

Inhaltsverzeichnis

Ressourcen.....	2
Vorwort.....	2
Vorgedanken.....	2
Workflow Integration.....	2
Hardware.....	3
Upscaling Methoden.....	3
Extern.....	3
Supersimpel.....	3
Mit einem AI Upscale Model.....	4
Supersimpel + AI Upscale Model.....	4
Supir und CSSR.....	5
KSampler Upscaling mit AnimateDiff.....	5
Zusammenfassung.....	9

Ressourcen

Die Beispielworkflows zum Artikel findet ihr hier:

https://www.tomgoodnoise.de/data/upscale_workflows_tomgoodnoise.zip

Vorwort

Vorgedanken

Ich hatte eigentlich vor mir einen ultimativen Video Upscaling Workflow für ComfyUI zu bauen. Einfach Video reindropfen, auf Queue drücken und glücklich sein. Und habe dann sehr schnell festgestellt dass es diesen Workflow so gar nicht geben kann. Wie man vorgeht hängt einfach von zu vielen Faktoren ab. Und jedes Input Video und jeder Workflow benötigt im Grunde eine etwas andere Behandlung. Andere Einstellungen. Oder auch eine komplett andere Methode.

Einen Cartoon kann man anders hochskalieren als eine Naturdokumentation. Die gewünschte Endgrösse spielt eine Rolle. Das Model das man verwendet. Ob man es Quick und Dirty haben will, oder alle Zeit der Welt hat. Und natürlich wie gut der PC ist auf dem das rechnet.

Sehr wichtig ist auch wie gross das Inputvideo ist, und wie es erstellt wurde. Videos die mit AnimateDiff erstellt wurden haben schon ein Grössenproblem. AnimateDiff wurde mit Bildern in Grösse von 512x512 trainiert. Um ein 16:9 Video zu erstellen muss man für sinnvolle Ergebnisse sogar noch weiter runterskalieren damit AnimateDiff noch brauchbare Ergebnisse liefert. Ich habe in meinen AnimateDiff Workflows meist mit 624 x352 oder 568 x 320 Pixel angefangen. Und dann wird eben so manch Detail wie ein Gesicht einfach nur noch ein undefinierbarer Pixelbrei.

So ein Video würde ich deswegen nicht einfach nur hochskalieren. Da ist dann einfach nicht genug Bildinformation vorhanden die man hochskalieren könnte. Sondern ich würde es über einen zweiten (K)Sampler hochrechnen und quasi noch mal erstellen.

Viele Faktoren also. Und deswegen kann es den einen Workflow gar nicht geben. Deswegen gibt es stattdessen eine kurze Übersicht über die im Moment gängigen Methoden. Die Workflows sind in einem Zipfile beigelegt.

Ich rede hier absichtlich nicht über Details oder erkläre Nodes, das würde den Rahmen sprengen. Darüber gibt es schon genug Artikel. Sondern gebe nur eine grobe Übersicht über die Methoden. Und beschreibe einige Fussangeln.

Workflow Integration

Ob ihr das Upscaling in einem eigenen Upscaling Workflow machen wollt, oder in einen existierenden Workflow integriert müsst ihr selber entscheiden. Dazu gibt es viele Meinungen.

Der Vorteil das zu integrieren ist dass ein Zwischenschritt entfällt. Und das Bildmaterial wird nicht in ein Video komprimiert. Dabei geht ja immer ein wenig Qualität verloren. Der Nachteil nennt sich OOM. Out of Memory. Das kann funktionieren, kann aber auch an fehlendem Ram scheitern. Ich persönlich habe gern alles in einem Workflow. Sofern es funktioniert. In den Beispielen hier ist die Methode natürlich vereinzelt.

Hardware

Ich habe derzeit einen schon wieder veralteten 5800 X3d Prozessor, 32 Gb Ram und eine 4060 TI mit 16 Gb Vram. Meine alte Grafikkarte, die ich für meine Animatediff Videos verwendet habe, war eine 3060 TI mit 8 Gb Vram. Ich kenne noch das Leid des täglichen OOM's. Und dass manche Sachen einfach nicht zu realisieren sind weil der Ram ausgeht. Viele Workflows haben da schlicht nicht funktioniert. Wenn ihr mit AI arbeiten wollt tut euch einen Gefallen und holt euch mindestens eine Karte mit 12 Gb Vram.

Upscaling Methoden

Vorab noch eine Anmerkung. Ich habe die Methoden nicht erfunden. Sie sind schon eine ganze Weile verfügbar. Ich beschreibe sie hier nur, als Übersicht.

