
Einflussnahme von bildgenerativer Künstlicher Intelligenz auf den Künstlerberuf und den Kunstbegriff

Bachelorarbeit zur Erlangung des akademischen Grades

Bachelor of Science

im Studiengang Informatik (Bachelor-Studiengang, Vollzeitstudium)

an der Fakultät für Informatik und Ingenieurwissenschaften

der Technischen Hochschule Köln

vorgelegt von: Vanessa Maria Ommer
Matrikel-Nr.: 11139658
Adresse: Mosacherweg 48A
92224 Amberg
vanessa_maria.ommer@smail.th-koeln.de

eingereicht bei: Prof. Dr. Wolfgang Konen

Zweitgutachter*in: Prof. Dr. Lutz Köhler

Gummersbach, 04.02.2025

Kurzfassung/Abstract

In dieser Bachelorarbeit soll herausgearbeitet werden, inwiefern moderne KI-Bildgeneratoren, wie DALL-E 3 [Betker] [OpenAI23] und Stable Diffusion [Esser24], Kunst und Künstler beeinflussen könnten. Dazu wird die Funktionsweise von KI-Bildgeneratoren, die Entwicklung, Anwendung und deren Limitierungen mit Hilfe von Literatur und Dokumentationen beleuchtet. Hierbei werden auch verschiedene Arten dieser KI-Systeme beschrieben und ein Modell wird durch Parametertuning noch detaillierter analysiert. Des Weiteren soll, mit Hilfe von einem maschinellen und einem mit menschlichen Probanden durchgeführten Test, annähernd herausgefunden werden, wie nah KI-Bilder an menschliche Kunst herankommt beziehungsweise, ob diese noch unterscheidbar ist. Zudem wurden persönlich Künstler zu ihrer Meinung zu KI-Bildgeneratoren befragt und auch Meinungen und Kritikpunkte aus der Literatur und Artikeln gesammelt. Anschließend werden die gesammelten Erkenntnisse zusammengetragen und miteinander abgewogen, um die Forschungsfrage zu beantworten.

Schlagwörter/Schlüsselwörter: Künstliche Intelligenz, Ethik, Kunst, Deep-Learning, generative KI, Bildgenerierung

Inhalt

Inhalt.....	II
Tabellenverzeichnis.....	III
Abbildungsverzeichnis	IV
1 Einleitung.....	1
2 Methodik.....	4
3 KI-Bildgeneratoren.....	7
3.1 Begriffserklärung	7
3.2 Meilensteine	7
3.3 Modelle	9
3.3.1 Variational Autoencoders	9
3.3.2 Generative Adversarial Networks	10
3.3.3 Diffusionsbasierte Modelle	11
3.4 Anwendungsgebiete.....	12
3.5 Limitierungen.....	13
4 Demokratisierung von Kunst durch KI.....	15
4.1 Begriffsdefinitionen.....	15
4.1.1 Kunst.....	15
4.1.2 Künstlerberuf.....	15
4.2 Möglichkeiten der Demokratisierung	16
5 KI-Bilder von Zeichnungen unterscheiden	21
5.1 Schwierigkeiten bei der Erkennung	21
5.2 Existierende Tools und Methoden	22
5.3 Untersuchung zur KI-Erkennung.....	23
6 Kritik an der Verwendung von KI-Bildgeneratoren.....	29
6.1 Gespräche mit Künstlern.....	29
6.2 Vorurteile durch Trainingsdatensätze	30
6.3 Umweltaspekt.....	31
6.4 Urheberrecht.....	32
7 Diskussion	34
8 Fazit	39
Literaturverzeichnis.....	41
Anhang.....	46
Erklärung	47

Tabellenverzeichnis

Tabelle 1 Prozentuale Einschätzung vom Hive Moderation Tool der KI-Wahrscheinlichkeit für gezeichnete Bilder	23
Tabelle 2 Prozentuale Einschätzung vom Hive Moderation Tool der KI-Wahrscheinlichkeit für tatsächliche KI-Bilder	23
Tabelle 3 Anzahl der richtigen Erkennungen von KI-Bildern im Test zur Untersuchung menschlicher Strategien für KI-Bild-Erkennung. Die Tabelle gibt an wie oft die einzelnen KI-Bilder als diese richtig identifiziert und wie oft diese fälschlicherweise für Zeichnungen gehalten wurden.	27
Tabelle 4 Anzahl der richtigen Erkennung von Zeichnungen im Test zur Untersuchung menschlicher Strategien für KI-Bild-Erkennung. Hierbei beschreibt „Richtig“ die Anzahl von Zeichnungen, die als solche erkannt wurden und „Falsch“ beschreibt die Anzahl an Malen, bei denen Probanden die Zeichnungen für Produkte eines KI-Bildgenerators hielten.	27

