

# Imágenes y videos infinitos con IA local

Crea una imagen. Dale movimiento.  
Todo en tu computadora.

## DE CERO A TU PRIMER ROBOT

Instalación paso a paso

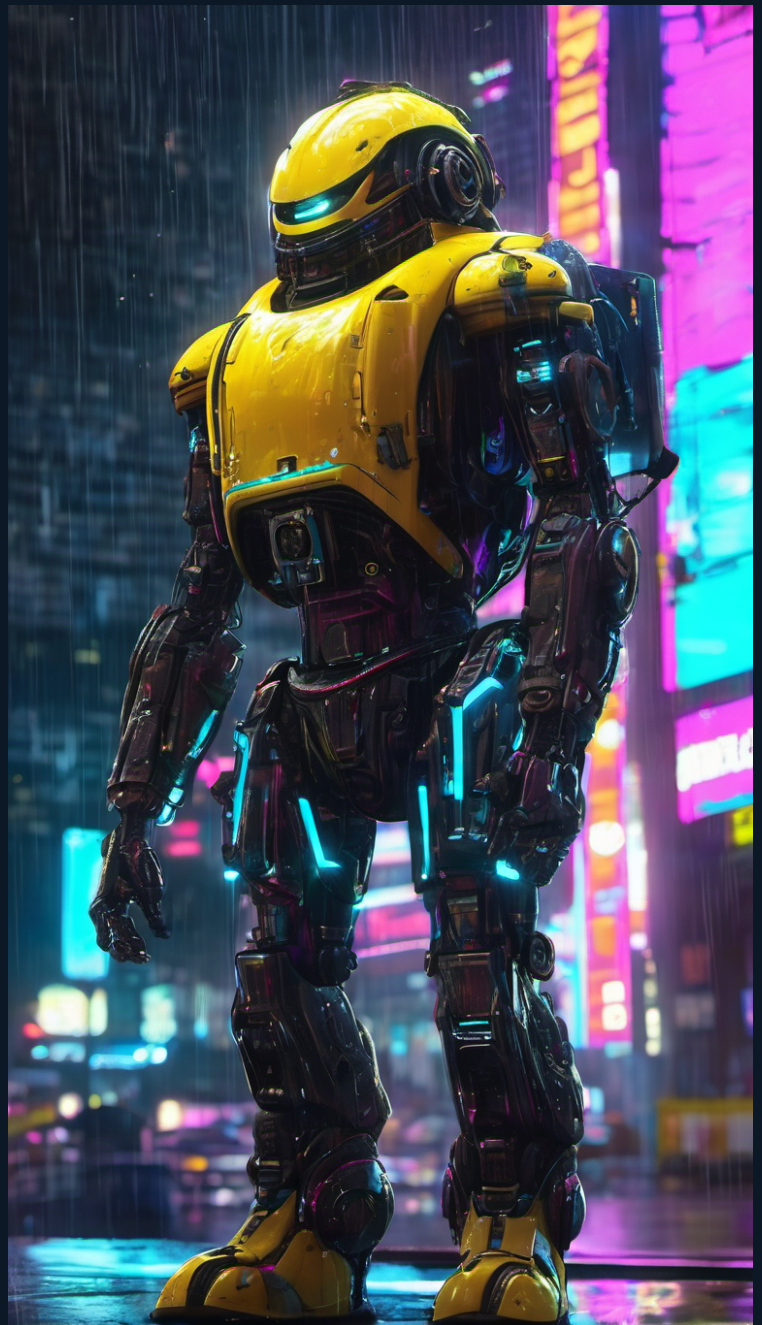
Dos workflows listos

Prompts para copiar

Ejemplos reales

## ComfyUI + SDXL + LTX-Video

Sin créditos por generación.  
El límite lo pone tu hardware,  
tu tiempo y tu almacenamiento.



## MI COMUNIDAD EN SKOOL

# Aprende Homelabs e IA local conmigo.

Tengo una comunidad en Skool para aprender sobre **Homelabs e IA local**.

Si quieres seguir explorando cómo montar tu propio laboratorio y ejecutar IA en tu computadora, te invito a conocerla.

## HOMELABS

Explora el mundo de los servidores y los laboratorios en casa.

## IA LOCAL

Descubre lo que puedes crear con inteligencia artificial en tu propio hardware.

## SIGUE APRENDIENDO CONMIGO

**CONOCE MI COMUNIDAD EN SKOOL**

[www.skool.com/homelabs-9089](https://www.skool.com/homelabs-9089)

La guía paso a paso comienza en la siguiente página.