Für meine Tests verwende ich hier ein Inputvideo in der Größe 720 x 480, erstellt mit Cogvideo. Mit einer Länge von 6 Sekunden bei 24 Bildern pro Sekunde. Und skaliere es um den Faktor 2 hoch.

Die Methoden sind wie folgt.

Extern

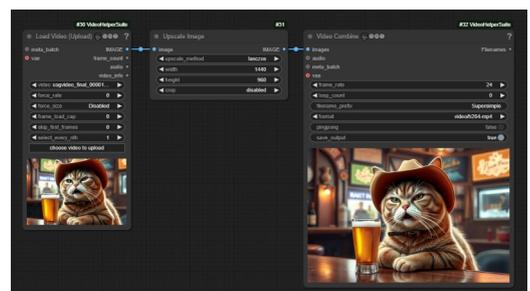
Los geht es schon damit das Upscaling erst gar nicht in ComfyUI zu machen. Sondern zum Beispiel im Branchenprimus Topaz Video Upscaler. Oder mit dem Super Scale Feature in der kommerziellen Version von DaVinci Resolve. Beides kostenpflichtige Lösungen (das Super Scale gibts in DaVinci erst mit der Studio Version) mit einem recht guten Ergebnis. Beide arbeiten mit AI Upscaling Methoden.

<https://www.topazlabs.com/topaz-video-ai>

<https://www.blackmagicdesign.com/de/products/davinciresolve>

Supersimpel

In ComfyUI besteht die simpelste Methode darin einfach eine Video Input Node an eine Upscale Node zu hängen, und von da aus direkt in die Video Combine Node zu gehen. So

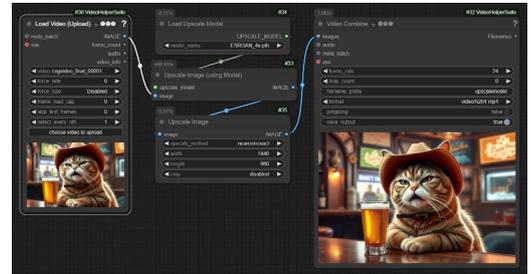


bekommt man natürlich keine Details dazu. Das skaliert einfach nur die Pixel grösser. Als wenn man in Photoshop ein Bild grösser skaliert.

Mein Beispielvideo mit 6 Sekunden Länge und 720x480 Auflösung war so in 10 Sekunden auf das Doppelte hochskaliert.

Mit einem AI Upscale Model

Es gibt eine Upscale Image Node mit der man ein AI Upscale Model verwenden kann. Eins der älteren und bekannteren Modelle ist ESRGAN. Upscale Modelle rechnen Details dazu. So werden zum Beispiel Haare viel höher aufgelöst dargestellt. Oder bei einem Stoff kann man plötzlich das Webmuster erkennen. Und das passiert je nach Modell auf eine etwas andere Weise. Es gibt Upscale Modelle die speziell auf Cartoons trainiert sind. Oder speziell auf Gesichter.



Für viele Videos ist das hochskalieren mit einem Upscaling Modell schon ausreichend. Besonders wenn man bedenkt dass die neuen Videogeneratoren wie Hunyuan oder CogVideo inzwischen recht hochauflösend arbeiten. Die Minimumgrösse für CogVideo liegt bei 720x480. Und die neueste Version kann die doppelte Grösse.

Das Ganze dauert natürlich schon etwas länger als einfach nur die Bilder hochzukalieren. Aber die Qualität hat sich verbessert. Das Fell der Katze ist nun viel detaillierter.

Der Workflow hat für mein 6 Sekunden Beispielvideo 7.4 Minuten gebraucht. Das könnte man noch etwas reduzieren, indem man keinen 4x-Upscaler, sondern nur einen 2x-Upscaler verwendet und so das Downscaling vermeidet. Aber Oversampling steigert die Qualität. Und mit einer einfacher Upscale Node kann dann auf die benötigte Auflösung herunterskalieren. Damit kann man nicht nur hochskalieren.

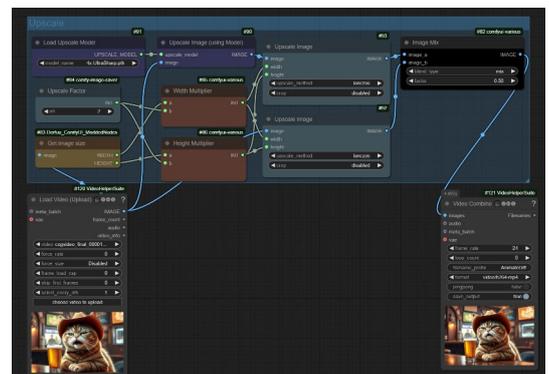
Das Problem ist, dass einem bei einem 4x Modell auch gern mal der Vram ausgehen kann. Das Problem hatte ich mit meiner alten 8-GB-Grafikkarte. Ich habe da das 2x RealESRGan Upscale Modell verwendet.

