Anuncios infinitos con IA: cambia persona y producto

@StephanoBrooke · Palabra clave: JUTSU

1. Lo que necesitas

- Una cuenta de Higgsfield con créditos. Mis costos reales: cada imagen de persona cuesta 2.75 créditos (GPT Image 2.5, calidad alta, 2K), el video completo de 15.5 s en 720p costó 112 créditos y el cambio de voz costó 2. El camino final fueron unos 125 créditos, sin contar mis pruebas.
- Un clip tuyo en horizontal (16:9), de unos 15 segundos, hablando a cámara en selfie y con la cámara en mano lo más estable posible. Mientras hablas del producto, abre la mano hacia un lado: ahí es donde la IA pondrá el producto.
- Un producto simple, sin marca ni logo visibles (yo usé una botella lisa de colores). Con logos la IA los deforma y tienes problemas de derechos.
- Un editor de video (CapCut sirve) para recortar el video, poner subtítulos y armar el reel.
- El archivo oficial del logo de Higgsfield si vas a ponerlo en el video. Pídelo a Higgsfield, no lo recrees.

2. Qué hace este video

Grabas un solo clip hablando. Con Genjutsu de Higgsfield lo conviertes en cuatro personas distintas con su producto, cada una en su cuarto y haciendo exactamente tus gestos, y al final le pones voz de mujer. Es la idea de los anuncios infinitos: una grabación, muchas versiones.

Se usan tres piezas de Higgsfield: un modelo de imagen para crear las personas con producto, Genjutsu en su modelo Motion Control para pasarles tu movimiento, tus gestos y tus labios, y la herramienta de cambio de voz para que la voz coincida con las mujeres.

Motion Control es distinto de Replace Object: no repinta tu video, toma el movimiento de tu clip y se lo pasa a las personas de tus imágenes de referencia. Por eso aquí cambia toda la persona, no solo el fondo.

3. Paso 1: graba tu clip

Graba unos 15 segundos hablando a cámara, en horizontal, con el celular en mano como una selfie. Mi clip duró 15.5 s (1280x720) y tenía este guion: Hola niñas, ¿cómo están? La realidad es que no existo, pero tengo aquí esta botella preciosa, tiene muchísimos colores y es perfecta para mí. Ahora puedes hacer anuncios en menos de 10 minutos. Comenta JUTSU y te mando la guía de cómo lo puedes hacer. Piensa en bloques: cada persona nueva aparece en un cambio de frase. Deja una pausa corta entre una idea y la siguiente, porque ahí caen los cortes. Y mueve la mano abierta hacia un lado en las frases del producto.
Si tu grabación empieza con ruido o con un error de grabación (a mí me sobraron los primeros 6 segundos), recórtalo y quédate solo con la parte útil. Ese recorte es tu @video1.

4. Paso 2: crea las personas con su producto

En Higgsfield elige el modelo de imagen GPT Image 2.5, formato 16:9, calidad alta y resolución 2K. Genera una imagen por persona. Cambia lo que va entre corchetes en el Prompt 01. Los Prompts 02 a 05 son las cuatro que usé yo.

Prompt 01 (Plantilla para generar una persona con tu producto):

> *"Photorealistic still frame from a handheld smartphone selfie video, 16:9 horizontal, a woman in her mid 20s speaking to camera, waist-up, natural skin texture, small black clip-on lavalier microphone on her top, soft natural light, no text, no logos. [hair and outfit], [room]. She holds [your product] in her raised hand toward the camera, plain with no label."*

Prompt 02 (Mi chica 1: rubia, sala luminosa, botella blanca (@image1)):

> *"Photorealistic still frame from a handheld smartphone selfie video, 16:9 horizontal, a woman in her mid 20s speaking to camera, waist-up, natural skin texture, small black clip-on lavalier microphone on her top, soft natural light, no text, no logos. Long straight ash-blonde hair, black sports crop top, bright Scandinavian-style living room with a white sofa and large windows. She holds a matte white insulated water bottle with a loop cap in her raised right hand toward the camera, plain bottle with no label."*

Prompt 03 (Mi chica 2: morena, luz dorada, botella rosa (@image2)):

> *"Photorealistic still frame from a handheld smartphone selfie video, 16:9 horizontal, a woman in her mid 20s speaking to camera, waist-up, natural skin texture, small black clip-on lavalier microphone on her top, soft natural light, no text, no logos. Long dark wavy brunette hair, white tank top, living room at golden hour with a warm floor lamp and a sofa. She holds a pastel pink insulated water bottle in her raised hand toward the camera, plain bottle with no label."*

