Dein Foto als Kunstwerk (Style Transfer)

Kapitel online

Lade ein eigenes Foto, wähle einen Malstil — van Gogh, Hokusai, Munch, Kandinsky oder dein eigenes Bild — und sieh zu, wie zwei kleine Netze daraus ein Gemälde machen. Direkt im Browser, ohne Upload. Und du siehst das Erstaunlichste dabei: Ein kompletter Malstil passt in 100 Zahlen.

Interaktiv · Kapitel 50 von 50 · ca. 25 Min. · 300 XP

Einfach gesagt: Du legst dein Foto neben ein Gemälde, und ein Netz verwandelt es in dessen Malstil — der komplette Stil passt dabei in nur 100 Zahlen.

Du kennst diese Apps: Selfie rein, ein Klick auf „van Gogh“ — und heraus kommt ein wirbelndes Ölgemälde von dir. Millionenfach heruntergeladen, meistens mit dem Kleingedruckten „Ihr Foto wird zur Verarbeitung an unsere Server übertragen“.

Heute läuft genau dasselbe — **auf deinem Gerät, in diesem Tab**. Und du siehst dabei etwas, das die Apps sorgfältig verstecken. Denn wenn dieses System gleich „Sternennacht“ auf dein Foto legt, dann hat es van Goghs Malstil vorher zusammengefaltet in einen einzigen kleinen Zahlenblock:

Ein kompletter Malstil passt in 100 Zahlen.

Roter Faden: Erst spielen, dann runter zum Tensor. Nach der ersten Verwandlung zeige ich dir die Zahlen, aus denen dein Gemälde entstanden ist — sie stehen dann wirklich da.

Was du hier lernst

- Wie zwei kleine Netze zusammenarbeiten: eins **liest den Stil**, eins **malt damit**
- Was ein **Stil-Vektor** ist — und warum 100 Zahlen für „van Gogh“ reichen
- Warum du den Stil per Regler **halb** auftragen kannst (Spoiler: Latent Space, Kapitel 310)
- Warum das hier in Sekunden läuft, während Bild-KIs wie in Kapitel 340 Rechenzentren brauchen
- Und ganz praktisch: dein eigenes Bild, dein eigener Stil, dein Ergebnis zum Behalten

Für den Unterricht

Einstieg (5 Min): Kapitel gemeinsam lesen, Alltagsbeispiel aufgreifen.

Erkunden am Gerät (15 Min): die interaktiven Module auf herrki.de/480_style_transfer ausprobieren.

Sichern mit Quiz (5 Min): „Kurzer Check" besprechen, „Das nimmst du mit" festhalten.