Upscaling Models findet ihr zum Beispiel hier: <https://openmodeldb.info/>

Supersimpel + AI Upscale Model

Die AI Upscale Model Methode kann manchmal etwas des Guten zu viel tun. Je nach verwendetem Modell. Es kann unschöne Artefakte geben, oder das Ganze kann zu scharf geraten. Dem kann man etwas entgegensteuern indem man die obigen zwei Methoden mischt.

Eine Linie arbeitet mit einem Upscaling Model. Eine Linie arbeitet mit simplem hochskalieren. Und dann mischt man die zwei Bilder zusammen.



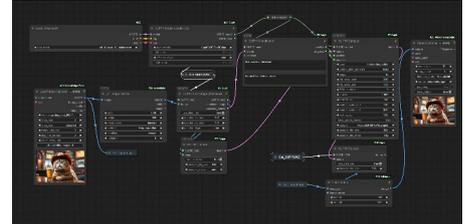
Der Workflow hat für mein 6 Sekunden Beispielvideo in 720x480 Auflösung etwas mehr als 8 Minuten gebraucht.

Supir und CSSR

<https://github.com/kijai/ComfyUI-SUPIR>

<https://github.com/kijai/ComfyUI-CCSR>

Supir und CSSR sind beides Custom Nodes für ComfyUI die speziell für das hochskalieren von Bildern und Videos gemacht sind. Sie versprechen eine Qualität die dicht an das Ergebnis vom Video Upscaling Branchenprimus Topaz Video herankommt. Beide haben aber ein grosses Performanceproblem.



Ich habe zwar beide Workflows zum Laufen bekommen. Aber beides mal den Workflow dann gestoppt und nicht zu Ende laufen lassen. Ich kann also noch nicht mal was zur Qualität sagen. Zehn Minuten und mehr für ein Bild bei 24 Bildern pro Sekunde machte mir einfach keinen Spass. Und bei Supir kommt noch ein Lizenzproblem oben drauf. Es ist nicht für den kommerziellen Gebrauch lizenziert.

CCSR muss man über die Github URL installieren. Supir kann man vom ComfyUI Manager installieren. Vorsicht Falle, der Ordner in ComfyUI\custom_nodes wird mit Kleinbuchstaben erstellt. Und dann funktioniert die Download Node nicht mit der man das Supir Modell im Workflow automatisch herunterläd. Den Ordner also unbedingt in *ComfyUI-Supir* umbenennen bevor ihr loslegt.

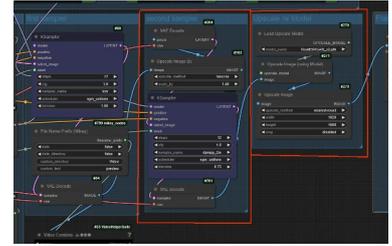
Beide Workflows sind die Standardworkflows aus dem jeweiligen Examples Ordner der Custom Nodes. Bei denen ich einfach die Image Nodes durch Video Nodes ersetzt habe. Die Workflows schenke ich mir deswegen. Und auch wegen der unglaublich langen Kalkulationszeit. Das ist für die allermeisten Offline Zwecke einfach viel zu langsam.

KSampler Upscaling mit AnimateDiff

Das Herz jeder AI Bilderzeugung ist ein Sampler. Für Stable Diffusion ist das ein KSampler. Für die Bilder AI Flux der Flux Sampler. Für die Video AI Cogvideo der Cogvideo Sampler, und so weiter.

Und hier fängt es an etwas komplizierter zu werden. Denn man kann das Video nicht mal einfach so noch mal durch einen KSampler laufen lassen. Die Bilder haben dann keinen Zusammenhang mehr. Denn jedes Bild wird dann zufällig "verbessert". Das flackert dann. Was man braucht ist ein weiterer Durchgang durch ein Videomodell. Für die Konsistenz. Und dafür gibt es meines Wissens nach im Moment nur AnimateDiff. Und das läuft dann über ein Stable Diffusion Weight und einen KSampler. Und um die Sache weiter zu verkomplizieren, es gibt im Moment drei AnimateDiff Modelle. Gängigste Version ist AnimateDiff 2 LCM. Einfach weil es dafür auch Loras gibt und es sehr schnell arbeitet.