4. Paso 2: crea las personas con su producto (continuación)

Prompt 04 (Mi chica 3: cabello negro con fleco, cuarto acogedor, termo lila (@image3)):

> *"Photorealistic still frame from a handheld smartphone selfie video, 16:9 horizontal, an East Asian woman in her mid 20s speaking to camera, waist-up, natural skin texture, small black clip-on lavalier microphone on her top, soft natural light, no text, no logos. Shoulder-length black hair with bangs, beige knit sweater, cozy room with a table lamp and warm light. She holds a lavender tumbler with a straw in her raised hand toward the camera, plain with no label."*

Prompt 05 (Mi chica 4: rizos, cocina verde, botella menta (@image4)):

> *"Photorealistic still frame from a handheld smartphone selfie video, 16:9 horizontal, a Black woman in her mid 20s speaking to camera, waist-up, natural skin texture, small black clip-on lavalier microphone on her top, soft natural light, no text, no logos. Voluminous curly hair, olive-green camisole, bright kitchen with pale green cabinets, a wooden counter with stools and a window. She holds a mint-green water bottle in her raised hand toward the camera, plain with no label."*

Qué pasa después: obtienes cuatro imágenes de personas distintas, cada una sosteniendo un producto de un color diferente en un cuarto diferente. Estas imágenes serán @image1 a @image4 en el siguiente paso. Revisa que la mano que sostiene el producto no esté deformada y que no salga ninguna marca ni letra en el producto.

5. Paso 3: una sola generación en Genjutsu (Motion Control)

Entra a Higgsfield, elige Genjutsu y el modelo Motion Control. Sube las referencias en este orden y pega el prompt de abajo:

- @image1: la chica 1 (la rubia con botella blanca).
- @image2: la chica 2 (botella rosa).
- @image3: la chica 3 (termo lila).
- @image4: la chica 4 (botella menta, cocina).
- @video1: tu clip completo recortado.
- Haz primero una prueba en 720p. Si todo se ve bien y quieres la versión final, genera en 1080p (cuesta más, revisa el costo en la app antes de generar).

Prompt 06 (Cambiar a 4 personas con producto en un solo video):

> *"Use @video1 as the master video. Keep its exact motion, gestures, lip movements, timing and handheld camera movement. Replace the man with a different woman in each part, same framing and scale as @video1. Hard cuts only, no morphing. Each woman keeps the small clip-on lavalier microphone on her top. 0-3s: the woman from @image1, same face, hairstyle and outfit, in the room from @image1, her hands empty with no bottle. 3-4.5s: hard cut, the same woman from @image1 in the same room, now holding the white bottle from @image1 in her raised open hand, clearly visible. 4.5-6s: hard cut, the woman from @image2 in the room from @image2, holding the pink bottle from @image2 in her raised open hand. 6-7.5s: hard cut, the woman from @image3 in the room from @image3, holding the lavender tumbler with a straw from @image3 in her raised open hand. 7.5-9s: hard cut, the woman from @image4 in the kitchen from @image4, holding the mint bottle from @image4 in her raised open hand. 9s to the end: the same woman from @image4 in the same kitchen, her hands empty with no bottle and no objects. Relight each woman to match her room. Photorealistic handheld smartphone look, no flicker, no extra people, no text, no logos."*

5. Paso 3: una sola generación en Genjutsu (Motion Control) (continuación)

Cambia los tiempos (3, 4.5, 6, 7.5 y 9 segundos) por los de tus propias palabras. Si usas una sola persona con una sola botella, usa el Prompt bonus 01. La chica 1 va dos veces porque en los primeros 3 segundos tiene las manos vacías y después sube la botella: por eso solo hay 4 imágenes para 5 cortes.

Qué pasa después: en unos 8 minutos obtienes un solo video con el mismo movimiento, gestos y labios que el tuyo, pero cada tramo con una mujer y un producto distintos. En mi caso salió de 16 segundos (el tuyo puede variar un poco) y los cortes cayeron cerca de los 3.4, 4.7, 6.0 y 9.0 segundos.