Deine Fotos bleiben auf deinem Gerät

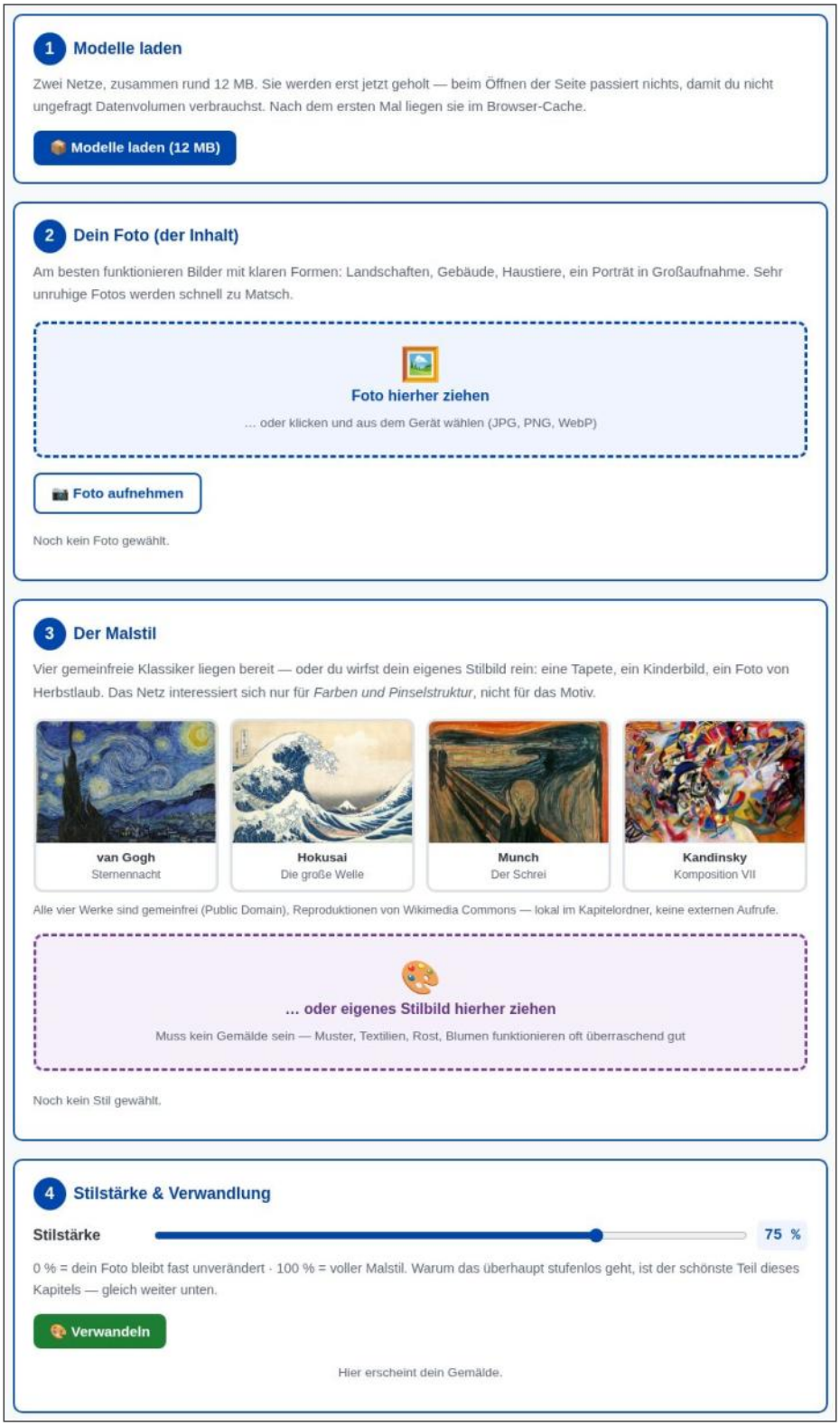
Alles läuft in deinem Browser: die Modelle werden einmal heruntergeladen wie ein Bild auf einer Webseite, danach rechnet *dein* Prozessor. Dein Foto wird **nirgendwohin hochgeladen** — es verlässt diesen Tab nicht, es wird nicht gespeichert und niemand sieht es außer dir.

Das ist genau das Argument, das du brauchst, wenn du Familienfotos oder Kinderbilder durch so einen Filter schicken willst. Bei den bunten Handy-Apps landet dein Bild auf fremden Servern und in fremden AGB. Hier nicht. Zum Prüfen: Nimm dieses Kapitel mit ins Flugzeug — **ohne Internet funktioniert es weiter**, sobald die Modelle einmal geladen sind.

Die Werkstatt

Am Gerät ausprobieren: herrki.de/480_style_transfer#abschnitt-3

Direkt zum Modul



Interaktives Modul – am Gerät ausprobieren

⚠ **Hm, da hakt etwas:**

Runter zum Tensor: Wie viel ist „van Gogh“?

Jetzt der Kernmoment. Als du auf „Verwandeln“ gedrückt hast, ist das Stilbild **zuerst durch ein eigenes Netz** gelaufen — den *Stil-Leser*. Es hat das Gemälde nicht gespeichert, nicht ausgeschnitten, nicht kopiert. Es hat es **zusammengefaltet auf 100 Zahlen**. Und mehr weiß das malende Netz danach über deinen Stil auch nicht.

Die ersten 10 dieser 100 Zahlen — aus deinem gewählten Stil, gerade eben berechnet:

```
// Noch nichts berechnet – wähle oben ein Stilbild und drück auf „Verwandeln“.
```

Und hier alle 100 Zahlen als Balken — das ist der komplette „Stil-Fingerabdruck“:

Die Shape-Reise deines Bildes

Wie in jedem Kapitel: Am Ende sind es Tensoren. Hier sind es genau drei Stationen.

Stilbild
[256, 256, 3]
196.608 Zahlen

→ Stil-Netz →

Stil-Vektor
[100]
100 Zahlen

Aus fast 200.000 Zahlen werden 100. Das ist eine Verdichtung auf 0,05 % — und trotzdem reicht es aus, um „van Gogh“ zu malen.