ANTES DE EMPEZAR

# Tu computadora es el estudio.

Con este kit pasarás de una descripción a una imagen y de esa imagen a un clip corto.

## Qué hace cada pieza

**ComfyUI** conecta y ejecuta los pasos. **SDXL** genera la imagen. **LTX-Video** genera el movimiento. **T5** interpreta el texto del video.

## Qué necesitas

- **Mac:** Apple Silicon M1 o posterior, macOS 13+.
- **Windows:** Windows 10+; una GPU dedicada compatible es recomendable.
- Internet para instalar y descargar. Después, estos flujos pueden generar sin conexión.
- Unos **23 GB** para los modelos. Recomendación práctica: **60 GB libres** para instalación, descargas y resultados.

## La referencia de esta guía

La demo se probó en un **Mac M4 Pro con 48 GB de memoria unificada**. Que ComfyUI abra en tu equipo no significa que este video vaya a caber en memoria. Si tienes menos recursos, empieza por la imagen y una prueba de video pequeña.

## Tu ruta

|                             |               |
|-----------------------------|---------------|
| Instalar y preparar         | Páginas 4-6   |
| Crear la imagen             | Páginas 7-8   |
| Animar la imagen            | Páginas 9-11  |
| Prompts, resultados y ayuda | Páginas 12-16 |

Requisitos de la app: [macOS](#) · [Windows](#). Los 60 GB son una recomendación de trabajo, no un mínimo oficial.

## 01 / INSTALACIÓN

# Instala ComfyUI en local.

La ruta más sencilla es Comfy Desktop. Prepara Python y las dependencias por ti.

## 1 Descarga desde la página oficial

[Abrir comfy.org/download](https://comfy.org/download). Elige el instalador de tu sistema.

## 2 Instala la aplicación

**Mac:** abre el .dmg y arrastra Comfy Desktop a Aplicaciones.

**Windows:** abre el .exe y sigue el asistente. Luego inicia la app.

## 3 Selecciona Local

En el primer inicio, elige **Local**. Revisa los términos y decide si los aceptas. El envío de estadísticas y las funciones beta son opcionales.

## 4 Completa la instalación

Puedes activar **Express Install** para usar los ajustes recomendados. Si eliges la configuración manual del asistente, selecciona la versión estable y una unidad con espacio.

## 5 Abre tu instancia

Espera a que termine y abre la instalación desde el panel de Comfy Desktop. Debes ver el lienzo de nodos y el botón **Run**.

### Si ya tenías ComfyUI

Puedes usar tu instalación actual y continuar en la página 5. Las capturas de esta guía provienen de una instalación local abierta en el navegador. Desktop usa el mismo motor de workflows.

# Descarga estos tres archivos.

Pulsa cada enlace. Conserva el nombre exacto y espera a que termine la descarga.

## 01 SDXL Base 1.0

sd\_xl\_base\_1.0.safetensors

Genera la imagen y contiene sus codificadores de texto y VAE.

**6,94 GB** · Carpeta: checkpoints/

[Descargar archivo oficial](#)

## 02 LTX-Video 2B 0.9.8 Distilled

ltxv-2b-0.9.8-distilled.safetensors

Genera los fotogramas del video e incluye su VAE.

**6,34 GB** · Carpeta: checkpoints/

[Descargar archivo oficial](#)

## 03 T5-XXL FP16

t5xxl\_fp16.safetensors

Convierte la descripción del movimiento en información para LTX.

**9,79 GB** · Carpeta: text\_encoders/

[Descargar archivo oficial](#)

### Usa las versiones del kit

El video usa **LTX-Video 2B 0.9.8 Distilled**. LTX-2 y otros templates tienen requisitos distintos. No necesitas Qwen, Ollama, otro VAE ni un refiner para estos dos workflows.

Tamaños aproximados en GB decimales. Los enlaces y SHA256 también están en MODELOS\_Y\_ENLACES.txt.

## 03 / PREPARACIÓN

# Pon cada archivo en su lugar.

El directorio correcto es el que está configurado en tu instalación activa.

## 1 Localiza tu carpeta de modelos

En **Comfy Desktop**: menú superior > **Desktop Settings > Storage > Shared Models**.  
Pulsa la ruta para abrirla en Finder o el Explorador.

