



Abraham Olvera | Inteligencia Artificial

@abrahamolvera.ai

El video de la caja: hazlo con tu producto

Dos métodos desde cero: referencias en ChatGPT y animación en Higgsfield, manual o por API. También tienes el video base del choque y la caja cerrada en el piso, con y sin audio, para añadir después la apertura y tu producto en CapCut.

Newsletter

Gracias por estar aquí. Eres parte del 14.5%. En este momento eres parte del 14.5% de la población mundial que usa inteligencia artificial. Dentro de eso está el 0.6% que son avanzados. De los cuales el 0.1% están al día todo el tiempo. Si quieres ser parte del 0.1%, suscríbete gratis a mi newsletter.

Suscribirme gratis

Elige cómo hacerlo

Si quieres empezar rápido, descarga la plantilla y ve directo a CapCut. Si quieres cambiar la ciudad, el camión o toda la escena, crea tus imágenes y anima dos clips.

Método 1: ChatGPT + Higgsfield manual. Para quien quiere trabajar desde el navegador, sin código.

Método 2: ChatGPT + API de Higgsfield. Para quien prefiere ejecutar el proceso desde su computadora. La API tiene facturación separada; no asumas que usa los créditos de tu suscripción web.

La plantilla original es vertical, 480 × 854, 24 fps y dura 6.04 s. Termina con la caja pequeña cerrada en el piso del puente; no incluye la apertura ni una hoja dentro de la caja. Añade la apertura con la referencia 3 o con tu propio clip antes de colocar el producto. Exportarla a 1080p no recupera detalle que no existe.

Descargas

Video con audio: <https://abrahamolvera.ai/caja/caja-plantilla-con-audio.mp4>

Video sin audio: <https://abrahamolvera.ai/caja/caja-plantilla-sin-audio.mp4>

Kit completo (referencias, prompts y script): <https://abrahamolvera.ai/caja/kit-caja.zip>

Paso común · Crea las referencias en ChatGPT

1. Abre ChatGPT y empieza una conversación nueva. Adjunta una foto clara de tu producto si quieres preparar el final desde ahora.
2. Pega el prompt de la referencia 1. Indica 9:16. Abre la imagen y usa Guardar/Descargar; guarda el archivo original, no una captura de la pantalla.
3. En el mismo chat, adjunta la imagen anterior y pide la referencia 2. Revisa continuidad: camión completo, mismo puente y cajas en el suelo solamente después del golpe; la caja pequeña debe quedar cerrada.
4. Pide la referencia 3 adjuntando la imagen posterior. Esta referencia sí muestra la apertura posterior: la caja debe seguir en el piso y sus manos, solapas y hoja deben ser coherentes.
5. Guarda 01-antes.png, 02-despues.png y 03-abierta.png. Si algo no coincide, corrige la imagen antes de gastar una generación de video.

ChatGPT permite editar una imagen adjunta y limitar el cambio a una zona, aunque puede alterar áreas cercanas. Por eso revisa cada referencia antes de pasar a Higgsfield. El kit incluye nuestras referencias como ejemplo; puedes empezar con ellas.

Prompts para las imágenes

1. Referencia 1 · Antes del choque

Pide esto

Crea una imagen vertical 9:16 que parezca un fotograma de un celular viejo. Vista desde arriba de un puente peatonal verde de Ciudad de México, pintura descascarada, suciedad y barandal en primer plano. Abajo viene un camión mexicano viejo rojo y crema, cabina y plataforma completas visibles, cargado con cajas de cartón apiladas más altas que el paso libre del puente. La carga todavía está intacta: NO hay cajas tiradas, NO hay choque todavía. Calle urbana real, coches detrás. Luz nublada, ruido digital moderado, composición espontánea. El camión está a punto de pasar directamente debajo del puente. Sin texto, sin logos inventados.

