

Criação de Influenciadores Virtuais com Inteligência Artificial

Marcelo Corrêa dos Santos¹; Cecília Alves Corrêa²; Arthur Nascimento Crosara³; Mateus de Sousa Valente⁴

^{1, 2, 3, 4} Universidade de Uberaba

E-mail: marceloads624@gmail.com; mateus.valente@uniube.br

Resumo

Este artigo aborda a criação de influenciadores virtuais por meio de Inteligência Artificial Generativa, buscando consistência facial e preservação de identidade da persona. O objetivo é demonstrar a estruturação de um pipeline criado no Obsidian que utiliza distintos Grandes Modelos de Linguagem (LLMs), responsáveis pela geração dos prompts em formato JSON, e modelos de difusão, os quais são responsáveis pela geração das imagens e vídeos. A metodologia possui uma abordagem iterativa e empírica: imagens de referência de poses, vestimentas e cenários são processadas por um LLM, que devolve parâmetros estruturados; esses parâmetros, somados a uma colagem de identidade que entrega mais precisão aos traços da influenciadora, alimentam os modelos de geração de imagem. Em seguida, podemos selecionar as melhores imagens e convertê-las em vídeos por modelos de image-to-video, ou realizar alguns ajustes de câmera em um workflow importado que roda via ComfyUI. Como resultado, o pipeline indica uma metodologia operacional e capaz de gerar uma persona em diferentes contextos sem perder seus traços originais, sendo validado por uma influenciadora base criada como prova de conceito. Entretanto, é relevante destacar que os modelos ainda transferem características indesejadas das imagens de referência, o que exige certa curadoria manual e dificulta uma automação total. Conclui-se que a integração de ferramentas acessíveis permite a atuação no mercado de influencers digitais sem exposição pessoal, democratizando a produção de conteúdo e oferecendo às marcas uma alternativa viável para divulgação de produtos e serviços.

Palavras-chave: preservação de identidade; influencer digital; engenharia de prompts; integração de ferramentas; persona virtual.

1 Introdução

O avanço das redes sociais transformou a figura do influenciador digital em um dos principais canais de marketing e comunicação da atualidade. Plataformas como Instagram e TikTok possibilitam que criadores de conteúdo construam audiências expressivas e estabeleçam parcerias comerciais com marcas de diversos segmentos. Paralelamente, o surgimento de influenciadores virtuais (personagens gerados por computação gráfica ou inteligência artificial) tem ganhado relevância no mercado, com casos como a Lu do Magalu, que possui mais de 6 milhões de seguidores e movimenta valores significativos por publicação (Sorosrungruang et al., 2024).

No entanto, muitas pessoas que desejam atuar como criadores de conteúdo enfrentam barreiras como a necessidade de exposição pessoal, timidez ou falta de recursos para produção de imagens profissionais. Da mesma forma, pequenas marcas e empreendedores que buscam promover seus produtos por meio de influenciadores muitas vezes não dispõem de orçamento para contratar criadores de conteúdo humanos. Além disso, a criação de personagens virtuais com aparência consistente ainda é vista como um processo complexo, que geralmente demanda conhecimentos técnicos específicos e

recursos computacionais significativos, dificultando o acesso de pessoas comuns a essa alternativa.

Diante desse cenário, o objetivo deste trabalho é demonstrar a utilização de uma esteira de ferramentas acessíveis de IA Generativa para a criação de personagens virtuais com consistência facial, utilizando técnicas de preservação de identidade, visando possibilitar a atuação no mercado de influência digital sem a necessidade de exposição pessoal.

A realização deste projeto se justifica pelo crescimento expressivo do mercado de influenciadores virtuais, avaliado em US\$ 4,6 bilhões em 2024, com estimativa de crescimento de 38,9%. Esse cenário demonstra uma demanda real e crescente por soluções nessa área. Contudo, a exigência de exposição pessoal ainda exclui uma parcela significativa de pessoas que possuem potencial criativo, mas não desejam expor sua imagem. Nesse contexto, a criação de influenciadores virtuais por meio de ferramentas acessíveis representa uma alternativa viável, permitindo que qualquer pessoa produza conteúdo visual profissional e consistente sem custos elevados e sem a necessidade de aparecer.

