "New lightning appears."
 - Describe any movements you want to occur during the video. For example: "The sorceress smiles and waves her hands."
 - Click the "Generate" button. The model will interpret your image and prompts to produce the best possible video results.

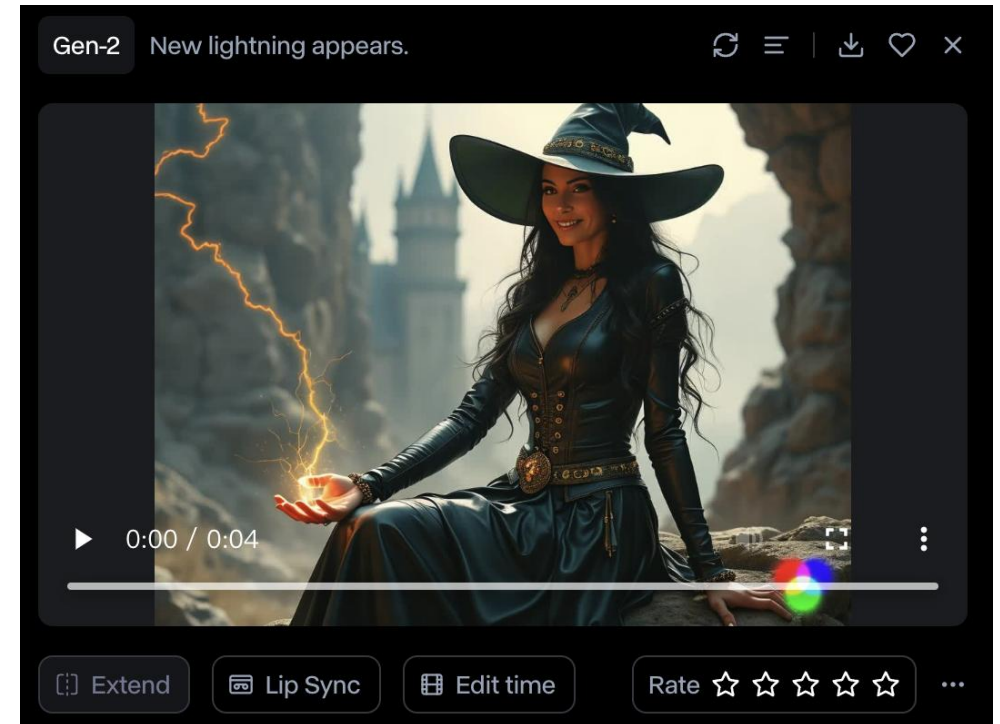


Image-to-video Generation

- A Step-by-Step Guide to Fine-grained Motion Control
 - Select "Generative Video" from your dashboard.
 - Select "Gen-2" from the model dropdown menu. Upload or select an image to use as the starting point for your video.
 - Click on the "Motion Brush" tool, and paint over the areas of the image where you want to control motion.
 - Each brush stroke represents a distinct motion area.
 - Painting over an existing stroke replaces it on the canvas, allowing for refinement of motion areas.
 - Control the speed and direction of motion for each brush stroke.
 - Adjust Horizontal movement
 - Adjust Vertical movement
 - Adjust Proximity (depth) movement
 - Click "Generate" to generate a video, and click on "prompt" to view the generated results.