[Guarda esta imagen como 01-antes.png. Si el camión sale cortado o ya hay cajas en el suelo, corrige antes de avanzar.](#)

2. Referencia 2 · Después del choque

Pide esto

Usa la imagen anterior como referencia y conserva exactamente el mismo puente, camión, colores, calle y punto de vista. Crea el fotograma posterior al choque de la carga demasiado alta contra el borde del puente. El camión está detenido directamente debajo del puente, NO lejos. Su cabina y plataforma pertenecen al mismo vehículo y siguen completas. Parte de la carga quedó deformada y algunas cajas cayeron a la calle; los coches de atrás frenan. Una caja pequeña cerrada aterrizó en el piso del puente, cerca de los pies de dos amigos que se apartaron. Vista de celular vertical 9:16, encuadre amplio, cámara ligeramente inclinada, suciedad y luz iguales. Sin personas heridas y sin texto.

[Adjunta también 01-antes.png. Guarda 02-despues.png. El origen y destino deben tener el mismo camión y geometría.](#)

La caja abierta

3. Referencia 3 · Caja abierta

Pide esto

Usa la referencia posterior al choque. Mismo piso, barandal verde, luz, manos y caja pequeña. La cámara ahora se acerca a la caja que sigue apoyada en el piso del puente. Dos manos abren las solapas. La caja está abierta y su interior contiene una hoja de papel lisa SIN texto, dejando un espacio amplio para colocar un producto después. El camión se mantiene detenido debajo del puente y los coches frenados. Encuadre vertical 9:16 de celular, realista; la caja NO flota y NO cambia de tamaño. Conserva la perspectiva y muestra las solapas completas.

[Adjunta 02-despues.png](#); [guarda 03-abierta.png](#). Esta es tu referencia de llegada para el clip de apertura.

Método 1 · Higgsfield manual

1. Entra a Higgsfield y abre el generador de video. Elige Seedance 2.5 y la modalidad de imagen a video o referencias que esté disponible en tu cuenta.
2. Para el clip A, sube 01-antes.png como inicio y 02-despues.png como final si la modalidad permite imagen final. En referencias múltiples, explica en el prompt qué imagen corresponde a cada momento.
3. Pega el prompt A. Elige 6 s, formato vertical heredado de las referencias y 720p si está disponible. Activa audio si quieres intentar la reacción generada. Revisa el costo mostrado antes de generar.
4. Reproduce a velocidad normal. Debe verse contacto primero y caída después, y el clip debe terminar con la caja pequeña cerrada en el piso. El camión tiene que detenerse debajo del puente.
5. Repite con 02-despues.png → 03-abierta.png y el prompt B. Descarga ambos clips. Si tu modalidad solo permite inicio, usa cada imagen como punto de partida y verifica el final; no es equivalente a fijar una imagen de llegada.

Los prompts de animación

4. Prompt A · Choque y caída

Pide esto

Animate from the intact-load reference to the aftermath reference. Vertical handheld phone recording from the pedestrian bridge in Mexico City, real-time action. The complete old red-and-cream truck moves forward directly under the fixed green bridge. Its over-height cardboard cargo visibly contacts the bridge first, compresses, then cartons tumble under gravity. No cartons on the road before contact. The truck brakes and stops directly beneath the bridge; following cars brake because cartons block the lane. One small closed carton lands on the bridge deck and stays on the floor. Two friends recoil and move aside; camera shakes naturally. Preserve truck, bridge and street geometry. No slow motion, no reverse motion, no flying box, no teleportation, no captions. Off-camera Mexican Spanish voices: “Ahhh, no manches... ahí viene... ¡quítate, quítate!” The warning happens immediately BEFORE impact.

Inicio: 01-antes.png. Final: 02-despues.png. Genera primero solo este clip (6 s). Si la física falla, corrige referencias o divide contacto y caída en dos clips; un prompt no garantiza el resultado.

5. Prompt B · Apertura

Pide esto

Start with the small closed carton resting on the pedestrian bridge deck and finish at the open-box reference. Same hands, carton, floor, dirty green railing and lighting. Handheld phone camera approaches quickly; a friend crouches, grasps the carton on the floor and opens the flaps with natural hand motion. The carton remains supported on the floor, no teleportation. The stopped truck remains directly below the bridge and traffic stays braked. Normal-speed action, no slow motion. Finish on a clearly visible open interior, no generated writing. Preserve proportions and perspective. Ambient city sound only, no speech, no music, no captions.

Inicio: 02-despues.png. Final: 03-abierta.png. Genera 6 s y recorta la espera al editar. No le pidas al modelo choque, carrera, apertura y producto en un solo salto.