2 Referencial Teórico

O presente referencial teórico aborda as bases tecnológicas e mercadológicas que fundamentam o desenvolvimento de influenciadores virtuais. A discussão está estruturada em três eixos principais: o uso de Inteligência Artificial Generativa em pipelines multimodais, os impactos do fotorrealismo na credibilidade das personas virtuais e os mecanismos algorítmicos que regem o engajamento nas redes sociais.

2.1 Inteligência Artificial Generativa e Pipelines Multimodais

A Inteligência Artificial Generativa tem evoluído de modelos isolados para pipelines complexos e multimodais, capazes de integrar texto, imagem e vídeo com alto grau de precisão. Apesar do avanço significativo dos modelos de difusão de texto para imagem, essas arquiteturas frequentemente encontram dificuldades em seguir instruções espaciais estritas.

Segundo Lian et al. (2024), “a adoção de Grandes Modelos de Linguagem como planejadores de layout resolve a incapacidade dos modelos de difusão de seguir instruções espaciais complexas, garantindo o posicionamento exato dos elementos na imagem”. Dessa forma, os Grandes Modelos de Linguagem (LLMs) assumem o papel de raciocinar sobre a estruturação da cena, gerando parâmetros organizados que orientam a geração visual.

A compreensão dessa dinâmica é fundamental para o presente projeto, pois fundamenta o fluxo de trabalho técnico adotado: delegar a extração de características e a engenharia de prompts a um LLM em formato JSON garante o rigor necessário para manter a consistência visual, o enquadramento e a preservação de identidade da influenciadora virtual em múltiplos cenários e vídeos.

2.2 Influenciadores Virtuais: Fotorrealismo e Credibilidade

A consolidação do mercado de marketing de influência digital abriu espaço para o surgimento de personas não humanas, impulsionadas pelo barateamento e acessibilidade das ferramentas de criação visual. Marcas de diversos setores, como moda e entretenimento, “recorreram a influencers virtuais como solução para suas necessidades

de marketing de influência” (Sorosrungruang et al., 2024), evidenciando a viabilidade comercial desses personagens.

No entanto, a eficácia dessas personas depende diretamente do seu nível de realismo e da sua capacidade de gerar conexões autênticas. Conforme apontam Jayasingh, Sivakumar e Vanathaiyan (2025), “um dos principais desafios dos influencers de IA é a falta de emoções genuínas e autenticidade, o que pode comprometer a percepção de credibilidade e afetar o engajamento”. Isso significa que a perfeição estética não é suficiente; a ausência de expressões naturais pode levar a persona ao “Vale da Estranheza” (Uncanny Valley), causando repulsa no público.

Essa discussão teórica tem impacto direto no desenvolvimento deste trabalho, justificando a adoção de técnicas rigorosas de lapidação de identidade e controle de expressões nos modelos de difusão, assegurando que a influenciadora gerada possua tanto fotorrealismo quanto credibilidade narrativa.

2.3 Engajamento e Mecanismos Algorítmicos nas Redes Sociais

O sucesso de qualquer criador de conteúdo, seja ele humano ou virtual, está intrinsecamente ligado à sua capacidade de navegar pelas regras dos algoritmos de recomendação das mídias digitais. Esses algoritmos são desenhados para reter a atenção do usuário explorando respostas psicológicas rápidas.

De acordo com Metzler e Garcia (2024), “as métricas de engajamento promovem principalmente conteúdos que se encaixam nas preferências afetivas e cognitivas imediatas dos usuários, em vez de conteúdo de qualidade”. Em complemento a essa dinâmica, o estudo de Yu et al. (2024) comprova que “a exibição de felicidade, tristeza, nojo e surpresa são significativas no acionamento do engajamento dos usuários ao promover produtos com conteúdo visualmente atraente”. Portanto, a viralização de uma persona digital exige uma engenharia de conteúdo baseada em estímulos afetivos.