Abbildungsverzeichnis

Abbildung 1 Meilensteine in der Entwicklung der Bildgeneratoren	7
Abbildung 2 Vereinfachte Darstellung einer VAE-Architektur. [Eugenio21]	9
Abbildung 3 GAN-Architektur vereinfacht dargestellt. [Corpnce].....	10
Abbildung 4 Von DALL-E generierte Bilder mit ihren Prompts darunter , als Beispiele für häufige Fehler des Modells. So sind im ersten Bild Hund und Katze vertauscht und die Katze schaut nicht wie gewollt den Hund an. Im zweiten Bild beschreibt der Prompt ein Pferd, das einen Astronauten reitet, aber das Modell hat ein Bild mit der umgekehrten Beschreibung geliefert. Im dritten Bild halluziniert DALL-E eine Blumenspezies, statt ein Bild der Blume mit dem vorgegebenen lateinischen Namen zu generieren. [Betker]	13
Abbildung 5 Echtes Bild einer Arum Dioscoridis. Zum Aufzeigen der Unterschiede zur Halluzination in Abbildung 4. [Meeuwissen]	14
Abbildung 6 Ergebnisse des GANs mit Standardparametern nach 1000 Epochen...	18
Abbildung 7 Generierte Bilder bei um 0,0002 erhöhter Diskriminator-Lernrate zu verschiedenen Zeiten des Trainings	19
Abbildung 8 Verlust- und Genauigkeitsverlauf von Trainingsdurchgang 2	19
Abbildung 9 Generierungen bei einer Dropout-Rate des Diskriminators von 0,2 und einer Generator-Lernrate von 0,0003	19
Abbildung 10 Beispielhaft mit Stable Diffusion [SDOnline] generiertes Bild, welches zeigen soll, das KI-Systeme oft Bilder von Menschen mit falscher Fingeranzahl generieren	21
Abbildung 11 Beispiel von Augen einer mit Stable Diffusion [SDOnline] generierten Animefigur, für das ganze Bild siehe im angehängten [Repository] unter AIRecognitionTest\popartImages\ai_images\ai_25.png.....	22
Abbildung 12 Diagramm zur Visualisierung der Sicherheit beim Erkennen von KI-Bilder durch Hive Moderation. In dem Diagramm ist dargestellt zu wie viel Prozent Hive Moderation Zeichnungen (blaue Punkte) und tatsächliche KI-Bilder (orange Punkte) für von KI generiert gehalten hat.....	24
Abbildung 13 Gezeichnete Bilder von der gleichen Künstlerin, welche Hive Moderation stark unterschiedlich eingeschätzt wurden	24
Abbildung 14 Mit DALL-E generiertes Bild, was von Hive Moderation nur zu 97,3% als KI-generiert erkannt wurde.....	25
Abbildung 15 Grafische Darstellung der Fehlerquote von Hive Moderation beim Bestimmen der verwendeten Modelle. In diesem Diagramm sind dabei nur falschen Modelleinschätzungen aufgezeigt und nicht die richtigen. Diese Fehlerquote hat auch keinen Bezug auf die Erkennung von KI an sich	25

Abbildung 16 Ergebnisse des Tests, bei dem Menschen mit unterschiedlicher Kunst-Erfahrung KI-Bilder von nicht KI-Bildern unterscheiden sollten. Im Diagramm sieht man, wie viele Teilnehmer, welche Genauigkeit im Test erreicht haben... 26	26
Abbildung 17 KI-Bild welches im Test ausnahmslos von allen Probanden als von einer KI generiertes Bild erkannt wurde..... 26	26
Abbildung 18 KI-Bild das am häufigsten für eine Zeichnung gehalten wurde..... 27	27
Abbildung 19 CO ₂ -Verbrauch verschiedener KI-Anwendungen in der Inferenz im Vergleich. [Luccioni24]..... 31	31
Abbildung 20 CO ₂ -Verbrauch verschiedener KI-Anwendungen in der Inferenz mit Bezug auf die Parameteranzahl der Modelle im Vergleich. [Luccioni24]..... 31	31
Abbildung 21 Ausgaben von DALL-E 3 für den Prompt: "An individual enjoying a leisurely picnic in the park, with an array of snacks spread out on a checkered blanket" (Eine Person genießt ein gemütliches Picknick im Park, mit einer Auswahl an Snacks, die auf einer karierten Decke ausgebreitet sind) am Anfang der Entwicklung und bei Veröffentlichung des Modells. [OpenAI23] 34	34
Abbildung 22 Eigens mit DALL-E 3, welches in [ChatGPT]integriert ist, generierte Testbilder zum Prompt aus Abbildung 21. Dabei wurde der Prompt in (b) und (c) von [ChatGPT] erweitert und bei (c) wurde explizit nach einer Illustration gefragt 34	34
Abbildung 23 Beispiel von Generierung eines Bildes durch Stable Diffusion von einer Einsatzkraft der Feuerwehr, die eine Katze rettet. Einmal ohne Geschlechtsangabe und einmal mit weiblicher Geschlechtsangabe. Die genauen Prompts stehen unter den Bildern 35	35