Dentro de ese directorio, usa estas subcarpetas:

```
checkpoints/  
sd_xl_base_1.0.safetensors  
ltxv-2b-0.9.8-distilled.safetensors  
text_encoders/  
t5xxl_fp16.safetensors
```

Si instalaste desde el código fuente, estas carpetas están en **ComfyUI/models/**. No añadas otra carpeta "models" dentro de una ruta que ya apunta a los modelos. Reinicia tu instancia después de copiarlos.

## 2 Descomprime el Kit\_LOCAL

En el ZIP encontrarás **workflows, ejemplos, recursos**, la guía, PROMPTS.txt y MODELOS\_Y\_ENLACES.txt.

## 3 Abre el workflow de imagen

Crea una pestaña nueva con +. Usa **Workflow > Open** o arrastra **01\_Robot\_Imagen\_SDXL.json** al lienzo. Guarda con otro nombre para conservar tu base.

### Comprueba antes de ejecutar

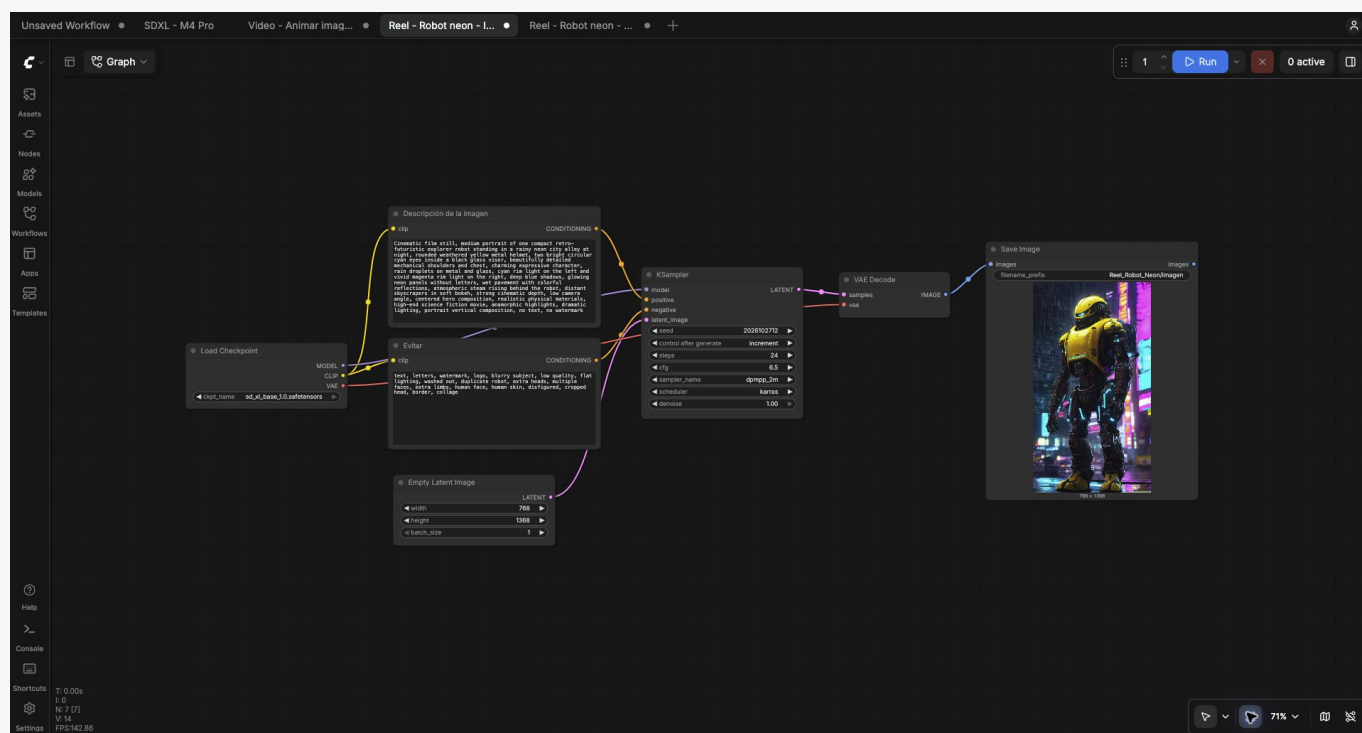
En Load Checkpoint debe aparecer **sd\_xl\_base\_1.0.safetensors**. El JSON contiene los nodos y ajustes; los modelos se descargan aparte. La carpeta del ZIP no sustituye la carpeta de modelos de ComfyUI.

Carpetas compartidas y rutas activas: [Documentación de Storage](#).

