



MAKE THE CAP POP OUT.

A step-by-step image-to-video workflow

THE ROUTE

Prepare two references → generate the oversized-phone editorial base → edit the on-screen post → replace the shoe with your cap → update the post text → create the pop-out frame → animate two clips → join in CapCut.

AUTHENTICITY / NON-PROMPT NOTE

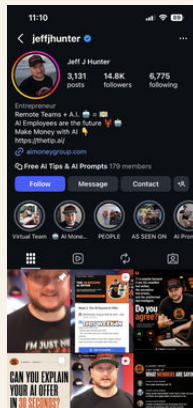
The verified New Balance page depicted in the source prompts is a fictional composite, not a real endorsement or authentic post. Label the result as AI-generated; do not pass it off as a genuine New Balance post.

01 / Prepare two references

The source calls for a close-up portrait with your face clearly visible as Reference 1 and a screenshot of your Instagram home page as Reference 2.



REF 1 / FACE



REF 2 / SCREEN



02 / Generate the editorial base

In Nano Banana Pro, upload Reference 1 first and Reference 2 second. Paste the original JSON prompt below without changing it.

COPY PROMPT 1 / EDITORIAL BASE · JSON

```
{
  "task": "high-end photorealistic composite image generation",
  "model_intent": "reference-driven scene reconstruction with dominant oversized object emphasis",
  "references": {
    "reference_1": {
      "role": "primary subject identity",
      "description": "Use this image as the sole identity source. Extract facial structure, body proportions, skin texture, hair, and overall likeness from this reference. Do not stylize or alter identity. Replace the person in the scene with the identity from this reference."
    },
    "reference_2": {
      "role": "screen UI layout reference",
      "description": "Use this image strictly as the phone screen content. Recreate the UI layout, spacing, scale, and proportions exactly as seen."
    }
  },
  "scene_description": {
    "environment": "stylish Italian street with strong fashion atmosphere, realistic Rome street with textured stone walls and cobblestone ground",
    "background_style": "high-end editorial street fashion setting, elegant European architecture",
    "composition": "the subject stands still in the middle of the street, facing the camera, holding an extremely oversized smartphone vertically, grounded on the pavement"
  },
  "subject_handling": {
    "identity_lock": true,
    "pose": "static confident stance, body facing forward, head aligned with camera, direct eye contact with the lens",
    "expression": "calm, self-assured, fashion-editorial seriousness",
    "interaction": "the subject holds the oversized phone with both hands while it rests firmly on the ground, as if the phone has real weight and mass",
    "scale_relationship": "the phone is intentionally massive, reaching from the ground up to the subject's chest or shoulder level, nearly comparable to the subject's height"
  },
  "wardrobe_styling": {
    "outfit_source": "Kith official website",
    "style": "modern premium streetwear",
    "outfit_description": "a realistic Kith outfit combination including top, bottom, footwear, and accessories inspired by current Kith lookbook aesthetics",
    "branding_rules": "do not show or mention any brand names or logos on clothing; branding is allowed only on footwear if present",
    "fabric_detail": "high-resolution textile texture, natural folds, realistic fabric weight and stitching"
  },
  "phone_handling": {
```



COPY PROMPT 1 / EDITORIAL BASE · JSON · CONTINUED

```
"device": "iPhone 17 Pro",
"scale_priority": "extreme oversized scale",
"physical_size_description": "the smartphone is gigantic, floor-standing, almost as tall as the subject's upper body, clearly larger than any handheld device",
"position": "standing vertically on the pavement, its bottom edge firmly touching the ground",
"interaction_detail": "the subject's hands grip the phone naturally from the sides and top edge, stabilizing it like a heavy physical object",
"orientation": "fully frontal and vertical, screen perfectly visible and readable to the camera",
"material_detail": "premium glass surface and metallic edges with realistic reflections, visible thickness, realistic weight perception"
},
"phone_screen_content": {
  "screen_state": "on and active",
  "content_mapping": "place the second reference image exactly onto the phone screen without resizing or cropping",
  "ui_realism": "accurate screen brightness, glass reflections, subtle glare, and realistic contrast",
  "perspective_alignment": "screen content aligns precisely with the phone's frontal orientation"
},
"lighting": {
  "key_light": "natural daylight with soft directional light",
  "ambient_light": "gentle city bounce light from surrounding buildings",
  "shadow_behavior": "the oversized phone casts a strong, grounded shadow onto the cobblestone pavement and partially onto the subject's legs"
},
"camera": {
  "angle": "frontal view at human eye level",
  "lens": "35mm street photography look",
  "framing_priority": "the frame centers both the subject and the oversized phone as equal visual anchors",
  "depth_of_field": "moderate depth of field with subject and phone sharply in focus, background softly separated",
  "realism": "photographic realism with editorial fashion tone"
},
"quality_constraints": {
  "photorealism": "ultra realistic",
  "resolution": "high resolution",
  "no_artifacts": true,
  "no_cartoon": true,
  "no_anime": true,
  "no_surreal_distortion": true
},
"final_instruction": "The final image must look like a real high-end fashion editorial photograph shot on an Italian street, where the subject faces the camera directly and stabilizes a massive, floor-standing smartphone that feels heavy, real, and physically grounded, blending luxury fashion with technology in a bold yet believable way."
}
```