Essa compreensão embasa teoricamente as escolhas práticas deste projeto, validando o esforço de utilizar ferramentas de animação e transferência de movimento para adicionar ações, falas e respostas emocionais aos frames estáticos, adaptando o conteúdo da influenciadora virtual aos gatilhos de engajamento priorizados pelas plataformas atuais.

3 Metodologia e Descrição do Sistema

3.1 Metodologia de Desenvolvimento

O desenvolvimento deste projeto fundamenta-se em uma abordagem iterativa, incremental e empírica. Diferentemente de projetos de software tradicionais, a criação de personas virtuais por IA exige ciclos contínuos de testes práticos, avaliação de resultados e refinamento técnico (engenharia de prompts).

Utiliza-se a ferramenta Obsidian como repositório de documentação ágil e para a modelagem visual do fluxograma de processos. Como os modelos generativos podem apresentar variações não intencionais. Por exemplo, herdar características indesejadas de uma imagem de referência, como piercings, alterações capilares ou discrepâncias de cenário. O processo exige curadoria ativa. A cada iteração, as saídas estruturadas em JSON são corrigidas e reprocessadas até a obtenção do fotorrealismo e da consistência de identidade ideais.

3.2 Tecnologias Utilizadas

A solução emprega um ecossistema de ferramentas multimodais. O ChatGPT atua como o Grande Modelo de Linguagem (LLM) responsável pela engenharia reversa de

imagens de referência, extraindo características visuais e devolvendo-as estruturadas no formato JSON.

Para a geração de imagens estáticas e a lapidação da identidade facial, utiliza-se a plataforma Google Flow integrada ao modelo generativo Nano Banana Pro. Na etapa de animação e síntese audiovisual, empregam-se os modelos Veo 3 (disponível no Google Flow) e Kling 2.6, encarregados de aplicar transferência de movimento e transformar frames estáticos em vídeos dinâmicos.

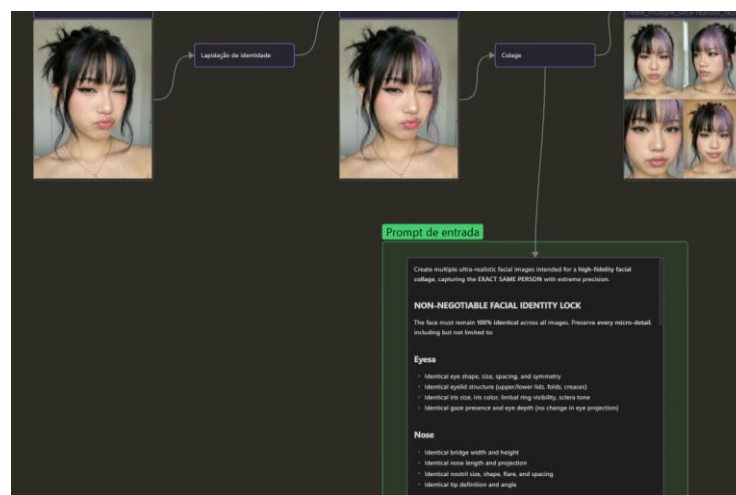
Complementarmente, o ComfyUI é utilizado de forma secundária para rotinas específicas de controle espacial, como manipulação de enquadramentos (360 graus) e execução de workflows locais de motion control.

3.3 Modelagem do Sistema

A arquitetura do projeto não segue padrões tradicionais como MVC, mas organiza-se como um Pipeline de Processamento Multimodal Generativo, estruturado em quatro etapas sequenciais:

- Geração de Identidade Base: criação de uma “colagem” de referência (identity lock) contendo os traços biométricos imutáveis da influenciadora.