1 Einleitung

Im Netz liest man immer mehr Kritik zu KI-Bildgeneratoren, so gibt es ganze Profile in den sozialen Medien, die „Nein zu KI-Kunst“ sagen, wie zum Beispiel [X1]. Des Weiteren kritisieren Künstler auch in einzelnen Posts die Nutzung von KI-Bildgeneratoren. So protestierten Künstler zum Beispiel auf der bekannten Portfolio-Seite ArtStation gegen KI-Bilder und überschwemmten die Seite mit Anti-KI-Bildern, siehe hierzu [Edwards22]. Bei dem Protest kritisierten die Künstler die Nutzung ihrer eigenen Werke zum Trainieren von KI und auch das Teilen von KI-Bildern seitens mancher Künstler in ihren Portfolios auf der Webseite. Die Protestaktion erreichte ein so großes Ausmaß, dass die Portfolio-Webseite reagieren musste. Diese bannte KI-Bilder zwar nicht von ihrer Webseite, fügte jedoch einen Tag hinzu, mit dem Künstler selbst entscheiden können sollen, ob ihre Kunst zum Trainieren von KI-Modellen verwendet werden darf.

Die Kritiker sehen KI-generierte Bilder nicht als gleichwertig zu von Menschen gezeichneten Bildern an. Dies begründen sie vor allem damit, dass der Arbeitsaufwand bei gezeichneten Bildern so viel höher sei, und äußern im Rahmen der Kritik ihre ethischen und rechtlichen Bedenken zu KI-Bildgeneratoren und deren Nutzung. Die sogenannten Text-zu-Bild-Modelle werden derweilen immer fortschrittlicher und finden auch immer öfter Verwendung durch Unternehmen, vor allem im Marketing. So starteten 2024 unter anderen auch viele große Marken, wie zum Beispiel Coca-Cola und Ferrero, Kampagnen mit von KI generierten Bildern, siehe dazu [DataFeedWatch24]. Diese Umstellung von der Nutzung von KI-Bildgeneratoren statt der Beauftragung von Illustratoren, nehmen einige Künstler als Bedrohung wahr. Besonders in dem Hinblick darauf, dass die verwendeten KI-Technologien immer ausgereifter werden und stetig weiterentwickelt werden. Hierbei sind diffusionsbasierte Modelle der aktuelle Stand der Technik und lassen sich in Anwendungen wie DALL-E [Betker] [OpenAI23] von OpenAI, Stable Diffusion [Esser24] von StabilityAI und Imagen [Imagen3] von Google wiederfinden.

Eine künstliche Intelligenz mit der Fähigkeit zur echten Kreativität auszustatten, sei nach [Foster20] ein wichtiger Schritt der Forschung. Doch auch wenn oft der Begriff KI-Kunst, beziehungsweise AI-Art im englischen für KI-Bilder verwendet wird, ist strittig, ob es sich dabei nach Definition wirklich um Kunst handelt, und ob man dabei von einem kreativen Prozess sprechen kann. Kritiker sagen nein und werfen der KI vor lediglich gelerntes teilweise zu kopieren und neu zusammensetzen. Neben der Kritik steigt aber auch das Interesse an der vergleichsweise neuen Technologie. So schätzt [Everypixel24], dass insgesamt pro Tag über 3 Millionen Bilder mit KI generiert werden.

In dieser Bachelorarbeit soll es darum gehen, wie KI-Modelle Bilder generieren und wie diese Technologie die Künstlerszene revolutionieren könnte. Dabei soll versucht werden, die Frage zu beantworten, inwiefern moderne KI-Bildgeneratoren, wie Stable Diffusion [Esser24] und Dall-E [Betker][OpenAI23] das Verständnis der Begriffe von Kunst und Künstlern beeinflussen. Es soll annähernd ergründet werden, wie KI-Kunst die Kreativität erweitert oder vermindert und welche Auswirkungen das auf Künstler haben könnte. Dazu sollen Argumente aus der Literatur und von Künstlern zusammengetragen

werden und technische Möglichkeiten der Modelle untersucht werden. Diese Ergebnisse sollen anschließend gegeneinander aufgewogen werden, um zu ermitteln, in welchem Ausmaß die Kritik berechtigt ist und auch wie Künstler, die mit ihrer Kunst Geld verdienen sich auf die neue Situation einrichten müssen.