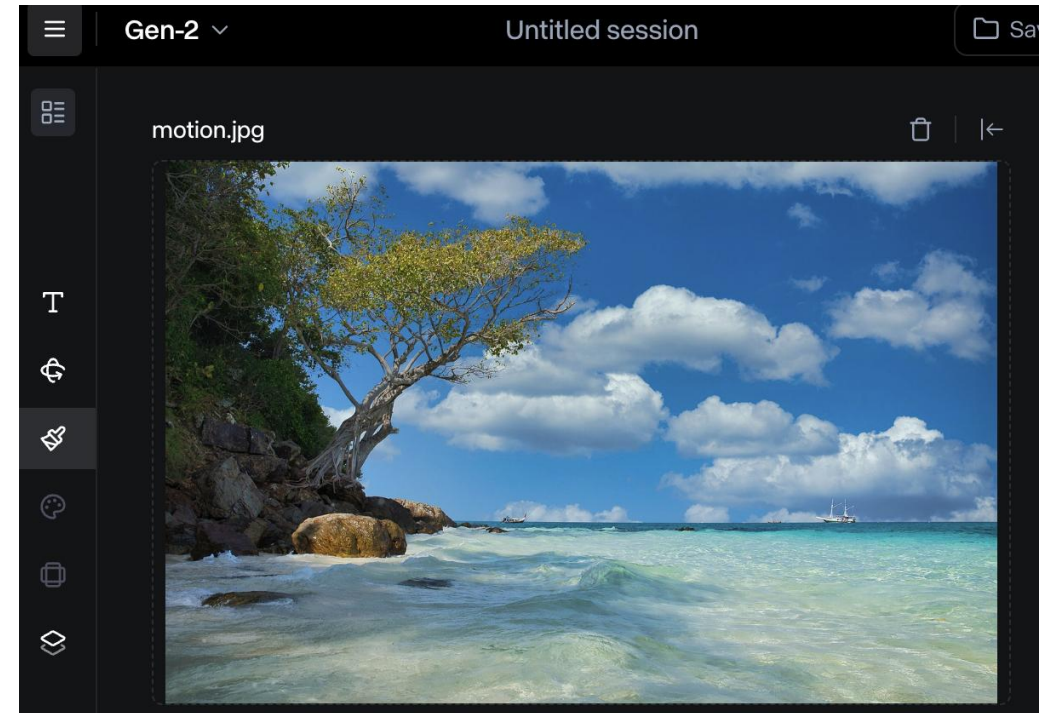


Image-to-video Generation

- A Step-by-Step Guide to Fine-grained Motion Control
 - Select "Generative Video" from your dashboard.
 - Select "Gen-2" from the model dropdown menu. Upload or select an image to use as the starting point for your video.
 - Click on the "Motion Brush" tool, and paint over the areas of the image where you want to control motion.
 - Each brush stroke represents a distinct motion area.
 - Painting over an existing stroke replaces it on the canvas, allowing for refinement of motion areas.
 - Control the speed and direction of motion for each brush stroke.
 - Adjust Horizontal movement
 - Adjust Vertical movement
 - Adjust Proximity (depth) movement
 - Click "Generate" to generate a video, and click on "prompt" to view the generated results.