## 04 / WORKFLOW DE IMAGEN

# Así nace la imagen.

Abre 01\_Robot\_Imagen\_SDXL.json. Los cables indican qué información recibe cada paso.



Captura real del workflow SDXL. Puedes ampliar el PDF para ver los nodos.

## 1 Load Checkpoint

Carga SDXL Base 1.0: modelo, codificadores de texto y VAE.

## 2 Descripción + Evitar

El texto positivo describe la escena; el negativo orienta qué evitar.

## 3 Empty Latent Image + KSampler

Define el tamaño. El sampler transforma ruido en una representación de la imagen, guiado por el texto.

## 4 VAE Decode + Save Image

Convierte esa representación en píxeles y guarda un PNG.

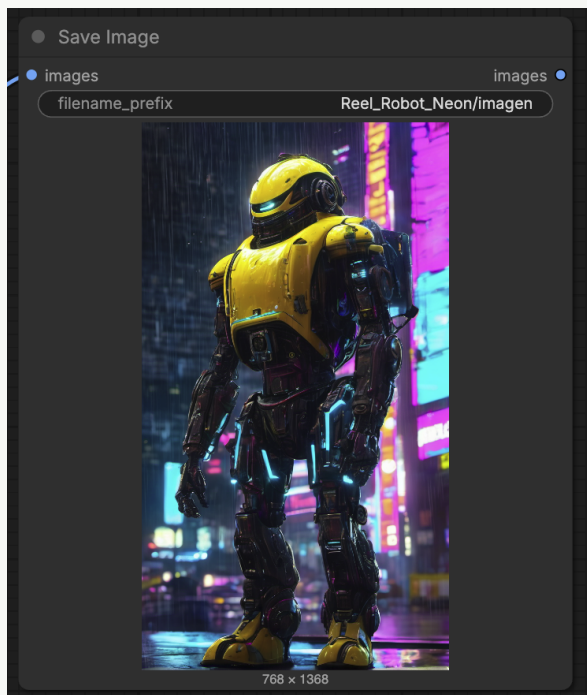
En SDXL, los componentes que leen el texto ya vienen dentro del checkpoint. No necesitas cargar T5 para este flujo.

## 05 / PRIMERA IMAGEN

# Genera tu robot.

El prompt ya viene escrito. Para la primera ejecución conserva los ajustes del kit.

|                     |                   |
|---------------------|-------------------|
| Resolución          | 768 × 1368        |
| Batch size          | 1                 |
| Steps / CFG         | 24 / 6.5          |
| Sampler / Scheduler | dpmpp_2m / karras |
| Denoise             | 1.0               |



## 1. Pulsa Run

Espera a que termine la cola. La primera ejecución puede tardar más al cargar los modelos.

## 2. Guarda el resultado

Save Image escribe el PNG en la carpeta de salida activa, dentro de **Reel\_Robot\_Neon**. También puedes descargarlo desde Assets.

## 3. Prepara el siguiente paso

Usa este PNG en el workflow de video. También puedes usar la imagen de ejemplo del ZIP.