Zunächst sollen die Methodiken, die für diese Arbeit angewendet wurden, erklärt werden. Anschließend soll im theoretischen Rahmen dieser Arbeit zunächst die Technik hinter KI-Bildgeneratoren erläutert werden. Dazu wird zunächst der Begriff einmal grob erklärt. Dann werden verschiedene Meilensteine aufgelistet. Dies soll verdeutlichen, wie schnell die Technik hinter KI-Bildgeneratoren sich in den letzten Jahren entwickelt hat. Danach werden die drei wichtigsten Modellarten von KI-Bildgeneratoren erklärt. Die Variational Autoencoders als Grundlage, die Generative Adversarial Networks als großen Durchbruch und der heutige Stand der Technik, diffusionsbasierte Modelle. Nachdem die grundlegenden Modelle vorgestellt sind, werden Beispiele für Anwendungsmöglichkeiten der Systeme aufgezählt und abschließend Limitierungen der Technik beschrieben.

Nach den technischen Rahmenbedingungen von Text-zu-Bild-KI-Modellen, sollen die Möglichkeiten der Demokratisierung von Kunst durch diese Systeme erläutert werden. Hierzu folgt zunächst eine Definition von Begriffen, die für weitere Erläuterungen dieser Arbeit wichtig sind. Als erstes soll die offizielle Definition von Kunst geklärt werden, um verstehen zu können, ob KI-Bildgeneratoren überhaupt Kunstwerke erschaffen können. Danach wird der Begriff Künstler für diese Arbeit definiert und festgelegt, welche Leute gemeint sind, wenn dieser Ausdruck im weiteren Verlauf der Arbeit verwendet wird. Zuletzt werden in dem Abschnitt dann Argumente wiedergegeben, die für eine Demokratisierung von Kunst sprechen.

Um den Ressourcenaufwand und die Komplexität des Trainings eines KI-Bildgenerators besser nachvollziehen zu können, werden dann die Ergebnisse aus einer Parameter-tuning mit einem beispielhaftem GAN vorgestellt, welches im Rahmen dieser Arbeit durchgeführt wurde.

Im Anschluss daran folgt eine Beleuchtung hinter die Schwierigkeiten KI-Bilder von menschengemachten Illustrationen zu unterscheiden. Hierbei werden zunächst die auftretenden Probleme beschrieben. Fokus liegt dabei besonders auf Eigenschaften, die bei der Erkennung von fotorealistischen KI-Bildern anders sind. Nach der Vorstellung der Schwierigkeiten, sollen Möglichkeiten zur Erkennung von KI-Illustrationen erläutert werden. Dabei wird, sowohl auf technische Möglichkeiten, wie auch auf Strategien zum Erkennen von KI-Bildern mit dem bloßen Auge, die aus Interviews mit Künstlern hervorgegangen sind, eingegangen. Der Abschnitt endet mit dem Testen dieser Methoden. Dabei wird ein Tool zur KI-Erkennung auf die Probe gestellt und außerdem wird ein Test mit Freiwilligen durchgeführt. Durch diese Tests soll festgestellt werden, wie zuverlässig das Tool und die Strategien sein können und auch geprüft werden, ob die befragten Künstler ihre Fähigkeiten zur Erkennung gut eingeschätzt oder übertrieben haben.

Daraufhin folgt die Sammlung von Kritik an Modellen, wie DALL-E [Betker] [OpenAI23] und ähnlichen Anwendungen. Hierzu werden zunächst Kritikpunkte aus persönlichen Gesprächen gezogen und besprochen. Darüber hinaus werden die drei Hauptkritikpunkte aus den Gesprächen und der Literatur genauer betrachtet. So soll erläutert werden, wie und welche Vorurteile aus Trainingsdatensätzen von den KI-Modellen übernommen werden und welche Auswirkungen daraus folgen können. Außerdem soll der Einfluss auf die Umwelt durch das Training und die Nutzung von KI-Bildgeneratoren beschrieben werden, und erklärt werden, warum der Einfluss auf die Umwelt dieser Modelle unter Umständen größer ist, als bei anderen KI-Systemen. Abgeschlossen wird das Kapitel mit den Kritikpunkten zum Urheberrecht und einer Erläuterung, weshalb gleich mehrere Urheberrechtsfragen bei der Nutzung und dem Vertrieb von KI-Bildgeneratoren aufkommen und was das Ganze so problematisch macht. In diesem Abschnitt sollen alle Kritikpunkte mit Fakten belegt werden, aber noch nicht weiter gewichtet werden. Dies geschieht im folgenden Kapitel, der Diskussion.

In diesem Abschnitt wird die Kritik positiven Argumenten entgegengestellt. Des Weiteren werden existierende und mögliche Lösungsansätze besprochen, sowie bewertet. Die Ergebnisse der Diskussion fließen dann abschließend in das Fazit mit ein, wo nach einer Zusammenfassung der im Laufe dieser Arbeit gesammelten Erkenntnisse, versucht werden soll die Forschungsfrage, inwiefern das Verständnis von Kunst und Künstler durch moderne KI-Bildgeneratoren wie DALL-E [Betker] [OpenAI23] und Stable Diffusion [Esser24] beeinflusst werden kann, zu beantworten. Außerdem soll eine mögliche zukünftige Entwicklung in Aussicht gegeben werden, um die Bachelorarbeit abzuschließen.