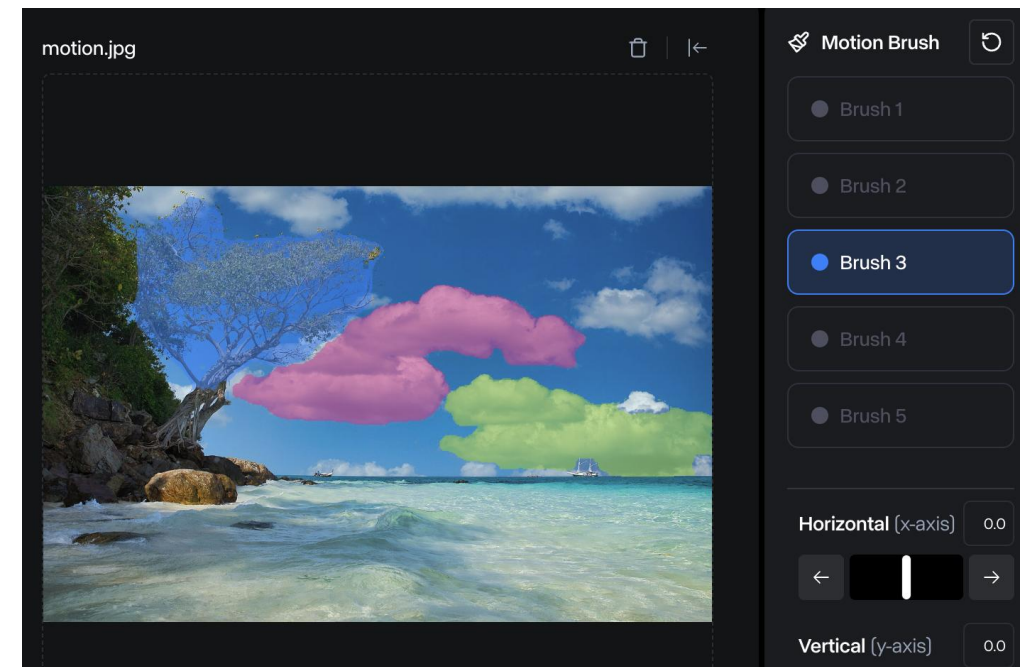


Image-to-video Generation

- A Step-by-Step Guide to Fine-grained Motion Control

- Select "Generative Video" from your dashboard.
- Select "Gen-2" from the model dropdown menu. Upload or select an image to use as the starting point for your video.
- Click on the "Motion Brush" tool, and paint over the areas of the image where you want to control motion.
 - Each brush stroke represents a distinct motion area.
 - Painting over an existing stroke replaces it on the canvas, allowing for refinement of motion areas.
- Control the speed and direction of motion for each brush stroke.
 - Adjust Horizontal movement
 - Adjust Vertical movement
 - Adjust Proximity (depth) movement
- Click "Generate" to generate a video, and click on "prompt" to view the generated results.

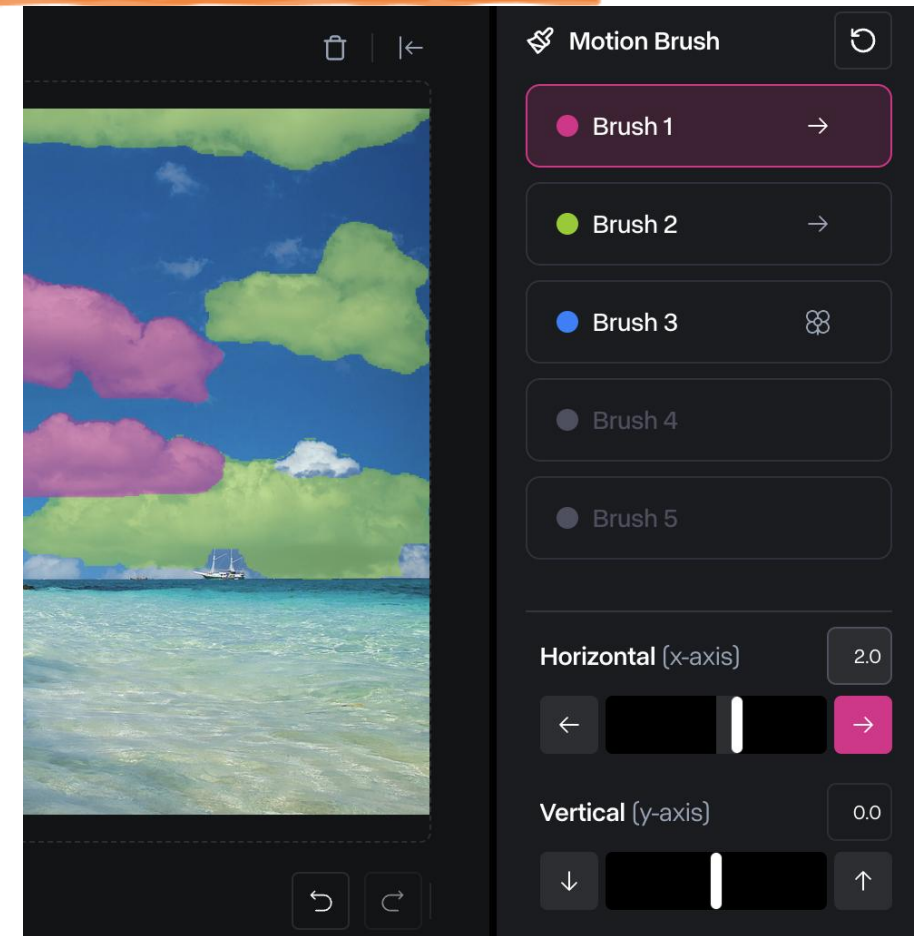


Image-to-video Generation

- A Step-by-Step Guide to Fine-grained Motion Control
 - Select "Generative Video" from your dashboard.
 - Select "Gen-2" from the model dropdown menu. Upload or select an image to use as the starting point for your video.
 - Click on the "Motion Brush" tool, and paint over the areas of the image where you want to control motion.
 - Each brush stroke represents a distinct motion area.
 - Painting over an existing stroke replaces it on the canvas, allowing for refinement of motion areas.
 - Control the speed and direction of motion for each brush stroke.
 - Adjust Horizontal movement
 - Adjust Vertical movement
 - Adjust Proximity (depth) movement
 - Click "Generate" to generate a video, and click on "prompt" to view the generated results.