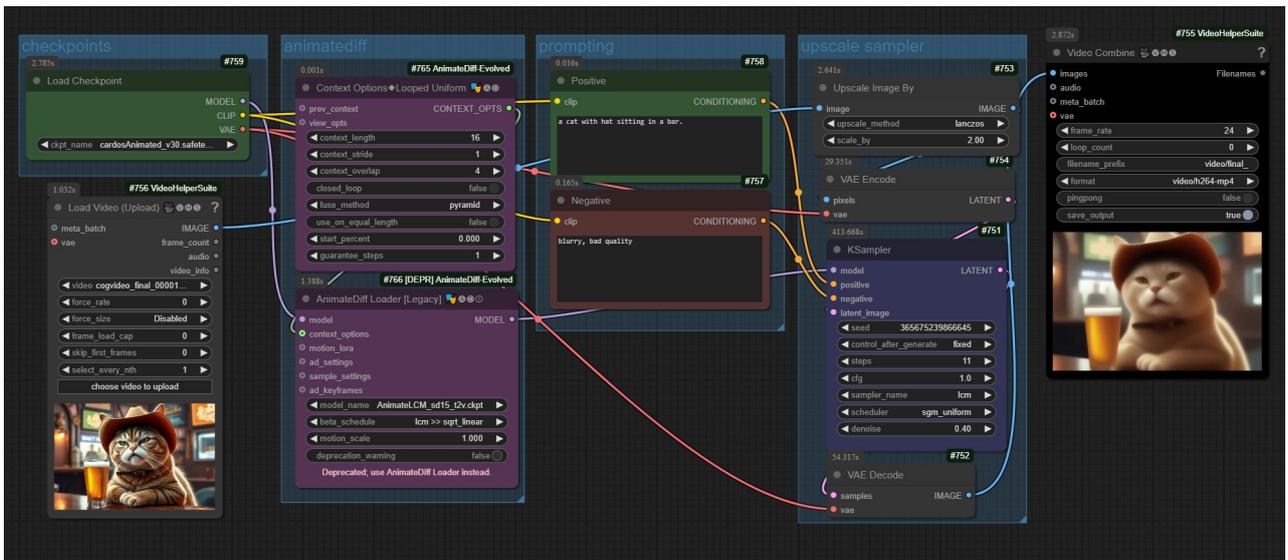
Über einen KSampler zu skalieren hat allerdings einen Preis. Das Ergebnis entspricht nicht 100% dem Input. Es wird quasi ein neues Video erstellt. Die Bilder werden neu berechnet, und es werden neue Details dazukalkuliert. Das kann die Farben verändern. Oder auch den gesamten Look. Jedes Stable Diffusion Weight gibt ja auch ein etwas anderes Ergebnis. Weswegen man am Besten das gleiche Weight für das Erstellen und hochskalieren verwenden sollte.



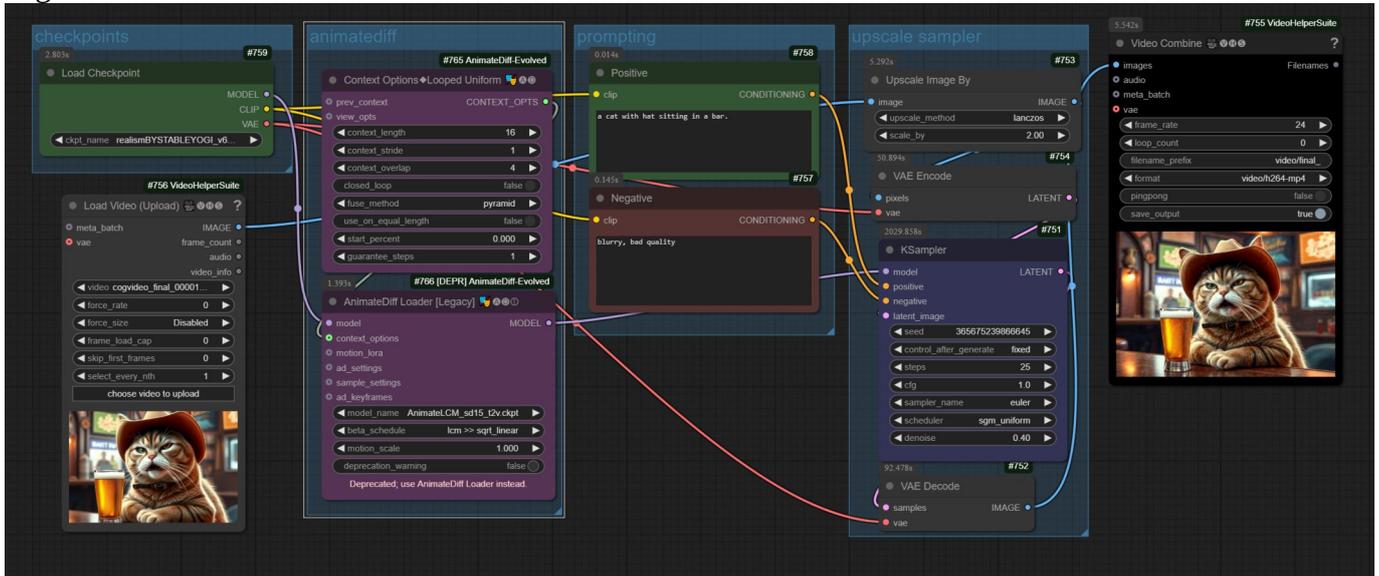
Und man muss irgendwie einen Prompt einbauen. Sei es in Textform oder über eine Bildanalyse. Am Besten wäre natürlich der Originalprompt. Womit wir wieder beim Thema wären ob man einen Upscale Workflow gleich in den Originalworkflow einbauen sollte. Was ich vielfach gemacht habe.

