

# Photo-to-Photo Recreation Tools with DALL-E 3: Current Limitations and Available Solutions

Based on the available search results, achieving strict 1:1 geometry and color fidelity for photo-to-photo recreation using DALL-E 3 presents significant technical challenges. The current landscape reveals fundamental limitations in DALL-E 3's architecture that constrain precise photo modification capabilities, though several repositories offer partial solutions for controlled image editing.

## Current Technical Limitations of DALL-E 3

The search results reveal critical constraints that impact photo-to-photo recreation capabilities. DALL-E 3 does not accept images as direct input, requiring users to rely on text descriptions for image generation<sup>[1]</sup>. Additionally, there is no mechanism to seed the model or feed previous images directly into DALL-E 3, which eliminates the possibility of fine-tuning for specific images<sup>[2]</sup>. Each generation creates a new prompt through moderation processes, preventing repeatable and reliable results even with identical parameters<sup>[2]</sup>. These architectural limitations fundamentally challenge the goal of maintaining strict 1:1 geometry and color fidelity.

## Available Repository Solutions

### DALL-E Image Editor with Selective Modification

The most comprehensive solution for controlled image editing is found in the DALL-E Image Editor repository<sup>[3]</sup>. This Python application built with Tkinter provides functionality to edit selected images by marking specific areas for modification and providing descriptive prompts. Users can open image files, select regions using brush tools including erase, restore, circle brush, square brush, and line brush options, then specify desired changes through text descriptions<sup>[3]</sup>. While this approach enables targeted modifications, it relies on DALL-E's generation capabilities rather than preserving original image geometry.

### ComfyUI DALL-E Integration for Inpainting

The ComfyUI DALL-E integration provides nodes for inpainting operations that could support selective modifications<sup>[4]</sup>. This extension integrates OpenAI's DALL-E 3 API for image generation and inpainting, automatically handling image resizing and format requirements with RGBA conversion<sup>[4]</sup>. The system supports text-to-image generation, inpainting with prompt and mask, and batch generation capabilities<sup>[4]</sup>. However, the inpainting functionality appears to be optimized for DALL-E 2 rather than DALL-E 3, limiting its applicability for the most current model capabilities.

## **API-Based Outpainting Framework**

The DALL-E API outpainting sample demonstrates how to leverage DALL-E's API for specific image modifications, though focused on extending rather than modifying existing content<sup>[5]</sup>. This repository shows techniques for manipulating images within "generation frames" by shifting original images and introducing filler pixels that are subsequently replaced<sup>[5]</sup>. The approach provides insight into API-based image manipulation but does not address the core requirement of preserving original geometry while making selective modifications.

## **Command-Line and Server Solutions**

### **CLI Tools for Basic Operations**

The dallecli repository offers command-line functionality for DALL-E operations, including generation, editing, and filtering capabilities<sup>[6]</sup>. The tool supports image editing through brightness, contrast, and sharpness adjustments, though these modifications represent basic image processing rather than content-based changes<sup>[6]</sup>. While useful for workflow automation, this solution does not address the specific requirements for photorealistic content modification with geometric preservation.

### **MCP Server Implementation**

The DALL-E MCP Server provides Model Context Protocol functionality for image generation using DALL-E 2 and DALL-E 3, with editing capabilities limited to DALL-E 2<sup>[7]</sup>. This server implementation supports image generation, editing of existing images, and creation of variations, though the editing features are not available for the latest DALL-E 3 model<sup>[7]</sup>. The server can be configured for integration with various applications but does not solve the fundamental limitation of DALL-E 3's inability to directly process input images.

## **Research and Development Context**

### **Advanced Editing Frameworks**

The HILITE platform represents a comprehensive approach to image editing by combining multiple diffusion-based models, achieving superior performance compared to DALL-E 3 and Google's Imagen 3 across quantitative metrics<sup>[8]</sup>. This platform implements a human-in-the-loop design that iteratively incorporates user feedback and provides a straightforward workflow from image input to output selection<sup>[8]</sup>. However, HILITE functions as a commercial platform rather than an open-source repository for photo-to-photo recreation.

### **Dataset and Training Approaches**

The HQ-Edit dataset demonstrates advanced approaches to instruction-based image editing using GPT-4V and DALL-E 3 for creating high-quality training data<sup>[9]</sup>. This research shows that models fine-tuned with such datasets can achieve state-of-the-art performance, even surpassing those trained with human-annotated data<sup>[9]</sup>. While this represents significant

progress in image editing capabilities, it focuses on training data creation rather than providing ready-to-use tools for photo recreation.

## Conclusion

Current open-source repositories and tools provide limited solutions for photo-to-photo recreation with strict 1:1 geometry and color fidelity using DALL-E 3. The most practical approaches involve selective editing through masking and inpainting techniques, though these methods cannot guarantee preservation of original image geometry and perspective. The fundamental limitations of DALL-E 3's architecture, particularly its inability to directly process input images and maintain consistent outputs, present significant barriers to achieving the desired functionality. Users seeking such capabilities may need to consider alternative approaches using other image editing models or hybrid solutions that combine multiple technologies for enhanced control and fidelity.



1. [https://www.reddit.com/r/ChatGPT/comments/177cveg/image\\_as\\_input\\_to\\_dalle3\\_is\\_there\\_a\\_trick\\_to\\_make/](https://www.reddit.com/r/ChatGPT/comments/177cveg/image_as_input_to_dalle3_is_there_a_trick_to_make/)
2. <https://github.com/JukoJones/dalle3-image-generation-test>
3. [https://github.com/xzxui/dalle\\_image\\_editor](https://github.com/xzxui/dalle_image_editor)
4. <https://github.com/cleanlii/comfyui-dalle-integration>
5. <https://github.com/SabatinoMasala/dalle-api-outpainting-sample>
6. <https://github.com/raiyanyahya/dallecli>
7. <https://github.com/Garoth/dalle-mcp>
8. <https://ieeexplore.ieee.org/document/10825916/>
9. <https://arxiv.org/abs/2404.09990>