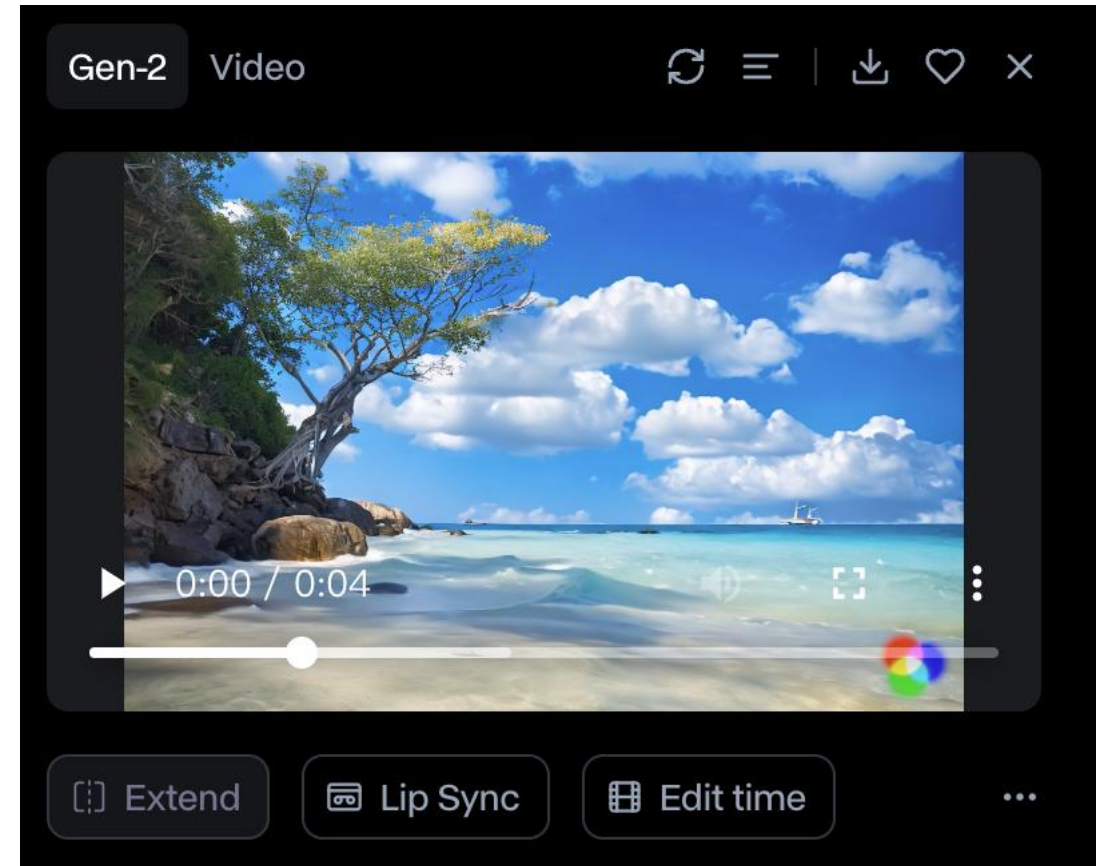


Image-to-video Generation

- A Step-by-Step Guide to Fine-grained Camera Control
 - Click "Camera control"
 - Choose one or multiple directions to move your camera
 - Horizontal, Vertical, Pan (left and right), Tilt (up and down), Zoom, and Roll
 - Adjust the "Speed" for camera motion.
- Click "Generate" to create a video. Click on "Prompt" to view the generated results.
 - This feature can be used in conjunction with the motion brush. When combined, the motion brush movement remains the same.