Dein Foto
[H, B, 3]
max. 512 px lange Kante

+ Stil-Vektor
[100]

→ Transformer-Netz →

Dein Gemälde
[H, B, 3]

Das zweite Netz bekommt also **zwei** Eingaben: das Bild, das es umbauen soll, und die 100 Zahlen, die ihm sagen *wie*. Es hat nie „gelernt, van Gogh zu malen“ — es hat gelernt, **aus 100 beliebigen Zahlen einen Malstil zu machen**. Deshalb funktioniert auch dein eigenes Stilbild, das kein Mensch je in das Training gesteckt hat.

Warum 512 px? Weil die Rechenzeit mit der Fläche wächst: ein Bild mit doppelter Kantenlänge ist *viermal* so teuer. 512 px ist der Punkt, an dem es auf einem Handy noch erträglich bleibt — genau die Abwägung, die auch die App-Entwickler treffen mussten.

Brücke zu Kapitel 310: Der Stil-Vektor ist ein Punkt im Raum

In „Der Latent Space“ hast du gesehen: KI-Systeme legen Bedeutung als **Punkte in einem hochdimensionalen Raum** ab, und ähnliche Dinge liegen nah beieinander. Genau das ist hier passiert — nur ist es diesmal ein **Stil-Raum** mit 100 Dimensionen. „Sternennacht“ ist ein Punkt darin, dein Urlaubsfoto ein anderer.

Und jetzt kommt der Trick mit dem Regler. Die Stilstärke schneidet nichts ab und blendet nichts über. Sie **wandert auf der geraden Linie zwischen zwei Punkten** in diesem Raum:

```
// Stilstärke a = 0.75 bedeutet: 75 % des Weges vom Foto-Punkt zum Stil-Punkt
vektor = stilVektor * a + inhaltVektor * (1 – a)
```

Ja — dein *Foto* läuft ebenfalls durch den Stil-Leser und bekommt seinen eigenen 100er-Vektor. Denn ein Foto hat auch einen „Stil“: seine Farben, sein Kontrast, seine Textur. Bei 0 % beschreibt der Vektor also den Stil deines eigenen Fotos, und das Netz malt es fast unverändert nach. Bei 100 % steht er auf van Gogh.

Dazwischen liegt kein Mischbild, sondern ein echter neuer Stil — einer, der so nie gemalt wurde und für den das Netz trotzdem ein sinnvolles Ergebnis liefert. Das ist die Eigenschaft, die den Latent Space so mächtig macht: Man kann *zwischen* den bekannten Punkten spazieren gehen, und es kommt immer noch etwas Sinnvolles heraus. Probier den Regler mal auf 30 %, 50 %, 90 % — dreimal derselbe Stil, drei verschiedene Punkte auf derselben Linie.

Brücke zu Kapitel 340: Warum das hier kein Diffusionsmodell ist

In „Wie Bild-KIs malen“ arbeitet das System völlig anders: Es startet bei **reinem Rauschen** und entfernt in Dutzenden Schritten nach und nach das Zufällige, bis ein Bild übrig bleibt. Jeder Durchgang ist ein neues Raten, jedes Ergebnis ein Unikat — und jeder dieser Schritte kostet einen kompletten Netz-Durchlauf.

Hier gibt es **kein Rauschen und kein Raten**. Dein Bild geht *einmal* durch ein Netz und kommt auf der anderen Seite stilisiert wieder heraus — eine direkte, gelernte Transformation, wie ein sehr, sehr komplizierter Fotofilter. Deshalb ist das Ergebnis reproduzierbar: gleiches Foto, gleicher Stil, gleicher Regler → exakt dasselbe Bild. Und deshalb reichen dafür 12 MB statt mehrerer Gigabyte.

Merksatz: Diffusion *erfindet* ein Bild aus Rauschen. Style Transfer *rechnet* dein Bild um. Beides heißt im Alltag „KI macht Bilder“ — technisch sind es zwei völlig verschiedene Welten.

