

KI in Medienanwendungen - Dokumentation

Flavio Carbone

Mein Thema war Image Generation mit Foundation Models. Zuerst habe ich mich über die verschiedenen Bildgenerierungsmodelle informiert und ihren zeitlichen Verlauf angeschaut. Ich habe zum Beispiel gelernt, wie Generative Adversarial Networks (GANs) funktioniert haben und wie sich später Diffusionsmodelle entwickelt haben.

Auch neuere Ansätze wie Flow-Matching-Modelle habe ich mir angeschaut. In meiner praktischen Arbeit habe ich versucht, verschiedene lokale Modelle auf meinem Rechner zu installieren und mit unterschiedlichen Prompts zu experimentieren. Dabei wollte ich besonders verstehen, warum Modelle manche Prompts nicht vollständig oder korrekt umsetzen können, das war auch meine Leitfrage.

Ich habe versucht, bekannte Modelle wie Stable Diffusion und Flux3 zu nutzen, allerdings funktionierten diese Modelle bei mir kaum oder gar nicht. Am Ende habe ich dann mit den Modellen zTurbo und Oven gearbeitet, weil sie stabiler liefen. Auch Flux2 habe ich ausprobiert, aber auch hier gab es Probleme.

Da ich nur einige Modelle lokal nutzen konnte, habe ich zusätzlich Gemini und ChatGPT verwendet, um den Unterschied zwischen lokalen Modellen und cloud- oder unternehmensbasierten Modellen zu sehen. Die Modelle, die bei mir nicht liefen, hatten vor allem das Problem, dass mein MacBook nicht genug Arbeitsspeicher hatte.

Das Generieren von Bildern auf meinem Mac hat bis zu 30–40 Minuten pro Bild gedauert, obwohl ich die Einstellungen so gewählt habe, dass nur kleine Bilder (512×512) erstellt werden. Ich habe schließlich mit verschiedenen Prompts experimentiert, die ich unten zusammen mit den erzeugten Bildern der jeweiligen Modelle aufgelistet habe. Am besten abgeschnitten hat ChatGPT, da dieses Modell am zuverlässigsten ein randvolles Glas Wein generieren konnte.

1. Installation von ComfyUI

Laden Sie ComfyUI in der Version **v0.5.0** über die offizielle Website herunter und installieren Sie die Software:

<https://www.comfy.org/download>

2. Installation der benötigten Modelle

Installieren Sie innerhalb von ComfyUI die Modellpakete **zTurbo** und **Oven**. Diese Pakete sind für die gewünschte Funktionalität erforderlich.

3. Konfiguration der Einstellungen

Übernehmen Sie die Einstellungen gemäß der vorgegebenen Konfiguration (siehe Abbildung bzw. vorheriger Abschnitt).

4. Eingabe der Prompts












Geben Sie anschließend die gewünschten Prompts in die vorgesehenen Eingabefelder ein und starten Sie den Generierungsprozess.










5. Weitere Unterstützung

Detaillierte Informationen, Troubleshooting und weiterführende Dokumentation finden Sie im offiziellen GitHub-Repository von ComfyUI:

<https://github.com/Comfy-Org/ComfyUI>

Benutze Prompts

Prompt	Z-Turbo	Gemini	ChatGPT	Ovis
Ein bis zum Rand gefülltes Weinglas				
Ein Übervoll Eingeschenktes Glas Wein				
Ein Weinglas Gefüllt Bis Ganz Oben				

Prompt	Z-Turbo	Gemini	ChatGPT	Ovis
ein bis zum randgefülltes Glas Wasser				
Ein Übervoll Eingeschenktes Glas Wein				
Ein Weinglas Gefüllt Bis Oben				

Ein Weinglas Gefüllt Bis Ganz weit Oben



Ein Wasserglas Gefüllt Bis Ganz weit Oben