## Para obtener otra versión

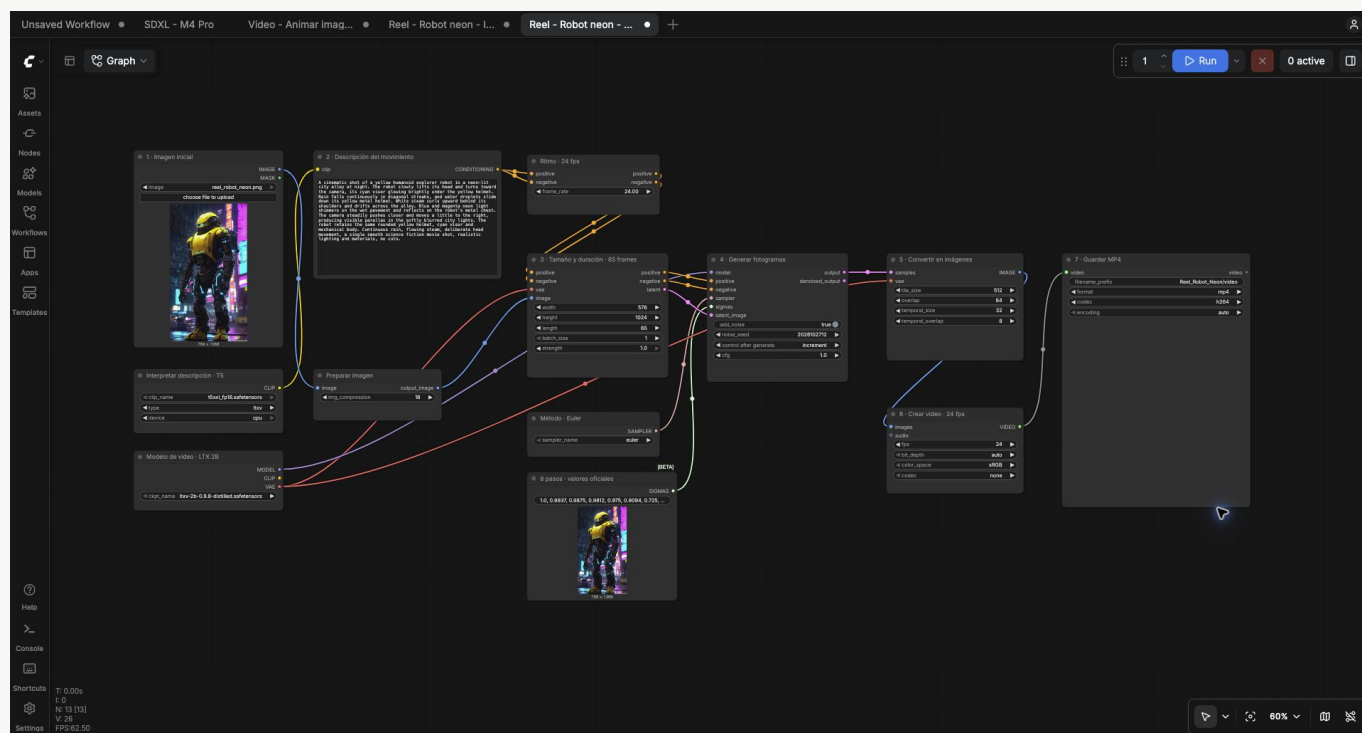
Cambia la descripción o la semilla (**seed**). El ajuste **increment** cambia la semilla después de generar. Una misma semilla ayuda a repetir pruebas, pero el resultado puede variar entre equipos y versiones.

Captura aportada por Harys. Imagen original incluida: ejemplos/reel\_robot\_neon.png.

## 06 / WORKFLOW DE VIDEO

# De una foto a una escena.

Crea otra pestaña y abre 02\_Robot\_Video\_LTX.json. Conserva el workflow de imagen por separado.



Vista completa del workflow de video. El detalle de entrada aparece en la página siguiente.

## 1 Imagen + descripción del movimiento

La imagen aporta la apariencia. T5 interpreta el texto que describe la acción y la cámara.

## 2 LTX-Video + sampler

LTX genera la secuencia guiándose por esa imagen y el texto. Los 8 pasos del sampler vienen configurados.

## 3 VAE Decode Tiled

Convierte la secuencia interna en imágenes visibles, procesando por bloques.