2 Verwendete Methodik

In dieser Arbeit soll die Frage beantwortet werden, inwiefern moderne KI-Bildgeneratoren, wie Stable Diffusion und DALL-E das Verständnis von Kunst und Künstler beeinflussen. Hierzu wurde sich zunächst einmal theoretisch mit dem Thema befasst. Dazu wurden Auszüge aus Fachliteratur und Dokumentationen zu KI-Bildgeneratoren und auch Artikel und Literaturanalysen vom Bezug dieser Technologien auf Kunst und Künstler gesammelt und durchgearbeitet. Die Quellen wurden zunächst über Google Scholar gesucht und dann die Schneeballmethode angewandt. Die Webseiten von etablierten KI-Bildgeneratoren boten auch einen guten Anhaltspunkt, vor allem, um die Dokumentationen dieser Modelle zu finden. Das aktuelle Interesse an der Forschung erleichterte die Suche und sorgte dafür, dass viele aktuelle Quellen herangezogen werden konnten. Bei den kritischen Artikeln zum Thema konnte schnell festgestellt werden, wie das Arbeitsfeld der Autoren deren Meinung zu beeinflussen schien. So wirkten aus der Technik stammenden Autoren meist eher positiv gegenüber KI-Bildgeneratoren gestimmt und Künstler schienen die Sache oft eher kritisch zu betrachten. Die Ergebnisse werden in verschiedenen thematisch getrennten Abschnitten dieser Arbeit zunächst zusammengefasst und im Fall der Erkenntnisse zur Bewertung von KI-Bildgeneratoren noch eingeordnet und mit anderen Erkenntnissen, die im Laufe dieser Arbeit gewonnen wurden, verglichen. Die Literaturerkenntnisse zur Funktionsweise von KI-Bildgeneratoren findet sich in Kapitel 3. In Abschnitt 4 wird zunächst der Bezug zur Kunst und die daraus entstehenden positiven Möglichkeiten beschrieben. Mit der Kritik zu Bildgeneratoren und die negativen Einflüsse auf die Kunst wird in Kapitel 7 eingegangen.

Um die Theorie hinter KI-Bildgeneratoren besser zu verstehen und auch, um zu prüfen, wie viel Aufwand das Training eines solchen Modells bedeuten würde, wurde dazu ein Parametertuning eines einfachen KI-Bildgenerators durchgeführt. Dieses wird in Abschnitt 5 beschrieben und ausgewertet. Das für das Experiment verwendete Modell stammt von [Foster20] und ist ein GAN welches mit einer Untermenge des Quick, Draw! Datensatz von [Google17] trainiert wird. Hierbei wurde die ungefähre Dauer des Trainings und die Trainingsergebnisse festgehalten. Letztere wurden sowohl numerisch, als auch grafisch aufgenommen. Außerdem wurden die verwendeten Parameter für verschiedene Trainingsdurchläufe notiert und zusammen mit den Ergebnissen gespeichert. Natürlich wurden hierbei auch die erzeugten Bilder gesichert. Das Parametertuning wurde auf einer Nvidia Geforce RTX 3070ti durchgeführt.

Als weiteres Experiment wurde Hive Moderation, ein Tool zur Erkennung von KI-Bildern, getestet, da sich während der Recherchen die Frage stellte, wie man KI-Bilder von Nicht-KI-Bildern zuverlässig unterscheiden kann. Dabei wurde die Hive Moderation Seite [HiveTool] gefunden und es blieb zu testen, wie gut die Anwendung funktionierte. Zum Testen des Tools wurden 54 KI-Bilder mit DALL-E [Betker] [OpenAI23] [Betker][Betker]und Stable Diffusion [Esser24] in den kostenlosen Versionen erstellt und 50 Bilder von menschlichen Künstlern gesammelt. Erstellt werden konnten pro Tag mit den kostenlosen Versionen der KI-Modelle maximal 13 Bilder. Die Generierung zog sich also

über 4 Tage. Nachdem alle über 50 KI-Bilder generiert waren, wurde von allen insgesamt 104 Bildern eine Bildschirmaufnahme gemacht, um die Metadaten zu entfernen. Hierbei wurden bei den KI-Bildern auch vorhandene Wasserzeichen entfernt. Bei Stable Diffusion befinden sich die Wasserzeichen in der rechten, unteren Ecke, diese wurde also einfach abgeschnitten. Bei den Illustrationen von Künstlern ließen sich Unterschriften nicht immer abschneiden, da sie sich vereinzelt in der Mitte des Bildes befanden. Die verwendeten Bilder sind öffentlich auf Instagram zu finden oder von befreundeten Künstlern, für die Quellenliste siehe bitte das angehängte [Repository]. Die gesammelten Screenshots wurden dann in einer zufälligen Reihenfolge angeordnet und in das Hive Moderation Tool hochgeladen. Die Ergebnisse des Tests wurden tabellarisch notiert und händisch ausgewertet, dies wird in Abschnitt 6.3 beschrieben.