Ehrlich bleiben: Was dieses Netz nicht kann

- **Es weiß nicht, was auf deinem Bild ist.** Es sieht Farbverläufe, Kanten und Texturen — keine Menschen, keine Katzen, keine Gesichter. Deshalb malt es einem Porträt gnadenlos einen Wirbel quer über die Augen, wenn der Stil das so vorgibt. Gesichter werden häufig unheimlich. Das ist kein Fehler, das ist der Kern des Verfahrens: *Textur ja, Bedeutung nein.*
- **Die Auflösung ist klein.** 512 px lange Kante — genug für ein Handy-Hintergrundbild, zu wenig für einen Druck. Größer geht technisch, dauert aber quadratisch länger.
- **Manche Stile funktionieren einfach nicht.** Sehr feine Bleistiftzeichnungen oder Bilder mit nur einer Farbe ergeben oft nur einen flauen Farbstich. Der Stil-Vektor ist eine Zusammenfassung — und manche Stile lassen sich in 100 Zahlen schlecht zusammenfassen.
- **Es ist ein Verfahren von 2018.** Damals war das ein Durchbruch (die Vorgänger brauchten pro Bild eine minutenlange Optimierung, und pro Stil ein eigenes Netz). Heutige Bild-KIs aus Kapitel 340 können unvergleichlich mehr — dafür laufen sie nicht in deinem Browser, nicht offline, nicht in Sekunden und nicht ohne dein Bild an einen Server zu geben. **Klein und spezialisiert schlägt groß und universell, wenn die Aufgabe eng genug ist.** Das ist die Lektion, die durch die ganze 4xx-Reihe geht.

Zurück zur Anwendung

Erinnerst du dich an die Selfie-Apps vom Anfang — ein Klick auf „van Gogh“, und im Kleingedruckten steht „Ihr Foto wird zur Verarbeitung an unsere Server übertragen“? Du hast dasselbe eben ohne diesen Satz gemacht: dein Foto, dein Gerät, dein Ergebnis. Und du siehst jetzt, was die Apps verstecken — den Stil-Vektor aus 100 Zahlen, den Regler als Wanderung zwischen zwei Punkten im Latent Space und den Unterschied zu einem Diffusionsmodell, das nicht umrechnet, sondern aus Rauschen erfindet.

Damit endet die Werkzeug-Reihe: Von deiner eigenen CSV-Tabelle bis zu deinem eigenen Gemälde lief alles im Browser, alles auf deinen Daten, alles auf Tensoren. Was du ab hier noch anschaust, entscheidest du selbst — zurück zur Übersicht, dort liegen die Kapitel der anderen Kategorien, die das *Warum* hinter dem Werkzeug erklären.

Zum Weiterlesen

- Gatys u. a. 2015, *„A Neural Algorithm of Artistic Style“* — arxiv.org/abs/1508.06576 — der Ursprung des Verfahrens, damals noch mit minutenlanger Optimierung pro Einzelbild.
- Johnson u. a. 2016, *„Perceptual Losses for Real-Time Style Transfer“* — arxiv.org/abs/1603.08155 — der Schritt, der aus Minuten einen einzigen Netzdurchlauf machte — allerdings noch ein Netz pro Stil.
- Ghiasi u. a. 2017, *„Exploring the structure of a real-time, arbitrary neural artistic stylization network“* — arxiv.org/abs/1705.06830 — genau das Verfahren aus diesem Kapitel: ein Stil-Leser, ein Maler, beliebige Stile.
- Reiichiro Nakano, *arbitrary-image-stylization-tfjs* — github.com/reiinakano (github.com/reiinakano/arbitrary-image-stylization-tfjs) — die Browser-Umsetzung, deren Modelle hier gerade dein Bild verwandelt haben.

Kurzer Check – mit Lösung

Lösungen für Lehrkräfte – für Lernende die Online-Version nutzen.

Frage 1: Das Stil-Netz macht aus dem Gemälde einen Vektor mit 100 Zahlen. Was steht da drin?

- Eine stark komprimierte Version des Gemäldes — also das Bild selbst in klein.
- **✓ Eine Beschreibung des Malstils: Farbwelt, Pinselstruktur, Kontrastverhalten — nicht das Motiv.**
- Die 100 wichtigsten Pixel des Gemäldes.

Aus dem Vektor lässt sich die Sternennacht nicht rekonstruieren — die Sterne, der Kirchturm, die Zypresse sind weg. Was bleibt, ist das Wie: die wirbelnden Striche, das Blau-Gelb, die dicke Farbe. Genau deshalb kann das zweite Netz diesen Stil auf ein völlig anderes Motiv legen.

Frage 2: Warum lässt sich der Stil per Regler stufenlos dosieren, statt nur an/aus?