## 4 Create Video + Save Video

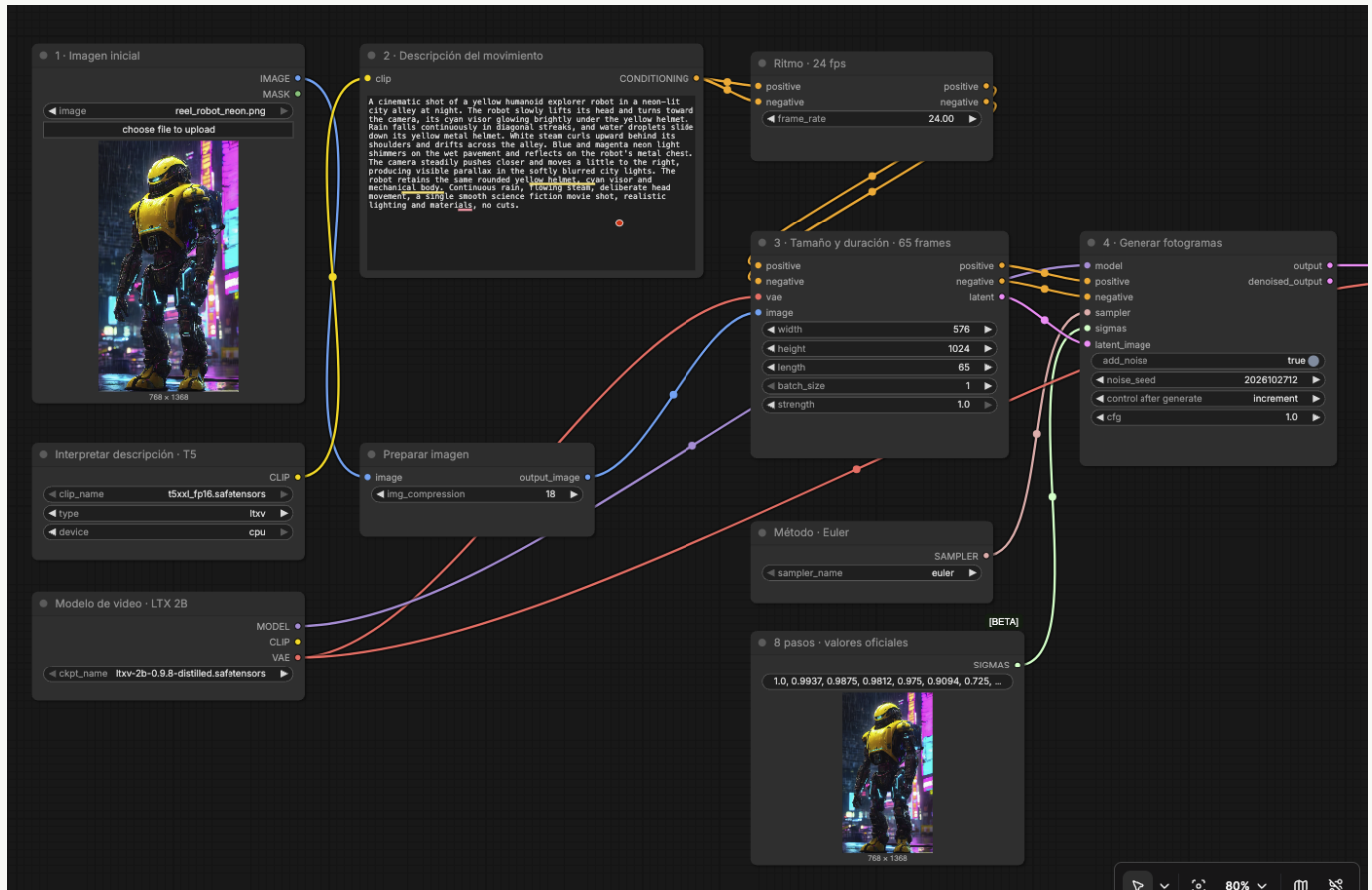
Une los fotogramas a 24 fps y guarda un MP4 sin audio.

Modelo utilizado: LTX-Video 2B 0.9.8 Distilled. El JSON incluye todos los cables y nodos nativos necesarios.

## 07 / ENTRADAS DEL VIDEO

# Carga la imagen y el movimiento.

Aquí controlas qué se anima y cómo quieres que se mueva.



Detalle real de la entrada, T5, LTX y los ajustes de generación.

## 1 Imagen inicial

Pulsa **choose file to upload**. Elige **ejemplos/reel\_robot\_neon.png** o el PNG que acabas de crear. Confirma que aparece la miniatura.

## 2 Descripción del movimiento

El prompt del kit pide lluvia, vapor, un giro suave del robot y un acercamiento de cámara. Puedes copiarlo desde PROMPTS.txt.

## 3 Revisa los dos cargadores

**Modelo de video:** `ltxv-2b-0.9.8-distilled.safetensors`.

**Interpretar descripción:** `t5xxl_fp16.safetensors`, tipo **ltxv**, dispositivo **cpu**.

T5 en CPU es la configuración usada en esta demo para reducir la carga de memoria en la GPU. En Mac, conserva FP16.

## 08 / GENERACIÓN

# Crea y guarda tu primer clip.

Estos son los valores usados para el robot del reel.

|                           |                             |
|---------------------------|-----------------------------|
| Tamaño                    | 576 × 1024, vertical 9:16   |
| Length / Batch size       | 65 fotogramas / 1           |
| Frame rate y Create Video | 24 fps en ambos nodos       |
| Sampler / CFG             | Euler / 1.0                 |
| Pasos                     | 8, con los sigmas incluidos |
| Salida                    | MP4, H.264, sin audio       |

## 1 Ejecuta el workflow

Pulsa **Run** una sola vez. Espera hasta que la cola termine. La carga de modelos puede ocupar buena parte de la primera ejecución.

## 2 Encuentra tu MP4

En **Assets**, abre el video y descárgalo. Save Video también lo guarda en **Reel\_Robot\_Neon**, dentro de tu carpeta de salida activa.