Ein Auszug der gleichen Bilder wurde für einen Test mit acht Probanden verwendet, die die Aufgabe hatten für jedes Bild zu entscheiden, ob es sich um ein digital gezeichnetes Bild oder um ein Bild, welches von einer KI generiert wurde, handle. Dazu wurde eine Google-Umfrage mit 10 ausgewählten Bildern, von denen 5 KI-generiert und 5 gezeichnet waren, erstellt. Diese Umfrage wurde an die Probanden geschickt, von denen die meisten selbst Künstler waren. Der Test wurde anonym durch die Quiz-Funktion von Google-Forms ausgewertet und eine Person wurde manuell befragt. Die Probanden beantworteten die Anfrage alle sehr zeitnah und viele gaben von sich aus Feedback zur Schwierigkeit des Tests zurück. Die Erkenntnisse des Versuchs finden sich auch zusammen mit dem Test von Hive Moderation in Abschnitt 6.3 wieder.

Zu Beginn der Durchführung dieser Arbeit wurden 4 semistrukturierte Interviews mit Künstlern zum Thema KI-Bildgeneratoren geführt. Ziel war es die Meinung der Künstler zu der noch neuartigen Technologie herauszufinden und festzustellen, inwieweit sich die Befragten von den KI-Modellen in ihrem Schaffen beeinflusst fühlen. Es wurden 15 Hauptfragen gestellt. Die Interviews wurden einzeln in Discord-Anrufen geführt und die Teilnehmer wurden gefragt, ob das Gespräch aufgezeichnet werden dürfe. Die Aufnahmen wurden anschließend transkribiert. Die bereinigten Transkripte befinden sich im angehängten [Repository] und die Auszüge der Interviews fließen in verschiedene Kapitel mit ein. Der Hauptpunkt der Interviews, die Kritik der befragten Künstler wird ausführlich in Kapitel 7.1 beschrieben. Befragt wurden zwei langjährige Online-Künstlerinnen, die schon lange in der Online-Kunstszene unterwegs sind, eine Designstudentin, die ebenfalls schon länger in der Online-Szene aktiv ist und ein Hobbykünstler, der erst vor kurzem angefangen hat, sich mit den richtigen Techniken auseinander zu setzen, aber auch schon eine Weile Künstlern in den sozialen Medien folgt. Besonders interessant bei der befragten Designstudentin ist, dass sie bald mit ihrem Studium fertig ist und voraussichtlich bald hauptberuflich mit ihren Zeichnungen Geld verdienen wird, was sie in direkte Konkurrenz mit KI-Bildgeneratoren setzt. Die Befragten wurden über Instagram und WhatsApp angefragt. Es wurden drei Einzelanfragen gemacht und eine Instagram-Story als Anfrage formuliert. Auf den Aufruf in den sozialen Medien meldeten sich zwei Künstler positiv, wie auch auf die Einzelanfragen. Die meisten Fragen wurden von den Künstlern sehr ausführlich beantwortet und die Interviews dauerten alle etwa 30 Minuten. Zur

Forschung für diese Arbeit wurden also durch verschiedene Methoden vor allem qualitative Daten gesammelt. Außerdem wurde deduktiv vorgegangen, um die Forschungsfrage zu beantworten.

3 KI-Bildgeneratoren

3.1 Begriffserklärung

Der Begriff KI-Bildgeneratoren wird für KI-Systeme verwendet, die durch Diffusion oder ähnliche Verfahren dazu in der Lage sind, durch das Trainieren mit Bildern, neuartige Bilder von sich aus zu generieren. KI-Bildgeneratoren können auf bestimmte Stile trainiert werden, dann erschafft die KI ein völlig neues Bild in dem Stil des Datensatzes. Weiterhin gibt es auch die Möglichkeit, dass aus sogenannten Text-Prompts, also einer textuellen, meist stichwortartigen Beschreibung, Bilder generiert werden. Im Internet finden sich zahlreiche kostenlose und kostenpflichtige Anwendungen dafür. Die bekanntesten Anwendungen sind wohl DALL-E 3 [Betker] [OpenAI23] von OpenAI, Stable Diffusion [Esser24] von Stability AI, Midjourney [Midjourney] und Imagen [Imagen3] von Google. Aktuelle Text-zu-Bild-Systeme, die wohl beliebteste Art der KI-Bildgeneratoren momentan, sind vermutlich alle diffusionsbasierte Modellsysteme und wurden mit Millionen Bildern trainiert. Da manche Unternehmen wie zum Beispiel Midjourney ihre Modellarchitektur und noch mehr Unternehmen ihre verwendeten Trainingsdaten geheim halten, lassen sich diese Aussagen nicht genau belegen. Es wird aber von vielen vermutet, dass ähnlich viele Trainingsdaten und diffusionsbasierte Modelle, wie sie bei dem Open-Source-Modell Stable Diffusion [Esser24] zum Einsatz kommen, auch für andere Anwendungen verwendet werden. Die Verwendung von diffusionsbasierten Modellen ist bei den meisten beliebten Modellen auch belegt. Neben diesen Modellen gibt es noch weitere Architekturen, die Bilder generieren können, nennenswert sind hierbei vor allem Variational Autoencoders (VAEs) [Kingma22] und Generative Adversarial Networks (GANs) [Goodfellow14]. Deren Technologie gilt heutzutage aber von diffusionsbasierten Modellen als aktuellen Stand der Technik abgelöst. Da die Technologie der KI-Bildgeneratoren sich insbesondere in den letzten 10 Jahren stark verbessert hat.

