Video viral con IA: tu Roadster volador en Higgsfield

@StephanoBrooke · Palabra clave: TESLA

1. Lo que necesitas

Una cuenta de Higgsfield con créditos. Cada imagen y cada video gasta créditos, así que revisa el costo del modelo antes de generar. Los borradores en 480p cuestan menos y los puedes subir a 1080p después.

Si todavía no tienes cuenta en Higgsfield, puedes entrar con mi enlace de referido: https://higgsfield.ai/s/higgsfield-genjutsu-infinite-ads-ig-stephanobrooke-QxbWcB . Es la misma plataforma con la que hice este video.

Tres fotos tuyas o con permiso de la persona: un selfie de frente con luz pareja y dos fotos de lugar (una calle normal con tráfico y unos cables de luz).

Una foto del vehículo si quieres usar uno real como referencia. En mi caso fue el Roadster, pero al final la toma 1 salió mejor sin esa foto.

Un editor para unir las 3 tomas: CapCut sirve perfecto. Yo lo armé con ffmpeg.

Los modelos que usé: GPT Image 2.5 para las imágenes base y Seedance 2.5 para los videos.

2. Cómo elegir tus fotos de referencia

Las fotos buenas hacen la mitad del trabajo. Estas son las reglas que me funcionaron:

- Selfie: de frente, cara grande, luz pareja, sin lentes oscuros ni filtros. Si la cara sale chica, la voz y los labios no se sincronizan bien.
- Calle: una calle normal de tráfico, con los coches pegados unos con otros. Sin marchas, sin bloqueos, sin policías, sin letreros legibles, sin escenas religiosas. Los modelos se confunden o bloquean si ven gente, texto o símbolos.
- Cables: una foto de un poste con muchos cables negros enredados. Entre más caótico, mejor se ve el gag.
- Vehículo (opcional): foto limpia y de lado. Ojo: si pones un coche real famoso junto con una imagen final, el filtro de propiedad intelectual puede bloquear el video (te explico cómo lo resolví en el paso 6).

Si tu foto de calle es horizontal, no pasa nada: el prompt 1 la extiende a formato vertical 9:16.

3. Sube tus fotos a Higgsfield

Sube las tres fotos (selfie, calle y cables) a tu biblioteca de Higgsfield. Si tu selfie es HEIC (formato de iPhone), conviértela a JPG o PNG antes. Cuando las uses como referencia, el orden importa: en cada prompt los llamo primera, segunda y tercera imagen.

Qué pasa después: ya tienes tus referencias listas para usarlas en los prompts de imagen del siguiente paso.

4. Las imágenes base (keyframes)

Primero generas imágenes y después las animas. Así controlas cómo se ve la cara y la calle antes de gastar créditos en video. Usa GPT Image 2.5, formato 9:16.

Prompt 01: Imagen base de la calle con poste y cables

> *"Vertical 9:16 realistic phone photo of the gridlocked Mexico City street from the first reference image: same bumper-to-bumper cars, same layout and daylight, extended naturally to a tall vertical frame. Add a concrete utility pole in the right foreground with a huge chaotic tangle of black electrical cables like the second reference image, cables also stretching across the street overhead. Empty sidewalk in the lower foreground. No people close to camera, no readable signs, no readable license plates, no text, no logos."*

Referencias: la foto de calle (primera) y la foto de cables (segunda). Calidad media. Qué pasa después: obtienes tu calle en vertical con un poste lleno de cables en primer plano. Esta imagen es el primer cuadro de la toma 1.

Prompt 02: Imagen final: el coche colgado en los cables

> *"The exact same street, pole, cars and framing as the first reference image. The sleek white two-door electric hypercar from the second reference image now hangs upside down, wheels up, completely tangled and snagged in the huge mass of black power cables on the concrete utility pole in the right foreground, cables wrapped around the car body, a few small blue-white sparks at the cables. A few drivers in the stationary cars lean out of their windows looking up. Realistic phone photo, daylight, no text, no logos, no readable plates."*

Referencias: la imagen del prompt 01 (primera) y la foto del vehículo (segunda). Calidad alta. Qué pasa después: obtienes el coche colgado de cabeza en los cables, en la misma calle. Esta imagen es el último cuadro de la toma 1 y la base de la toma 2.

Prompt 03: Tu cara colgada dentro del coche

"Unretouched iPhone photo taken from the sidewalk on the same Mexico City street as the first reference image, looking up at the white two-door hypercar from the second reference image hanging upside down, wheels up, snagged in the tangle of black power cables on the concrete utility pole, cables wrapped around the car body, a few small sparks. The same gridlocked cars and the same buildings as the first reference image are visible below; no skyline, no palm trees. Through the open side window the man from the third reference image hangs upside down in his seat with his seatbelt on, face clearly visible at medium distance, his exact face, black wavy hair, brown eyes, white t-shirt, hair hanging down, calm deadpan expression, looking down toward the camera. Natural skin texture with visible pores, harsh midday sun with hard shadows, slight motion blur, phone camera noise, imperfect slightly off-center framing, vertical 9:16. No text, no logos, no readable plates."

