

Từ 1 bức ảnh ra video AI biết nói

7 bước + 4 câu lệnh dán sẵn · Anh Thành AI

TỪ 1 BỨC ẢNH RA VIDEO AI BIẾT NÓI — 7 BƯỚC

Làm lần đầu mất khoảng 45 phút. Từ lần thứ hai chỉ còn 10 phút một video.

Cần có: một ảnh chân dung của anh chị (hoặc của nhân vật đại diện) + một đoạn lời thoại 20-30 giây. Không cần máy quay, không cần phòng thu.

BƯỚC 1 · CHỌN ĐÚNG TẤM ẢNH

Ảnh sai là hỏng cả video, không công cụ nào cứu được. Tấm ảnh dùng được phải:

- Nhìn thẳng máy, thấy rõ cả hai mắt, miệng không bị che
- Đủ sáng, ánh sáng từ phía trước, mặt không bị bóng đổ chia đôi
- Không đeo kính râm, không khẩu trang, tóc không phủ lên môi
- Vai và cổ nằm trong khung, không cắt ngang cằm
- Ảnh rõ nét — ảnh chụp màn hình hoặc ảnh cũ mờ thì nên chụp lại

BƯỚC 2 · VIẾT LỜI THOẠI TRƯỚC, LÀM HÌNH SAU

Dán **câu lệnh 1** ở trang sau vào ChatGPT hoặc Gemini. Đọc thử thành tiếng, chỗ nào lú lẫm thì sửa cho ngắn lại.

BƯỚC 3 · DỌN TẤM ẢNH CHO "CHUẨN QUAY"

Dán **câu lệnh 2** vào phần tạo ảnh của Gemini kèm ảnh gốc: dọn nền, đều ánh sáng, giữ nguyên gương mặt.

BƯỚC 4 · TẠO TIẾNG

Hai đường:

- Giọng thật của anh chị* — thu bằng điện thoại trong phòng kín, hơn hẳn giọng máy khi bán hàng, khách nghe là tin.
- Giọng AI* — dùng ElevenLabs hoặc công cụ lồng tiếng tương tự. Chọn giọng vùng miền hợp khách, chỉnh tốc độ chậm lại một chút.

BƯỚC 5 · GHÉP MẶT NÓI THEO TIẾNG

Đưa ảnh ở bước 3 + file tiếng ở bước 4 vào công cụ lip-sync (Hedra hoặc tương đương) → xuất ra video mặt nói. Xem lại: môi có khớp tiếng không, mắt có chớp tự nhiên không. Lệch nhiều thì thường do tiếng bị ồn hoặc ảnh nghiêng.

BƯỚC 6 · CHO NHÂN VẬT CHUYỂN ĐỘNG (không bắt buộc)

Muốn cảnh sinh động hơn thì dựng thêm 3-5 giây từ ảnh bằng công cụ ảnh-ra-video (Veo, Kling, Hailuo) với **câu lệnh 3**, rồi ghép vào đầu video.

BƯỚC 7 · DỰNG THÀNH VIDEO ĐĂNG ĐƯỢC

Thêm phụ đề (quá nửa người xem để tắt tiếng), cắt gọn 3 giây đầu, kết bằng một câu mời nhấn tin. Dùng **câu lệnh 4** để từ một kịch bản ra nhiều biến thể — đó mới là chỗ "một bức ảnh ra hàng loạt video".

4 CÂU LỆNH DÁN SẴN

Chép nguyên câu, chỉ đổi chữ trong ngoặc vuông.

1 · LỜI THOẠI 30 GIÂY

"Viết lời thoại 30 giây cho video tôi tự nói về [sản phẩm], bán cho [khách: tuổi, nghề, nỗi lo], giá [giá]. Câu ngắn, nói như đang nói chuyện với khách quen, không dùng từ 'siêu', 'cực kỳ'. Mở đầu bằng một câu khiến người đang lướt dừng lại, kết bằng một câu mời nhắn tin. Chỉ trả về lời thoại, không mô tả cảnh."

2 · ẢNH CHÂN DUNG CHUẨN QUAY

"Giữ nguyên 100% gương mặt, màu da, kiểu tóc và trang phục của người trong ảnh. Chỉ dọn nền thành [nền xám trơn / góc làm việc gọn gàng], làm ánh sáng đều từ phía trước, bỏ bóng đổ trên mặt. Khung hình từ ngực trở lên, nhìn thẳng máy, ảnh dọc 9:16. Không làm mịn da quá tay, giữ nét chân thực."

3 · ẢNH THÀNH CLIP CHUYỂN ĐỘNG

"Từ ảnh này, tạo video 5 giây: nhân vật hơi nghiêng đầu, chớp mắt tự nhiên và mỉm cười nhẹ, vai chuyển động rất ít. Máy quay đứng yên, không zoom. Giữ nguyên gương mặt, trang phục và ánh sáng."

4 · MỘT KỊCH BẢN RA MƯƠI VIDEO

"Đây là kịch bản gốc: [dán lời thoại]. Viết lại thành 10 bản khác nhau, mỗi bản đổi góc tiếp cận: kể một tình huống của khách, chỉ ra một sai lầm hay gặp, so sánh trước-sau, trả lời một câu khách hay hỏi, nói về một chi tiết nhỏ của sản phẩm. Giữ nguyên độ dài và giọng văn. Đánh số từng bản."

3 điều nên nhớ

1. **Hình có thể dựng, lời thì không được bịa.** Công dụng, giá, bảo hành nói sai là mất khách vĩnh viễn, không AI nào gỡ lại được.
2. **Nên ghi một dòng "video có sử dụng AI"** khi anh chị dựng lại hình ảnh của chính mình. Người xem biết trước thì không thấy bị lừa.
3. **Không lấy ảnh và giọng người khác.** Tuyệt đối không dựng mặt người nổi tiếng — đó là chỗ mất tài khoản và dính kiện.

Từng làm chủ salon tóc, ảnh viện áo cưới, xưởng in — nay đưa AI vào cửa hàng Việt.

Tôi là Anh Thành AI. Chúc anh chị ra đơn. · Zalo 0818122339 · hoangthanh.pages.dev