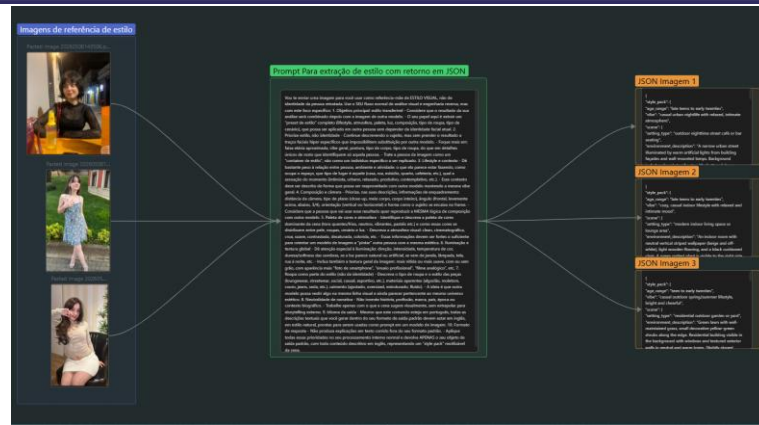
Figura 1 - Criação de colagem



Fonte: Elaborada pelo autor (2026)

- Extração e Estruturação de Dados: imagens de referência (poses, vestimentas, cenários) são processadas pelo LLM, que traduz o contexto visual em parâmetros textuais rigorosos (JSON).

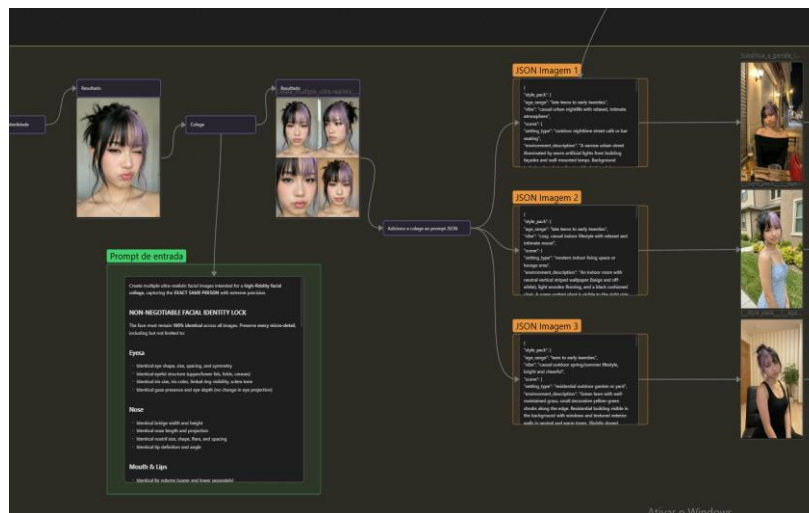
Figura 2 - Extração de contexto



Fonte: Elaborada pelo autor (2026)

- Geração Contextual: o modelo de difusão de imagem (Google Flow / Nano Banana Pro) recebe a colagem de identidade e o JSON de contexto, fundindo a face da persona com novos ambientes, sem perder os traços originais.

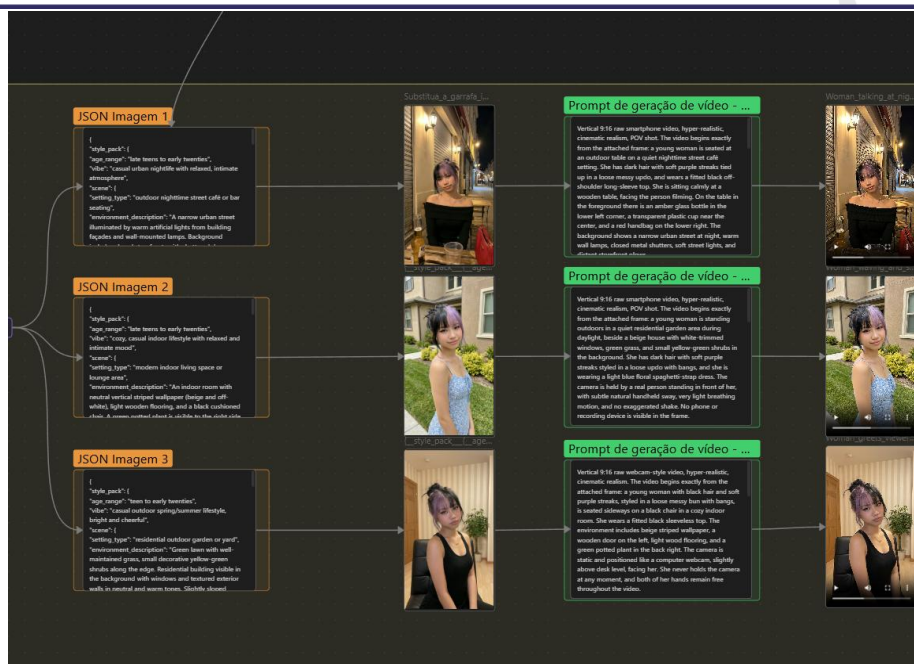
Figura 3 - Criação de imagem contextualizada