03 / Build the on-screen post

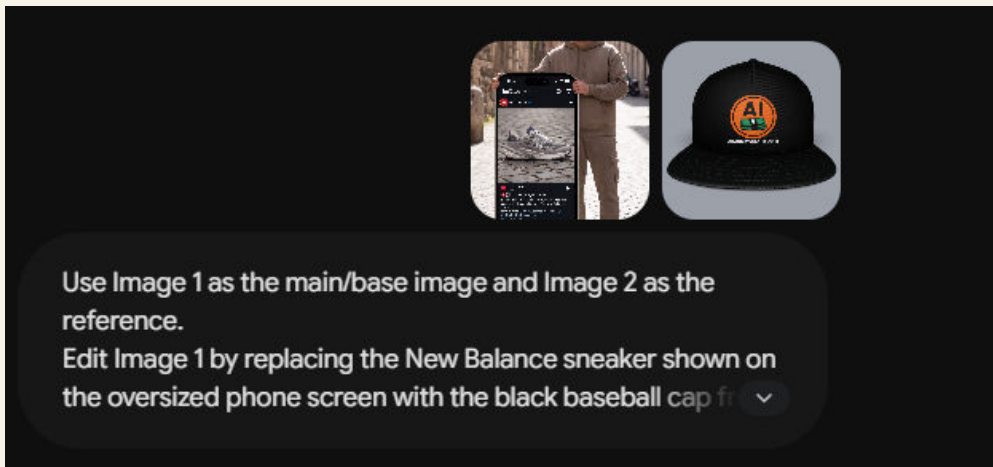
Use the first generated image as reference. Run the second source prompt to create the sneaker post.

COPY PROMPT 2 / SECOND PROMPT

"Make the Instagram on the screen a real, natural post from the blue tick New Balance page, with lots of likes and comments, and in the photo an amazing New Balance sneaker stands sideways, an incredible product photo."

WHAT COMES NEXT

For the cap replacement, upload two images: the result from this second prompt as Image 1 and your hat/product reference as Image 2. The source screenshot shows this two-image setup.



04 / Swap in the cap

First pass: replace the shoe on the phone screen with the cap while preserving the original street and phone composition.

COPY PROMPT 3 / HAT PRODUCT REPLACEMENT

"The cap on the phone screen should clearly match the reference hat from Image 2
Keep the hat black, with the orange "AI" logo and "AIMONEYGROUP.COM" text visible
Place the hat in the same area where the shoe appears, as a clean and realistic product photo
Make it look like a natural Instagram post from the verified New Balance page
Keep the Instagram interface realistic, with lots of likes and comments
Preserve the man, the street background, the oversized phone, and the overall composition
Maintain realistic lighting, perspective, shadows, and proportions
Make the final image look polished, natural, and high quality
The result should look like the oversized phone is displaying a real Instagram post where the featured product is the black AI Money Group cap instead of the sneaker. "



04 / Update the post text

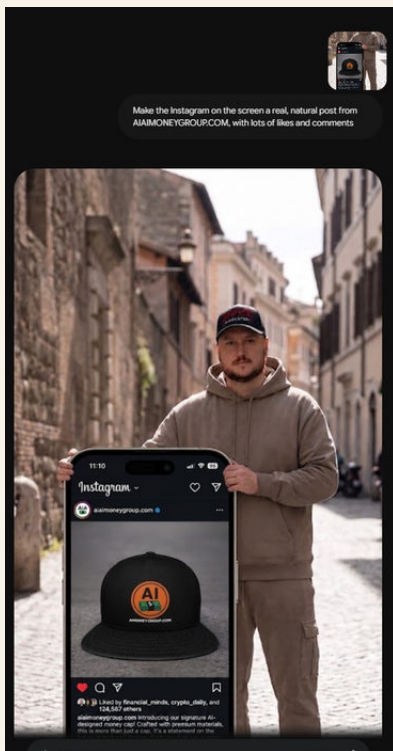
Second pass: after the product swap, update the on-screen account/content with the source prompt below. This is separate from the first pass.