Método 2 · Higgsfield por API

1. Crea las mismas tres imágenes en ChatGPT. La API anima las referencias; no omitas ese trabajo.
2. Entra a open.higgsfield.ai y crea una API key. Copia la clave completa. Revisa acceso, saldo y precio del modelo antes de enviar una generación.
3. Instala Node.js 18 o posterior desde nodejs.org. Descarga el kit y descomprímelo. Abre Terminal (Mac) o PowerShell (Windows) en esa carpeta.
4. Abre API-LEEME.md: explica cómo guardar tu clave en la variable HF_API_KEY y ejecutar generar-seedance.mjs. No pegues tu clave dentro de ChatGPT ni la publiques.
5. Ejecuta un clip A con 01-antes.png y 02-despues.png, y después el clip B con 02-despues.png y 03-abierta.png. Usa los mismos prompts de animación.
6. El script sube tus archivos, envía una sola solicitud, guarda el ID y consulta el resultado. Genera 6 s, 720p, MP4 y sin audio. Agrega la reacción en CapCut si la necesitas.

Ejemplo: `node generar-seedance.mjs 01-antes.png 02-despues.png "PEGA EL PROMPT A" clip-a.mp4`

Este ejemplo sigue el contrato oficial de Seedance 2.5 Image-to-Video. Se comprobó localmente sin generar un video de pago; el acceso real depende de tu cuenta. Un timeout no significa que la generación se canceló: conserva el request_id y revisa el estado antes de volver a enviarla.

Pon tu producto dentro de la caja

6. Pon tu producto en la caja · ChatGPT

Pide esto

Edita la primera imagen (caja abierta) usando la segunda imagen como referencia exacta de mi producto. Sustituye únicamente la hoja y el contenido del interior por **[PRODUCTO]**, apoyado en el fondo de la caja. Mantén intactos las manos, solapas, piso, barandal, encuadre y posición de la cámara. Respeta la perspectiva, escala, luz y sombras de contacto. Conserva fielmente la forma, color, empaque y logotipo visibles en mi foto; no inventes texto ni accesorios. El producto debe caber físicamente y quedar parcialmente oculto por los bordes de la caja cuando corresponda. Fotograma realista vertical 9:16, producto nítido y entorno de celular.

Adjunta [caja-abierta-referencia.png](#) y una foto clara de tu producto. Reemplaza **[PRODUCTO]**. Si tu etiqueta necesita exactitud, usa la foto o el video original superpuesto en CapCut.

Cómo unir la revelación

No extraigas el último fotograma esperando encontrar una caja abierta: la plantilla termina con la caja cerrada en el piso. Usa [03-abierta.png](#) o [caja-abierta-referencia.png](#) del kit como referencia de llegada, y súbela a ChatGPT junto con tu foto de producto.

Guarda el resultado y úsalo como inicio de otro clip de Higgsfield. En CapCut, corta la plantilla cuando la caja ya esté apoyada en el piso y añade después la apertura o tu video de producto. Procura que la caja, manos y luz coincidan; si hay un salto evidente, usa un corte directo a un video real de tu producto.

Anima el producto o cambia la hoja

7. Animar tu producto · Higgsfield

Pide esto

Use the uploaded open-box product image as the starting frame. Keep the product exactly the same shape, packaging, color and label, physically supported inside the box. A subtle handheld phone push-in reveals the product, preserving hands, carton flaps, perspective, shadows and lighting. No rotating product, no morphing, no new objects, no captions, no speech. Real-time movement.

Genera un clip corto de 4-6 s y usa solo 1-2 s al final. Únelo al punto donde se abre la caja; si el corte cambia demasiado, usa un plano de tu producto a pantalla completa.

8. Texto opcional escrito a mano

Pide esto

Edita SOLO la hoja dentro de la caja. Escribe con pluma azul, como una persona real: “Comenta [PALABRA] y te paso los pasos para hacerlo para tu producto”. Letra manuscrita sencilla, trazos con presión variable, espaciado irregular sutil, sin fuente tipográfica y sin caligrafía perfecta. Mantén el papel, las sombras, la perspectiva y todo lo demás. Prioriza que cada palabra sea legible.