Fonte: Elaborada pelo autor (2026)

- Animação e Síntese de Vídeo: o frame estático finalizado atua como ponto de partida (image-to-video) em modelos como Veo 3 ou Kling 2.6, nos quais prompts de ação ditam o movimento e o comportamento da influenciadora.

Figura 4 - Geração image-to-video



Fonte: Elaborada pelo autor (2026)

3.4 Desenvolvimento Inicial da Aplicação

Até o presente momento, o pipeline central de geração encontra-se validado e operacional. Como prova de conceito (PoC), foi gerado o modelo base de uma influenciadora virtual, demonstrando na prática a viabilidade da preservação de identidade.

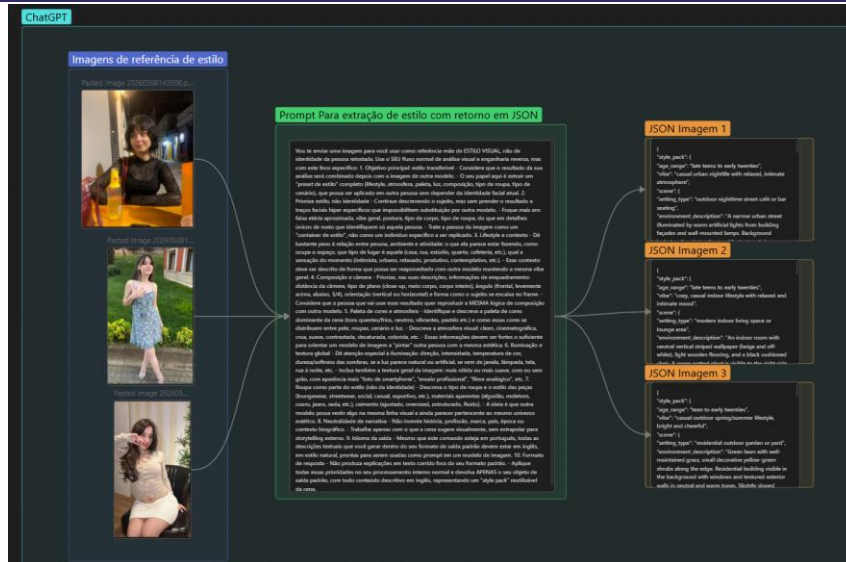
Os templates de prompts em JSON e as rotinas de extração visual via ChatGPT foram padronizados. Além disso, foram executados com sucesso os primeiros testes práticos de aplicação de lock facial em diferentes cenários simulados, bem como as validações iniciais de transferência de movimento audiovisual utilizando os modelos Veo 3 e Kling 2.6, e o ajuste tridimensional de câmera por meio do ComfyUI. A influenciadora já desenvolvida atua como o principal produto validado e modelo de demonstração desta esteira de produção integrada.

4 Resultados e Discussão

4.1 Funcionalidades Implementadas e Funcionamento do Sistema

O pipeline multimodal generativo proposto encontra-se plenamente operacional, estruturado em um fluxo contínuo que integra modelos de linguagem e de difusão. Até o momento, foram implementadas as funcionalidades de extração de parâmetros visuais estruturados (JSON) via ChatGPT e a consolidação do identity lock (colagem de identidade) da influenciadora virtual. O funcionamento do sistema inicia-se com a inserção de imagens de referência de contexto (cenários, poses, vestuário), das quais o LLM realiza a engenharia reversa para gerar o respectivo JSON. Esses dados, em conjunto com a colagem de identidade, alimentam as plataformas Google Flow e Nano Banana Pro, que renderizam a persona em novos ambientes mantendo os traços biométricos intactos.