6. Paso 4: ponle voz de mujer

El video que genera Genjutsu no trae tu voz, así que primero hay que pegársela y después cambiarla.
A) Pega tu audio al video. En CapCut importa el video generado y tu audio original (el mismo tramo del recorte). El video de Genjutsu sale unos 0.5 s más largo que tu audio (el mío: 16 s contra 15.5 s), pero ese sobrante está al final: el movimiento y los labios van a la misma velocidad que tu clip. No aceleres ni cambies la velocidad del video, porque eso los desfasa poco a poco y al final queda adelantado. Solo recorta el video al largo de tu audio, silencia el audio del video generado y exporta un solo archivo mp4.
B) Pásalo por el cambio de voz de Higgsfield. Sube ese mp4 a la herramienta de cambio de voz (Voice Change), elige una voz de mujer y genera. Yo usé la voz Marisol. Conserva tus palabras y los tiempos y lo devuelve con la nueva voz. Costó 2 créditos y tardó menos de un minuto.
Qué pasa después: obtienes el mismo video con voz de mujer. Escúchalo completo antes de usarlo, sobre todo las palabras clave (JUTSU y guía). En mi prueba la transcripción automática entendió idea en lugar de guía al final: puede ser un error de la transcripción o de la voz. Si pasa en tu video, prueba otra voz o regraba esa frase.

7. Paso 5: arma el reel en CapCut

Crea un lienzo vertical 9:16 de 1080x1920 con fondo negro. Medidas tomadas del reel de referencia: logo de Higgsfield arriba (y 164 a 216), título (y 247 a 303; yo lo llamé Anuncios infinitos), el video de la IA (y 332 a 937, ancho completo, 1080x606), subtítulos en blanco y negritas (y 975 a 997) y un texto fijo abajo (y 1678 a 1713).
En la referencia el espacio de abajo (y 1031 a 1638) es otro video. Puedes dejarlo vacío, poner un fondo de color o poner tu video real, pero si pones tu voz original ahí, no uses la voz de mujer encima del mismo video.
Subtítulos de 2 a 3 palabras por vez, con estos tiempos: 0.6 Hola niñas, / 1.3 ¿cómo están? / 2.1 La realidad es / 2.5 que no existo, / 3.6 pero tengo aquí / 4.2 esta botella preciosa. / 6.0 Tiene muchísimos / 6.8 colores y es / 8.0 perfecta para mí. / 9.6 Ahora puedes hacer / 10.5 anuncios en menos / 11.0 de 10 minutos. / 12.1 Comenta JUTSU / 12.9 y te mando la guía / 13.8 de cómo lo puedes hacer.
Texto fijo abajo: Comenta JUTSU para la guía. En los últimos 3 segundos agrega Sígueme en texto grande. Si tu audio no dice Sígueme, ponlo en pantalla.

8. Qué revisar antes de publicar

- La cara de cada persona: que no cambie de cara a mitad de su tramo.
- La mano y el producto: que la mano tenga cinco dedos y el producto no se deforme.
- El micrófono de solapa: que se vea en cada persona (es un detalle que le da realismo).
- El fondo de los primeros 3 segundos: en una de mis pruebas se coló el fondo original de mi video en ese tramo. Si te pasa, regenera o prueba el Prompt bonus 04.
- Los cortes: que caigan cerca de tus palabras. Si el termo o la botella aparece tarde, regenera o acepta el desfase si el video se entiende bien.
- La voz: que diga bien JUTSU y la palabra guía.
- La etiqueta: marca el video como hecho con IA donde la plataforma lo permita.

9. Prompts bonus

Los bonus 01, 02 y 05 los probé. El 03 y el 04 son variaciones que no he probado, así que haz primero una prueba corta en 720p. El bonus 05 sube las imágenes en este orden: @image1 chica 1 sin botella, @image2 chica 1 con botella blanca, @image3 chica 2 con botella rosa, @image4 chica 3 con termo lila, @image5 chica 4 con botella menta, @image6 chica 4 sin botella, y @video1 tu clip. En mi prueba la botella blanca apareció a los 4.5 s justo cuando se levantó mi mano y la plaza ya no se coló, pero los cortes de rosa, lila y menta salieron más tarde de lo pedido (5.8, 7.5 y 9.3 s) y la chica 4 siguió con la botella en la mano al final.

Bonus 01 (Una sola persona con tu producto en todo el video (probado)):

"Use @video1 as the master video. Keep its exact motion, gestures, lip movements, timing and handheld camera movement. Replace the man with the woman from @image1: same face, hairstyle and skin tone in every frame. She wears the outfit from @image1 and keeps the small clip-on lavalier microphone on her top. In her raised open hand she holds the product from @image1, clearly visible, plain with no label. Background: the room from @image1, same layout and lighting, for the whole clip. Relight her to match the room. Photorealistic handheld smartphone look, no flicker, no extra people, no text, no logos, no cuts."