Referencias: la imagen del prompt 02 (primera), la foto del vehículo (segunda) y tu selfie (tercera). Calidad alta y genera 2 opciones. Qué pasa después: escoges la imagen donde tu cara se ve más natural. Aquí es donde importan las palabras de iPhone sin retocar, poros y sol duro: sin ellas la cara se ve de plástico.

Prompt 04: El señor de la CFE con su torta

"Realistic vertical 9:16 phone photo, static medium shot at chest height. A Mexican man in his 50s with a plain white hard hat and a plain gray work uniform, holding a torta wrapped in paper, chewing, looking up with a relaxed unimpressed expression. Behind him a gridlocked Mexico City street with stationary cars, softly blurred, daylight. No logos, no company marks on the hard hat or uniform, no text, no readable signs."

Sin referencias. Calidad media. Qué pasa después: obtienes el personaje de la toma 3. Evita logos y marcas en el casco y el uniforme para no toparte con filtros ni con problemas de marca.

5. Las 3 tomas de video

Usa Seedance 2.5 en modo Omni Reference, formato 9:16, con audio activado y en borrador (480p). Genera 2 opciones de cada toma y quédate con la mejor. Una toma por generación: no intentes meter todo el chiste en un solo video.

Prompt 05: Toma 1: el coche llega volando y se enreda

"Realistic handheld phone video, vertical 9:16, bright daylight, comedy tone. The first frame is a gridlocked Mexico City street with a concrete utility pole covered in a huge tangle of black power cables in the right foreground; the last frame shows a white sports car hanging upside down, snagged in those cables. Keep the street, cars and layout identical. 0-0.8s: static wide shot of the stationary traffic, a low rumble and a rising whoosh building up. 0.8-1.9s: a white two-door electric sports car shoots into frame from the left, very fast and very low, just above the roofs of the stationary cars, large in the frame, heavy motion blur, two jets of white compressed air blasting backward from its rear leaving a long horizontal white trail behind it; it is not hovering and the gas does not point down. The air blast makes dust and papers fly, drivers flinch and turn their heads. The phone camera whip-pans right to follow the car. 1.9-3s: the car, nose slightly up and tilting, slams into the tangle of cables on the pole; cables stretch like rubber bands, blue-white sparks burst, the car spins over. 3-5s: it ends hanging upside down, snagged in the wires, swinging slowly with small sparks; drivers in the stationary cars lean out and look up. Camera: handheld with a fast whip pan, natural shake. Audio: loud whoosh, thruster hiss, car horns, heavy cable twang, electrical crackle, then creaking cables and a moment of silence. No music, no dialogue. Constraints: no explosion, no fire, no text, no logos, no readable signs or license plates."

Ajustes: 5 segundos. Imagen inicial: la del prompt 01. Imagen final: la del prompt 02. No pongas la foto del vehículo como referencia en esta toma. Qué pasa después: ves el coche cruzar muy rápido sobre el tráfico, chocar con los cables y quedar colgado de cabeza. Los tiempos por segundo son lo que evita que el coche se quede flotando.

Prompt 06: Toma 2: tú colgado diciendo la línea

"Realistic handheld phone video, vertical 9:16, deadpan comedy, same street as the previous shot. The starting frame shows a white sports car hanging upside down from the tangle of power cables on a pole, with a man hanging upside down inside, seen through the side window. Keep his face identical to image 2: same face, black wavy hair, brown eyes, white t-shirt. 0-2s: the handheld phone camera zooms in slowly and shakily toward the car window until his face fills about a third of the frame; cables sway, small sparks fall, the car swings slightly. 2-3s: he blinks once and looks down at the camera. 3-5s: he says in Mexican Spanish, polite, calm, slightly tired voice: "Buenas... ¿me podría bajar?" Lips synced to the line, minimal head movement. 5-6s: he stays still, one slow blink. Camera: handheld, natural shake, slight autofocus hunting. Audio: creaking cables, distant car horns and city traffic, clear dialogue. Constraints: the face must not change, natural skin texture, no text, no logos, no subtitles."

Ajustes: 6 segundos. Imagen inicial: la del prompt 03. Referencia: tu selfie. Qué pasa después: la cámara se acerca a tu cara, parpadeas y dices la frase. Revisa el audio: confirma que la voz salga en español y que los labios coincidan.

Prompt 07: Toma 3: el señor de la CFE

"Realistic phone-video look, vertical 9:16, deadpan comedy. The starting frame is this man: keep him exactly (white hard hat, gray work uniform, torta in paper) and the same gridlocked street behind him. 0-1s: he chews, looking up at something above the camera, unimpressed. 1-2.5s: he says in Mexican Spanish, flat bored tone: "Ahorita le llamo a la CFE." 2.5-3.5s: an off-screen male voice asks: "¿Y cuánto tardan?" 3.5-5.5s: he answers while chewing, with no change in expression: "Mañana. O el lunes." 5.5-7s: he takes another bite and keeps staring up. Camera: static medium shot at chest height, no movement. Audio: street ambience, clear Spanish dialogue, chewing. No music. Constraints: no text, no logos or company marks on the hard hat or uniform, no subtitles."