Figura 5 - Processo de engenharia reversa



Fonte: Elaborada pelo autor (2026)

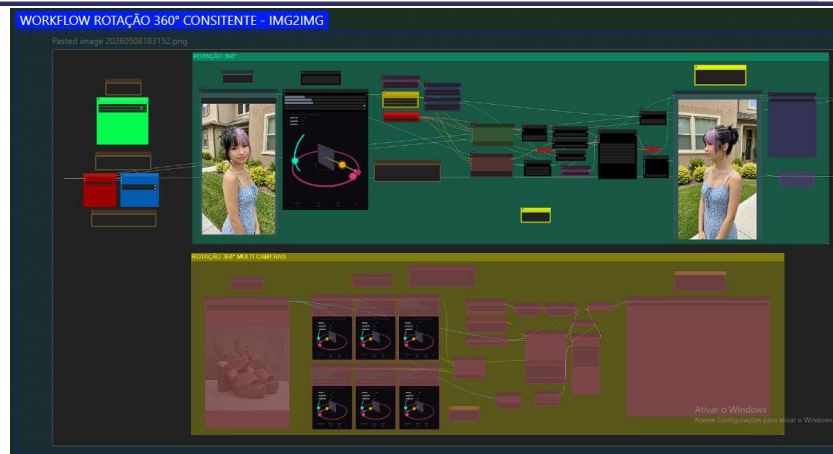
Figura 6 - Resultados das imagens



Fonte: Elaborada pelo autor (2026)

Em seguida, para gerar dinamismo e engajamento, os frames estáticos validados são exportados para os modelos Veo 3 e Kling 2.6. Nessas ferramentas de image-to-video, aplica-se a transferência de movimento e o controle de ações, permitindo que a influenciadora realize movimentos, expressões ou interações com produtos. O ComfyUI é acionado para realizar ajustes espaciais pontuais, como o rotacionamento de câmera em 360 graus, entregando os ativos finais para publicação nas redes sociais.

Figura 7 - Workflow Rotação 360° no ComfyUI



Fonte: Elaborada pelo autor (2026)

4.2 Discussão dos Resultados Obtidos e Limitações Técnicas

A análise das amostras geradas atesta o êxito da solução no que tange ao fotorrealismo e à preservação de identidade. A utilização de JSON para engessar os parâmetros visuais evitou variações extremas na fisionomia da personagem, enquanto o uso do Kling 2.6 proporcionou naturalidade nas exibições emocionais animadas. A principal vantagem desta abordagem reside no controle granular sobre a cena, algo que ferramentas automatizadas mais simples não permitem.

Entretanto, o desenvolvimento iterativo também evidenciou limitações inerentes aos modelos generativos atuais. O sistema está sujeito a “alucinações” pontuais de herança de características. Durante os testes, notou-se que atributos exclusivos das imagens de referência — como a presença de piercings, tatuagens específicas ou anomalias no cenário — frequentemente eram transferidos de forma indesejada para a influenciadora. Por conta desse gargalo técnico, a automação total do processo não é viável no momento, sendo indispensável a curadoria manual intensa, que envolve processos contínuos de tentativa e erro, correções de prompt e sucessivas novas gerações para extrair o ativo visual em nível profissional.

4.3 Relação com o Problema de Pesquisa

Apesar da exigência de curadoria ativa, a aplicação desenvolvida responde satisfatoriamente ao problema central levantado no início da pesquisa. A estruturação do pipeline multimodal generativo viabiliza, na prática, a criação e a gestão de uma persona influenciadora de altíssimo engajamento sem qualquer exposição humana real. A solução elimina barreiras financeiras relativas a locações, fotógrafos e contratos de imagem, cumprindo o objetivo de democratizar a atuação no mercado de influência digital e oferecendo a empreendedores uma via escalável, segura e flexível para a divulgação de marcas.