- Das Ergebnisbild wird am Ende halbtransparent über das Original gelegt.
- Es werden 75 der 100 Zahlen übernommen und 25 weggelassen.
- **✓ Weil zwischen dem Stil-Vektor und dem Vektor deines Fotos interpoliert wird — man wandert im Latent Space von einem Punkt zum anderen.**

Die Rechnung ist stil·a + inhalt·(1−a) auf allen 100 Zahlen gleichzeitig. Das Ergebnis ist ein neuer, gültiger Punkt im Stil-Raum — ein Stil, der nie gemalt wurde. Genau diese Eigenschaft hat dir Kapitel 310 versprochen: Zwischen sinnvollen Punkten liegen wieder sinnvolle Punkte.

Frage 3: Warum läuft das hier in Sekunden im Browser, während eine Bild-KI wie in Kapitel 340 ein Rechenzentrum braucht?

- **✓ Weil hier ein kleines, spezialisiertes Netz das Bild in einem einzigen Durchlauf umrechnet — statt ein Bild in vielen Schritten aus Rauschen zu erzeugen.**
- Weil der Browser insgeheim doch einen Server um Hilfe bittet.
- Weil moderne Handys inzwischen so schnell sind wie Rechenzentren.

Ein Diffusionsmodell durchläuft sein (riesiges) Netz pro Bild zwanzig- bis fünfzigmal und muss dabei alles Mögliche malen können. Dieses Netz hier ist winzig, kann genau eine Sache — Textur übertragen — und braucht dafür genau einen Durchlauf. Enge Aufgabe, kleines Modell, lokale Ausführung: die Rechnung, die die ganze Kategorie „Interaktiv“ trägt.

Geschafft — das war die Werkzeug-Reihe

Du hast dich durch acht Werkstätten gearbeitet, und in jeder einzelnen lief die KI **auf deinem eigenen Gerät, mit deinen eigenen Daten**:

- 410 — deine CSV-Tabelle, ein trainiertes Modell auf deinen Zahlen
- 420 — ein Netz, das du selbst trainiert hast
- 430 — Rahmen um alles, was die KI im Bild findet
- 440 — dein Körper als 17 Punkte
- 450 — deine Stimme als Bild aus Frequenzen
- 460 — deine eigene Bild-KI in zwei Minuten
- 470 — Suche, die Bedeutung versteht statt Buchstaben
- **480** — dein Foto als Gemälde

Von deiner CSV bis zu deinem Gemälde — alles im Browser, alles Tensoren. Jede dieser acht Anwendungen sah von außen aus wie Magie und war von innen dasselbe: Zahlen rein, Matrizen multiplizieren, Zahlen raus. Und du weißt jetzt, wo man hinschauen muss, damit das sichtbar wird.

Wenn du das *Warum* hinter dem heutigen Kapitel noch weiter ausreizen willst: 310 — Der Latent Space erklärt den Raum, in dem dein Stil-Vektor lebt, und 340 — Wie Bild-KIs malen zeigt den ganz anderen Weg, auf dem heutige Bildgeneratoren ans Ziel kommen.

→ Und für den Überblick, was noch alles wartet: zurück zur Kurs-Startseite. Bis dahin — HerrKI

Das nimmst du mit

- **Zwei kleine Netze teilen sich die Arbeit.** Eines liest den Stil aus dem Gemälde, das andere malt dein Foto damit neu.
- **Ein kompletter Malstil passt in 100 Zahlen.** Der Stil-Vektor ist eine Zusammenfassung — und deshalb funktionieren manche Stile gut und andere gar nicht.
- **Der Regler mischt im Latent Space.** Auch dein Foto bekommt einen eigenen 100er-Vektor; die Stilstärke wandert auf der Geraden zwischen beiden Punkten.
- **Das ist kein Diffusionsmodell.** Ein Durchlauf statt Dutzender Entrausch-Schritte — Diffusion *erfindet* ein Bild, Style Transfer *rechnet* deines um.
- **Textur ja, Bedeutung nein.** Das Netz sieht Farbverläufe, Kanten und Muster, keine Gesichter — deshalb legt es den Wirbel gnadenlos quer über die Augen.

So zitierst du dieses Kapitel

HerrKI: Dein Foto als Kunstwerk (Style Transfer). herrki.de/480_style_transfer, ein privates Hobbyprojekt von Terragon.de, abgerufen am 09.09.2026.