COPY PROMPT 4 / POST CONTENT REPLACEMENT

"Make the Instagram on the screen a real, natural post from AIAIMONEYGROUP.COM, with lots of likes and comments"

SOURCE FIDELITY

The unusual account spelling in the prompt is reproduced exactly from the supplied document, not corrected here.



SOURCE SCREEN EXAMPLE



SOURCE COMPOSITE EXAMPLE



05 / Make the cap pop out

Use the second generated image as the reference for the third source prompt.

COPY PROMPT 5 / THIRD PROMPT · 3D POP-OUT

"A hyper-realistic 3D pop-out effect of a black baseball cap with an orange "AI AIMONEYGROUP.COM" logo. The hat is floating stylishly in mid-air in the exact center of the frame, angled slightly and tilted towards the camera. The hat is in razor-sharp macro focus, revealing highly detailed, authentic fabric textures. The background is heavily blurred with a shallow depth of field. The blurred background shows a man holding a giant smartphone displaying an Instagram layout with a plain white image background. Dramatic, strong cinematic backlight illuminates the hat from behind, highlighting organic dust particles floating in the air. Ultra-high resolution, photorealistic, professional editorial photography style, completely realistic with no cartoonish or cheap graphical elements."



06 / Animate two Start / End clips

Kling 3.0: pair the first generated frame with the changed/scrollled-screen frame for Clip A. Then pair the second frame with the pop-out frame for Clip B. The source also mentions Veo 3 as an option.

CLIP A · SCREEN TRANSITION

First Image = Start. Second Image = End. Paste the first Kling prompt without changes.

COPY PROMPT 6 / KLING CLIP A

"Handheld shaky camera, the man grabs and swipes up the phone's touchscreen with his hand and the image on the screen starts changing very fast and after sliding in a blurry way for a long time it slows down and the new image slowly slides and settles into place. add real time motion, add camera shake"

CLIP B · CAP EMERGENCE

Second Image = Start. Third Image = End. Paste the second Kling prompt without changes.

COPY PROMPT 7 / KLING CLIP B

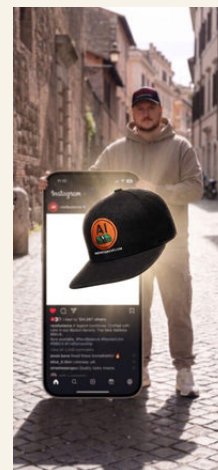
"Handheld shaky camera, The black baseball cap with an orange "AI AIMONEYGROUP.COM" logo on the screen flies out through the screen glass and glides toward the exact center of the frame with very soft, gently swaying movements. add real time motion, add camera shake"



START



SCROLLED SCREEN



CAP POP-OUT

07 / Join in CapCut

Place the two videos back-to-back in order: screen transition first, cap pop-out second. Review the transition and export.



THE NEXT STEP

You made one striking asset. Now build a marketing system.

WHAT YOU JUST BUILT

You turned a few references into a short creative video. That is a useful starting point. Save your source images, prompts and final clips so you can repeat the process with your own product or client brief.

ONE VIDEO IS NOT A MARKETING PLAN.

The next job is choosing who this video is for, what it should say, where it belongs in a campaign and what happens after someone sees it. That is the difference between playing with an AI tool and building a repeatable marketing machine for your business.

BUILD WITH JEFF INSIDE AI MONEY GROUP

Turn AI experiments into content, campaigns and useful systems.

Get training and support while you build assets you can actually use.

Take this hat video beyond the edit and give it a job in your marketing.

YES, I WANT TO BUILD AI MARKETING THAT WORKS

You can finish this project on your own. If you want to turn the next creative idea into a repeatable business asset, come build it with me.

Jeff J Hunter / AI Money Group