3.2 Meilensteine

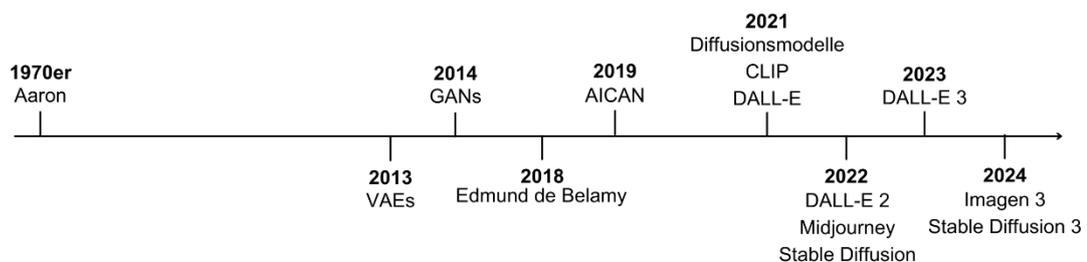


Abbildung 1 Meilensteine in der Entwicklung der Bildgeneratoren

Wie in dem Zeitstrahl aus Abbildung 1 zu sehen, geht die Geschichte von KI-Bildgeneratoren zurück bis in die 70er Jahre, als Harold Cohen das Programm Aaron vorstellte. Aaron war ein KI-System das zunächst nur einfache Schwarz-Weiß-Zeichnungen hervorbrachte, später aber von Cohen weiterentwickelt wurde, siehe hierzu [Rombach22], [Srinivasan21] und [Vass24]. [Grba22] beschreibt Cohen und andere Künstler, die mit

KI-Kunst experimentierten, als eine Minderheit ihrer Zeit. Dies lässt sich nach Grba auf den KI-Winter in dieser Zeit und der darauffolgenden geringen Verfügbarkeit von KI-Modellen und Forschungsgeldern zurückführen. Wie schon in Abschnitt 3.1 angerissen, lieferten aber vor allem die letzten Jahre entscheidende Fortschritte in der Entwicklung der KI-Bildgeneratoren. So stellten Kingma und Welling 2013 VAEs vor, eine neue Art der Bildgeneratoren, die auf DeepLearning basiert [Kingma22]. Nur ein Jahr später wurden die GANs von Ian Goodfellow vorgestellt [Goodfellow14], diese bilden die Grundlage der KI-Bildgeneratoren wie sie heute sind. GANs wurden in den folgenden Jahren ständig erweitert und verbessert. Eine dieser Erweiterungen ist das AICAN vorgestellt von [El-gammal19]. AICAN wurde nicht wie vorherige GANs auf einen Stil oder eine bestimmte Ästhetik trainiert, sondern mit 80.000 Bildern aus verschiedenen Stilen der letzten 5 Jahrzehnte gefüttert. Bei einem Experiment des Autors von AICAN konnten Menschen $\frac{3}{4}$ der Zeit nicht feststellen, dass Bilder durch AICAN generiert und nicht von einem Künstler gemalt wurden.