Bonus 02 (Quitarle el producto a una persona que ya generaste (probado)):

"Edit the reference image. Keep the exact same woman (identical face, hair, skin), the same outfit with the small clip-on microphone, the same room and the same soft light. Change only this: her hands are empty and relaxed, with no bottle and no objects anywhere. Photorealistic, 16:9, no text, no logos."

Bonus 03 (Lo mismo pero con un hombre, sin cambio de voz (sin probar)):

"Use @video1 as the master video. Keep its exact motion, gestures, lip movements, timing and handheld camera movement. Replace the man with the man from @image1: same face, hairstyle and skin tone in every frame. He wears the outfit from @image1 and keeps the small clip-on lavalier microphone on his top. In his raised open hand he holds the product from @image1, clearly visible, plain with no label. Background: the room from @image1, same layout and lighting, for the whole clip. Relight him to match the room. Photorealistic handheld smartphone look, no flicker, no extra people, no text, no logos, no cuts."

9. Prompts bonus (continuación)

Bonus 04 (Evitar que se cuele el fondo original en los primeros segundos (sin probar)):

> *"Add this sentence to the end of Prompt 06: The original background of @video1 must not be visible in any frame, including the first seconds: every part uses only the room from its own image."*

Bonus 05 (Cortes alineados a tu mano, 6 imágenes: una sin botella y otra con botella de la chica 1 y de la chica 4 (probado)):

> *"Use @video1 as the master video. Keep its exact motion, gestures, lip movements, timing and handheld camera movement. Replace the man with a different woman in each part, same framing and scale as @video1. The original street, plaza, statue and buildings of @video1 must never appear in any frame: every part uses only the room of its own image. Hard cuts only, no morphing. Each woman keeps the small clip-on lavalier microphone on her top. 0 to 4.5s: the woman from @image1 in the living room from @image1, her hands empty with no bottle. 4.5 to 5.6s: hard cut, the same woman from @image2 in the same living room, holding the white bottle from @image2 in her raised open hand, clearly visible. 5.6 to 6.8s: hard cut, the woman from @image3 in the room from @image3, holding the pink bottle from @image3 in her raised open hand. 6.8 to 7.9s: hard cut, the woman from @image4 in the room from @image4, holding the lavender tumbler with a straw from @image4 in her raised open hand. 7.9 to 8.8s: hard cut, the woman from @image5 in the kitchen from @image5, holding the mint bottle from @image5 in her raised open hand. 8.8s to the end: hard cut, the same woman from @image6 in the same kitchen from @image6, her hands empty with no bottle and no objects. Relight each woman to match her room. Photorealistic handheld smartphone look, no flicker, no extra people, no text, no logos."*

10. La parte honesta

Este formato no es mágico y tampoco es instantáneo. El video de 15.5 segundos tardó unos 8 minutos en generarse, las imágenes unos 2 minutos y la voz menos de uno. Eso son unos 10 minutos de espera, y después viene el montaje en CapCut. Por eso el decir anuncios en menos de 10 minutos es cierto para la generación, no para todo el proceso.

Mi primer intento fue cortar el clip en 6 pedazos y generar uno por chica. Falló: cada pedazo salió más largo que el original (4 segundos en vez de 3), el movimiento ya no coincidía con el audio y uno de los intentos falló sin explicación. Con un solo prompt y las 4 imágenes funcionó mucho mejor.

Genjutsu no respeta los tiempos del prompt al segundo. En mi video el termo lila duró de 6 a 9 segundos y la botella menta apareció de 9 a 11. Para controlar mejor los cortes, deja 3 segundos por persona al grabar.

Cada generación gasta créditos y los míos no incluyen las pruebas que hice antes del resultado final. Revisa el costo en Higgsfield antes de generar.

Esta guía es parte de una colaboración con Higgsfield.

Las personas, las voces y los cuartos de este video no son reales. Si los usas en un anuncio, no los presentes como clientes o personas reales: etiqueta el video como hecho con IA y revisa las reglas de anuncios de la plataforma, porque pueden pedirte que lo indiques. No uses fotos de personas reales sin su permiso y evita marcas y logos en los productos.

11. Cómo usarlo

Graba tu clip. Genera las cuatro personas con tu producto con el Prompt 01. Haz una prueba en 720p con el Prompt 06 y las imágenes en orden. Revisa la sección 8. Pega tu audio al video, pásalo por el cambio de voz y arma el reel en CapCut. Para variar, usa los prompts bonus uno por uno.

12. Pruébalo en Higgsfield

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