Ajustes: 7 segundos. Imagen inicial: la del prompt 04. Qué pasa después: el señor mastica, dice su línea, alguien le contesta fuera de cámara y remata con el chiste. Revisa el audio igual que en la toma 2.

5. Las 3 tomas de video (tip)

Tip general de los tres prompts: descríbelo como un video de celular (handheld, vertical, con temblor natural) y marca qué pasa en cada segundo. Los videos que piden cámara cinematográfica se ven más falsos, y justo lo que queremos es que parezca que alguien lo grabó con el teléfono.

6. Qué hacer cuando algo sale mal

Estos fueron los tropiezos reales que tuve y cómo los resolví:

- El coche flotaba lento y no terminaba colgado (toma 1, intento 1): el prompt tenía un chorro de aire hacia abajo y no tenía final. Solución: generar la imagen final del coche colgado (prompt 02) y usarla como imagen final, y reescribir la acción como algo muy rápido y bajo, con la cámara siguiéndolo.
- El video se bloqueó con estado de propiedad intelectual (ip_detected): pasó al combinar un coche real conocido con una imagen final. Solución: quita la foto del vehículo como referencia y deja solo imagen inicial e imagen final. Describe el coche como white two-door electric sports car, sin poner la marca en el prompt.
- Tu cara se ve falsa o de plástico (toma 2, intento 1): también salió una ciudad distinta, con rascacielos y palmeras. Solución: usar como referencia la imagen de la misma calle y agregar las palabras de realismo de iPhone, poros, sol duro y ruido de cámara. Funcionó.
- Higgsfield te sugiere efectos (presets) antes de generar, como IN THE DARK, DROWN IN MUSIC o un zoom a la Tierra. Si no los quieres, recházalos y vuelve a mandar tu generación para que use solo tu prompt.
- Los propulsores salen como fuego o explosión: escribe que son jets de aire comprimido blanco, no fuego, y agrega no explosion, no fire en las restricciones.

7. Arma el video completo

Une las tres tomas en orden: coche enredado, tú colgado, señor de la CFE. Todas son verticales 9:16 y duran 5, 6 y 7 segundos, así que el corte final queda en unos 18 segundos.

- Pon un corte directo entre tomas, sin transiciones.
- Agrega al final este texto en pantalla, en tres líneas: Tesla Roadster: 0 a 100 / en menos de 1 segundo. / CFE: mañana.
- En CapCut: importa las tres tomas, ponlas en línea de tiempo, agrega texto en los últimos segundos con letra gruesa y color amarillo, y exporta en vertical.

Si quieres hacerlo con ffmpeg, junta los clips a la misma resolución y 24 cuadros por segundo, y agrega el texto con el filtro drawtext. Funciona igual, pero CapCut es más fácil si es tu primera vez.

Qué pasa después: tienes tu video listo para subir. Antes de publicar, ponlo en 1080p (paso 8) si te gusta cómo quedó.

8. Súbelo a 1080p

Los borradores en 480p se pueden finalizar a 1080p dentro de los 7 días siguientes, sin volver a escribir el prompt. Hazlo solo con las tomas que decidiste usar, para no gastar créditos de más. Después reemplaza los clips en tu editor y exporta de nuevo.

9. Lo que debes saber antes de publicar

Los videos con IA no salen perfectos a la primera. En mi caso hice varias versiones de la toma 1 y de la toma 2 antes de quedarme con las buenas. Es parte del proceso, y por eso conviene empezar en borrador.

Revisa el audio siempre. Las voces en español y la sincronía de los labios pueden variar entre una generación y otra, aunque el prompt sea el mismo.

Usa solo tu cara, o la de alguien que te dio permiso por escrito. No uses la cara de otra persona sin su consentimiento.

Avisa que es IA. Dilo en el caption (por ejemplo: hecho con IA) o con la etiqueta de contenido generado de la plataforma. Es un chiste, no un hecho: ningún Roadster real se quedó colgado de ningún poste de la CDMX.

No pongas logos ni marcas reales en los uniformes o en los coches. Evitas bloqueos del modelo y problemas con las marcas.

Los nombres de los botones en Higgsfield pueden cambiar con el tiempo. Probé estos prompts con el conector de Higgsfield, así que si algo se llama distinto en tu pantalla, busca la opción equivalente.

10. Cómo usarlos (resumen del orden)

- Elige y sube tus fotos: selfie, calle y cables.
- Genera las 4 imágenes base con los prompts 01 al 04.
- Genera las 3 tomas con los prompts 05 al 07, en borrador y con 2 opciones cada una.
- Si algo sale mal, revisa el paso 6.
- Une las tomas, agrega el texto final y revisa el audio.
- Sube a 1080p solo lo que vas a usar, publica y avisa que es IA.