### ¿Cuánto dura?

$65 \text{ fotogramas} \div 24 \text{ fps} = \mathbf{2,71 \text{ s}}$ . Con 121 fotogramas: 5,04 s y mayor consumo. Usa longitudes  **$8n + 1$** : 33, 65, 97, 121...

### Una primera prueba más pequeña

Prueba **288 × 512 y 33 fotogramas**, con batch 1. Es una propuesta para reducir carga, no una garantía de rendimiento. Mantén los 8 pasos, CFG 1 y los demás ajustes del modelo.

Las dimensiones de este workflow LTX deben ser múltiplos de 32. El archivo de ejemplo ya usa los valores de la demo.

## 09 / PROMPTS PARA COPIAR

# La escena del robot.

Los textos completos también están en PROMPTS.txt, para copiarlos sin saltos del PDF.

## DESCRIPCIÓN DE LA IMAGEN

Cinematic film still, medium portrait of one compact retro-futuristic explorer robot standing in a rainy neon city alley at night, rounded weathered yellow metal helmet, two bright circular cyan eyes inside a black glass visor, beautifully detailed mechanical shoulders and chest, charming expressive character, rain droplets on metal and glass, cyan rim light on the left and vivid magenta rim light on the right, deep blue shadows, glowing neon panels without letters, wet pavement with colorful reflections, atmospheric steam rising behind the robot, distant skyscrapers in soft bokeh, strong cinematic depth, low camera angle, centered hero composition, realistic physical materials, high-end science fiction movie, anamorphic highlights, dramatic lighting, portrait vertical composition, no text, no watermark

## EVITAR / PROMPT NEGATIVO

text, letters, watermark, logo, blurry subject, low quality, flat lighting, washed out, duplicate robot, extra heads, multiple faces, extra limbs, human face, human skin, disfigured, cropped head, border, collage

### Cómo convertirlo en tu propia escena

Cambia el sujeto, el lugar, la luz y los materiales. Por ejemplo: reemplaza el robot amarillo por un explorador plateado, o la ciudad nocturna por un desierto al amanecer. Mantén una sola idea clara por imagen.

Prompt utilizado en la demo SDXL. La imagen resultante es una interpretación, no una reproducción literal de cada detalle del texto.

## 10 / PROMPT DE MOVIMIENTO

# Describe lo que sucede.

Usa frases sobre acciones visibles: sujeto, ambiente y cámara.

A cinematic shot of a yellow humanoid explorer robot in a neon-lit city alley at night. The robot slowly lifts its head and turns toward the camera, its cyan visor glowing brightly under the yellow helmet. Rain falls continuously in diagonal streaks, and water droplets slide down its yellow metal helmet. White steam curls upward behind its shoulders and drifts across the alley. Blue and magenta neon light shimmers on the wet pavement and reflects on the robot's metal chest. The camera steadily pushes closer and moves a little to the right, producing visible parallax in the softly blurred city lights. The robot retains the same rounded yellow helmet, cyan visor and mechanical body. Continuous rain, flowing steam, deliberate head movement, a single smooth science fiction movie shot, realistic lighting and materials, no cuts.

## Tres ingredientes que funcionan en esta escena

**Sujeto:** gira la cabeza suavemente.

**Ambiente:** caen gotas y se mueve el vapor.

**Cámara:** se acerca de forma lenta y continua.

## El ciclo para crear más

Cambia el prompt en SDXL > genera una imagen > súbela al workflow de video > describe un movimiento > genera el clip.

Empieza con cambios suaves. Movimientos extremos pueden deformar al personaje o alterar los detalles.

Este workflow produce un clip silencioso. La música, la voz y el montaje se añaden después en tu editor de video.

## 11 / RESULTADO REAL

# Esto salió de la prueba.

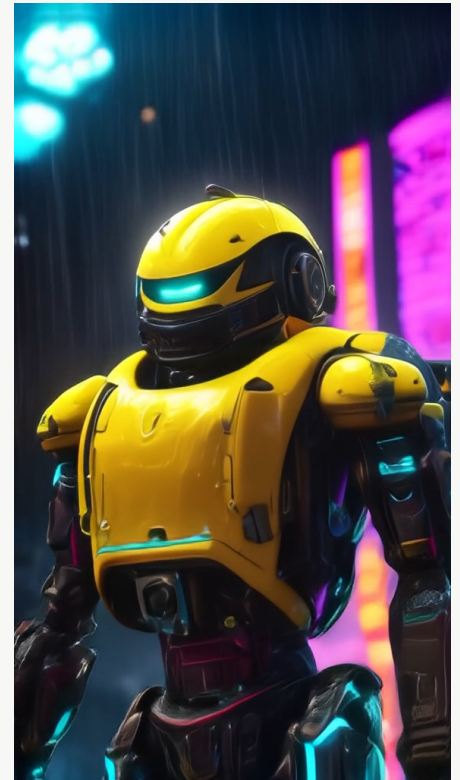
Robot amarillo bajo lluvia y neones. Fotogramas del MP4 incluido en el kit.