5 Conclusão

Este trabalho partiu de um problema simples: muitas pessoas querem criar conteúdo, mas não querem se expor, têm timidez ou não têm recursos para uma produção profissional. O mesmo vale para pequenas marcas, que nem sempre podem pagar um influenciador. Como solução, foi montado um pipeline de Inteligência Artificial Generativa que une Modelos de Linguagem (para gerar prompts em JSON) e modelos de difusão (para

criar imagem e vídeo), com foco em manter o rosto e a identidade da persona sempre consistentes.

O objetivo foi alcançado. O pipeline funciona e foi validado na prática: é possível criar uma influenciadora virtual com identidade consistente em diferentes cenários, roupas, poses e produtos, sem nenhuma exposição de uma pessoa real. Uma influenciadora base foi gerada como prova de conceito e serve de modelo da esteira, mostrando que dá para entrar no mercado de influência digital sem aparecer e para divulgar produtos por meio de uma persona virtual.

Os principais resultados foram: a extração das características em JSON pela LLM, a colagem de identidade que trava o rosto da persona, a geração das imagens em novos contextos sem perder os traços originais e a animação dos frames em vídeo (Veo 3 e Kling 2.6), além do ajuste de câmera pelo ComfyUI. O controle da cena pelo JSON foi o grande diferencial em relação a ferramentas mais automáticas.

A principal limitação é que os modelos às vezes “alucinam” e copiam detalhes indesejados da imagem de referência, como piercings, tatuagens ou erros de cenário. Por isso, ainda não dá para automatizar tudo: é preciso curadoria manual, com tentativa e erro até chegar a um resultado profissional. Algumas ferramentas, como o Kling 2.6, também limitam a quantidade de gerações por créditos.

Como trabalhos futuros, pretende-se reduzir a curadoria manual e tornar o processo mais automático e estável, além de melhorar a voz da persona, ampliar sua presença em mais plataformas e organizar o fluxo para que qualquer pessoa consiga reproduzi-lo sem conhecimento técnico.

Referências

JAYASINGH, Sudarsan; SIVAKUMAR, Arunkumar; VANATHAIYAN, Arputha Arockiaraj. Artificial Intelligence Influencers' Credibility Effect on Consumer Engagement and Purchase Intention. *Journal of Theoretical and Applied Electronic Commerce Research*, v. 20, n. 1, art. 17, 2025. DOI: 10.3390/jtaer20010017. Disponível em: <https://www.mdpi.com/0718-1876/20/1/17>. Acesso em: 27 mar. 2026.

LIAN, L.; LI, B.; YALA, A.; DARRELL, T. LLM-grounded Diffusion: Enhancing Prompt Understanding of Text-to-Image Diffusion Models with Large Language Models. In: *INTERNATIONAL CONFERENCE ON LEARNING REPRESENTATIONS (ICLR)*, 12., 2024, Viena. Anais [...]. Viena: ICLR, 2024. Disponível em: <https://openreview.net/forum?id=u4jiA5XmKj>. Acesso em: 14 maio 2026.

METZLER, Hannah; GARCIA, David. Social Drivers and Algorithmic Mechanisms on Digital Media. *Perspectives on Psychological Science*, v. 19, n. 5, p. 735–748, set. 2024. DOI: 10.1177/17456916231185057. Disponível em: <https://journals.sagepub.com/doi/10.1177/17456916231185057>. Acesso em: 27 mar. 2026.

SOROSRUNGRUANG et al. How real is real enough? Unveiling the diverse power of generative AI-enabled virtual influencers and the dynamics of human responses. *Psychology & Marketing*, 2024. DOI: 10.1002/mar.22105. Disponível em: <https://onlinelibrary.wiley.com/doi/10.1002/mar.22105>. Acesso em: 27 mar. 2026.

YU, Joanne et al. Artificial intelligence-generated virtual influencer: Examining the effects of emotional display on user engagement. *Journal of Retailing and Consumer Services*, v. 76, jan. 2024. DOI: 10.1016/j.jretconser.2023.103560. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0969698923003119>. Acesso em: 27 mar. 2026.