Para asegurar legibilidad, también puedes fotografiar una hoja escrita por ti y colocarla como referencia. Evita generar texto mientras el papel está moviéndose.

CapCut · El final más sencillo

1. Crea un proyecto 9:16. Importa la plantilla con audio y un video real de tu producto (1-3 s), grabado con buena luz y sin movimientos bruscos.
2. Pon la plantilla en la pista principal. Avanza hasta el final, alrededor de 6.0 s, cuando la caja pequeña ya está cerrada en el piso. Divide ahí y añade tu clip de apertura o producto; la plantilla no contiene una hoja que tengas que ocultar.
3. Añade inmediatamente el video de tu producto en la misma pista. Llena la pantalla, ajusta el encuadre y corta la espera del principio. Un corte directo suele verse mejor que una transición larga.
4. Conserva el audio del choque y baja el audio del video de producto si estorba. Con la plantilla sin audio puedes grabar tu propia voz y agregar efectos.
5. Añade nombre del producto y tu CTA sobre el plano final. Exporta vertical a 24 o 30 fps; comprueba el archivo antes de publicarlo.

Si quieres que aparezca dentro de la caja

En CapCut de escritorio coloca el video de producto en una pista superior. En móvil usa Superposición → Añadir superposición. Empieza en el momento en que el interior se vuelve visible. El video debe tener el producto centrado y un fondo fácil de recortar.

CapCut · Superposición dentro de la caja

1. Selecciona la superposición. Ajusta escala, posición y rotación para cubrir la hoja sin invadir las manos ni las solapas.
2. Aplica Video → Máscara (Mask), con un rectángulo que limite el producto al interior. Usa poco difuminado de borde. Si tu versión ofrece recorte de fondo, úsalo para aislar el producto.
3. Pon un fotograma clave (rombo) al principio con esa posición/escala. Avanza 3–5 fotogramas y corrige. Repite hasta el final para que acompañe la cámara. No asumas que el tracking automático hará bien las oclusiones.
4. Si una solapa o mano pasa por delante del producto, coloca una copia del clip original en una pista encima y usa máscaras pequeñas para conservar ese borde como primer plano.
5. Para un plano más fácil, congela el último fotograma de la caja abierta y extiéndelo 1–2 s. El producto puede seguir reproduciéndose encima, pero la caja ya no necesita seguimiento.
6. Revisa el resultado a velocidad normal: sin bordes flotantes, producto atravesando las manos o perspectiva imposible. Si la cámara cambia mucho de ángulo, el recorte rectangular no basta: usa el corte a pantalla completa.

Los nombres y herramientas dependen de la versión de CapCut. Posición, escala y rotación no corrigen por sí solas una perspectiva de cuatro esquinas; no fuerces una superposición que parezca un sticker.

Audio, ritmo y control final

Si generas todo desde cero, puedes grabar una reacción con acento mexicano: “Ahhh, no manches... ahí viene... ¡quítate, quítate!”. Coloca el último aviso justo ANTES del golpe. Los efectos de impacto y frenado entran en el contacto, no varios segundos después.

Mantén el choque a velocidad normal. Acelera solo el tramo de acercarse y abrir la caja si hay tiempos muertos; reserva 1-2 s para ver el producto. La reacción generada puede salir distinta: escucha y ajusta en CapCut.

Antes de publicar: carga intacta al inicio; contacto visible; cajas caen por gravedad; camión debajo del puente; manos y caja coherentes; producto fiel a tu foto; texto legible; audio sincronizado. No hay un prompt que garantice física perfecta a la primera.

Fuentes y material

ChatGPT: help.openai.com/en/articles/11084440-images-in-chatgpt

Seedance 2.5 API: open.higgsfield.ai/models/bytedance/seedance-2.5/image-to-video/api-reference

Inicio y subida de referencias: open.higgsfield.ai/quick-start

CapCut: capcut.com/resource/overlay-video y capcut.com/resource/how-to-use-capcut

Documentación revisada el 4 de octubre de 2026. La escena y el audio de la plantilla fueron generados con IA.

Sigue usando la guía

Puedes leer esta guía, copiar los prompts y descargar el PDF en:

<https://abrahamolvera.ai/g/caja/>