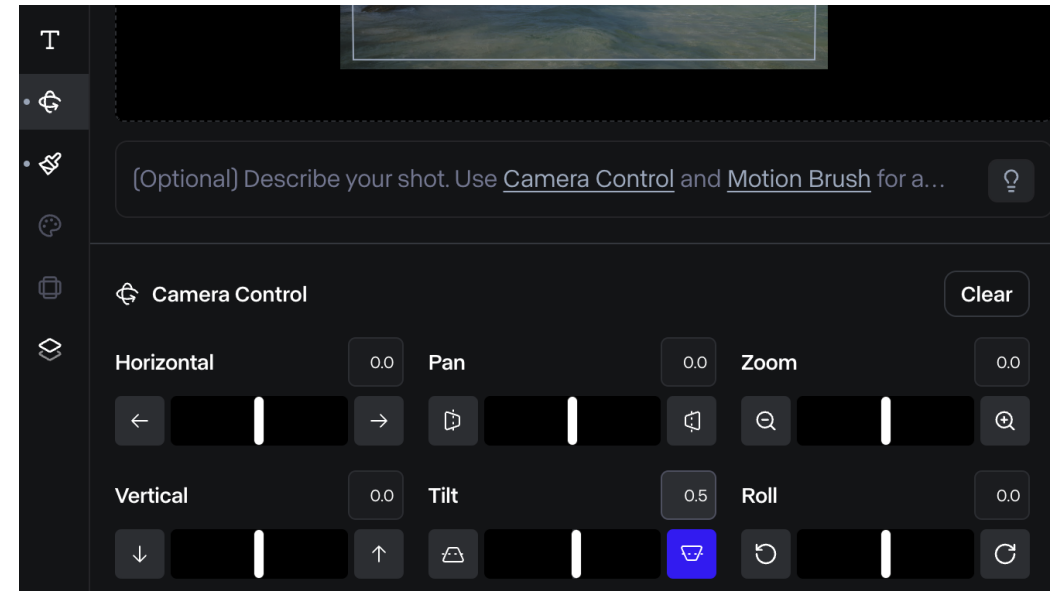
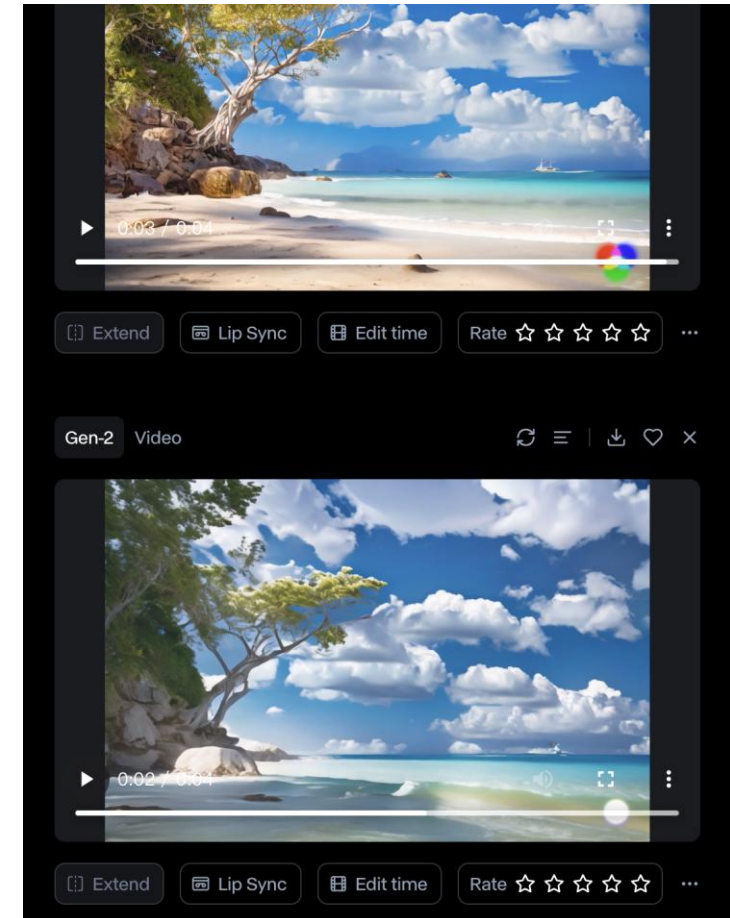


Image-to-video Generation

- A Step-by-Step Guide to Fine-grained Camera Control
 - Click "Camera control"
 - Choose one or multiple directions to move your camera
 - Horizontal, Vertical, Pan (left and right), Tilt (up and down), Zoom, and Roll
 - Adjust the "Speed" for camera motion.
 - Click "Generate" to create a video. Click on "Prompt" to view the generated results.
 - This feature can be used in conjunction with the motion brush. When combined, the motion brush movement remains the same.



Question Time!

Or See you next week!