INICIO



MITAD



FINAL

**Imagen SDXL**

768 × 1368 · 24 pasos · 82,74 s

**Video LTX**

576 × 1024 · 8 pasos · 80,63 s

**Clip guardado**

65 fotogramas · 24 fps · 2,71 s

**Equipo de referencia:** MacBook Pro M4 Pro, GPU de 20 núcleos y 48 GB de memoria unificada. ComfyUI 0.42.1, PyTorch 2.14.1 y Python 3.13.12; generación con MPS y T5 en CPU.

Son tiempos registrados de una ejecución, no promedios ni una promesa para otros equipos. La primera carga, otras aplicaciones abiertas y las versiones pueden cambiar el rendimiento. En el clip se aprecia movimiento; algunos detalles mecánicos cambian entre fotogramas.

Medición local de la demo. El tiempo de generación y la duración del video son cosas distintas.

## 12 / SI ALGO FALLA

# Resuelve lo habitual.

No hace falta reconstruir el gráfico: empieza por esta comprobación.

## “Model not found” o lista vacía

Comprueba carpeta, nombre exacto y que la descarga terminó. Reinicia ComfyUI y vuelve a elegir el archivo en el nodo.

## Te pide Qwen u otro modelo distinto

Probablemente abriste otro template. Crea una pestaña y carga uno de los JSON del kit. Estos flujos usan SDXL, LTX y T5.

## Nodos rojos o “missing nodes”

Actualiza el motor ComfyUI de tu instancia a una versión estable que incluya estos nodos y reinicia. El kit no necesita custom nodes.

## El equipo se queda sin memoria

Cierra apps pesadas. Usa batch 1. En video, baja a 288 × 512 y 33 frames; conserva T5 en CPU. Si sigue fallando, tu equipo puede necesitar más memoria.

## Carga lento o parece detenido

Revisa la cola y Console. Espera la carga inicial de modelos antes de volver a pulsar Run; repetirlo acumula trabajos.

## Video negro, error de tipo o imagen extraña

Confirma LTX 2B 0.9.8 Distilled y T5 FP16. En Mac evita sustituir T5 por FP8. Prueba de nuevo el JSON original y la imagen del kit.

## No encuentro el archivo final

Abre Assets o tu carpeta de salida. En Desktop revisa Settings > Storage > Shared Output; busca Reel\_Robot\_Neon.

## El robot cambia demasiado

Pide menos movimiento, una sola toma y una cámara suave. Cambia la seed para probar otra variación; la continuidad perfecta no está garantizada.

Si aparece un error, copia el mensaje de Console junto con tu modelo de GPU, memoria y versión de ComfyUI.

## LISTO PARA CREAR

# Tu primera sesión, en orden.

Una guía para volver al proceso cada vez que quieras probar una nueva idea.

1. ComfyUI abierto en **Local**.
2. Los tres modelos descargados y reconocidos.
3. Workflow SDXL abierto en su pestaña.
4. PNG generado y guardado.
5. Workflow LTX abierto en otra pestaña.
6. Imagen cargada y descripción del movimiento revisada.
7. MP4 generado, descargado y listo para editar.

## Sin créditos por generación

“Infinitos” significa que estos workflows locales no tienen una cuota de créditos. Cada generación consume tiempo, energía y espacio. Después de instalar y descargar los archivos, puedes usarlos sin conexión.

## Enlaces de referencia

[Descargar Comfy Desktop](#)

[Instalación oficial: Mac](#)

[Instalación oficial: Windows](#)

[Carpetas de modelos, entradas y salidas](#)

[ComfyUI: código y documentación](#)

[SDXL Base 1.0: ficha del modelo](#)

[LTX-Video: ficha del modelo](#)

Conserva este PDF junto al **Kit\_LOCAL**. Los archivos JSON, prompts, imagen y clip de ejemplo te permiten repetir el recorrido y cambiar la escena a tu gusto.