Zurück zum ursprünglichen Problem. Wir benötigen einen KSampler und AnimateDiff für kohärente Ergebnisse. Dafür nehmen wir am Besten AnimateDiff LCM, weil es schnell ist. Und die Version bessere Ergebnisse als AnimateDiff 2 und 3 liefert.

Hier gleich mal eine kleine Falle. Man sollte ja denken wenn man schon AnimateDiff LCM verwendet dass man das dann auch im KSampler die LCM Methode verwenden kann. Das funktioniert im Workflow in dem man das Video erstellt recht gut. Doch hier beim hochskalieren bleibt von der Katze mit der LCM Methode quasi nichts über.



Wenn ich hier nun den Ksampler von LCM auf Euler umstelle und ein anderes Weigth verwende ist das Ergebnis deutlich koheranter.



Dieser Workflow hat für mein 6 Sekunden Beispielvideo in 720x480 Auflösung etwas mehr als 36 Minuten gebraucht.

Es flackert zwar noch ein wenig. Nicht ganz so doll als wenn man es ohne AnimateDiff machen würde. Aber sichtbar. Wie ein schlechter Super 8 Film. Was man hier leider im Shot nicht sieht. Es könnte schon helfen die Samples weiter hochzuschrauben. Dafür sieht das Ergebnis schon sehr nach dem Eingabevideo aus.

Es braucht einfach noch ein paar weitere Schritte zum verfeinern.

Genau das hat Banodoco in seinem supermegabadassupscaler Workflow getan. In seinem Workflow findet sich ein IP Adapter der den Prompt ersetzt. Und noch eine Controlnet Instanz die für noch mehr Koherenz sorgt. Und er verwendet die Ultimate SD Upscale Node, was Tiling ermöglicht. So rennt man nicht so schnell in OOM's. Ein Upscale Modell ist auch noch im Spiel. Und sein Workflow liefert tatsächlich ein beeindruckendes recht koheräntes und vor allem flackerfreies Ergebnis.

Der Workflow hat für mein 6 Sekunden Beispielvideo in 720x480 Auflösung etwas mehr als 24 Minuten gebraucht. Ultimate Upscale ist trotz Tiling schneller als wenn man einen reinen Ksampler verwendet.

Das ist der Preis, den man für das Hochskalieren mit einem KSampler zahlt. Man kann ziemlich dicht rankommen. Aber abhängig vom gewählten AI Weight, dem gewählten Denoise Faktor und den sonstigen Einstellungen kann das Ergebnis auch recht deutlich abweichen. KSampler kalkulieren eben das Video neu.

Zusammenfassung

Den einen glücklichmachenden Upscaling Workflow gibt es nicht. Es kommt eben immer darauf an. Mit einem Video das mit AnimateDiff erstellt wurde, und in dem ihr vielleicht einen Pixelbrei habt, ist es empfehlenswert mit einem KSampler hochzuskalieren. Damit man die fehlenden Details bekommt. Mit anderen Methoden würde ich eher auf das Upscaling mit einem Model setzen.

Ich hoffe dass ich dafür ein paar Denkanstösse geben konnte.

Tom Goodnoise - 14.01.2025

<https://www.tomgoodnoise.de>