Letztlich wurden GANs aber von diffusionsbasierten Modellen als State-of-the-Art abgelöst, siehe hierzu auch [Dhariwal21]. In den 2010ern generierten auch einige KI-Kunstprojekte Aufsehen. So zum Beispiel das Kunstwerk „Edmund de Belamy“ von Obvious im Jahr 2018 siehe hierzu [Cetinic22] und [Christie's]. Das Bild wurde durch ein GAN generiert und für 432.500\$ versteigert. Diese Auktion löste, laut [Cetinic22] einen KI-Kunst-Hype aus und führte dazu, dass mehr Künstler, als vorher sich mit dieser Technologie befassten. Nicht verwunderlich also, dass auch die Forschung weiter vorangetrieben wurde. Große Konzerne wie Google und OpenAI arbeiteten die letzten Jahre an immer besser werdenden Methoden zur Bildgenerierung durch KI. So veröffentlichte OpenAI 2021 CLIP (Contrastive Language-Image Pretraining) [Radford21] [OpenAI21], ein multimodales KI-Modell, welches Text- und Bildinformationen verbindet und so eine wichtige Bereicherung zur Entwicklung von KI-Bildgeneratoren bildete. CLIP ist sowohl Bestandteil der Stable Diffusion 3, sowie der DALL-E 2 und 3 Architektur. Letztere enthält CLIP nur noch in unterstützender Funktion. CLIP bleibt aber essenziell für die Qualität der Ausgaben und war ein wichtiger Schritt zur Verbesserung von Text-zu-Bild-Modellen. 2021 markiert auch den Einzug von Diffusionsmodellen in das Forschungsfeld, die wie schon beschrieben, GANs schließlich als State-of-the-Art ablösten. Die bekanntesten KI-Bildgeneratoren, deren Architektur bekannt ist, verwenden Diffusion. Vermutlich enthalten auch moderne Modelle, deren Architektur nicht bekannt ist, wie Midjourney, diffusionsbasierte Elemente. Mittlerweile werden KI-Bildgeneratoren in andere Applikationen eingebettet, so findet sich DALL-E 3 [OpenAI23] in ChatGPT und auch das dieses Jahr veröffentlichte Imagen 3 [Imagen3] wurde von Google in ihren ChatBot, Gemini, integriert. Auch in Programmen wie Photoshop [Adobe] und [Canva] sind KI-Bildgeneratoren heutzutage eingebettet.

Die Entwicklung von KI-Bildgeneratoren ist also durch den Hype der letzten Jahre enorm schnell vorangeschritten. Von VAEs, über GANs zu den heute so beliebten diffusionsbasierten Modellen hat die Forschung in den letzten Jahren rasante Fortschritte erzielt. Die Modelle werden immer besser und sind teilweise vor allem für ungeübte Augen nur

noch schwer von echten Illustrationen unterscheidbar, somit nehmen auch die Auswirkungen auf die Kunstszene immer mehr zu. Im späteren Verlauf der Arbeit sollen diese Auswirkungen weiter diskutiert werden.

3.3 Modelle

3.3.1 Variational Autoencoders

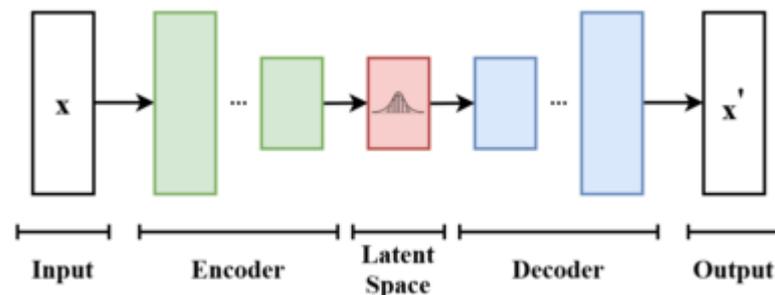


Abbildung 2 Vereinfachte Darstellung einer VAE-Architektur. [Eugenio21]

Wie schon in Abschnitt 3.2 beschrieben, wurden VAEs erstmalig 2013 von Diederik P. Kingma und Max Welling vorgestellt [Kingma22]. VAEs sind eine Untergruppe der Autoencoder. Im Unterschied zu gewöhnlichen Autoencodern sind VAEs probabilistische Modelle, die latenten Variablen werden also nicht als feste, diskrete Werte ausgedrückt, sondern als Wahrscheinlichkeitsverteilungen dargestellt. Daher rekonstruieren VAEs nicht die Originalbilder, wie gewöhnliche Autoencoder, sondern erschaffen neuartige Ergebnisse, die jedoch ähnlich wie die Trainingsdaten aussehen, so [Bergmann24]. Eine vereinfachte Darstellung eines VAEs ist in Abbildung 2 dargestellt.

VAEs bestehen, wie andere Autoencoder aus einem Encoder und einem Decoder und verwenden, wie Autoencoder zur Rekonstruktion von Bildern, den Rekonstruktionsverlust. Der Encoder verdichtet hierbei die Eingabedaten in einen Darstellungsvektor und der Decoder überführt diesen Vektor dann wieder auf die Ursprungsebene, siehe hierzu [Foster20, S.66]. Zusätzlich, erklärt [Bergmann24], enthalten VAEs einen Regularisierungsterm der Kullback-Leibler-Divergenz, die auch KL-Divergenz genannt wird. Dieser Term sorgt dafür, dass die latente Repräsentation der Variablen einer Normalverteilung entspricht, und somit Kontinuität und Vollständigkeit im latenten Raum gegeben sind, was Overfitting des VAEs verhindert und gewährleistet, dass nicht die Bilder aus dem Trainingsdatensatz kopiert werden, sondern ähnliche, neue Bilder entstehen. Für die Variationsinferenz ist die KL-Divergenz allerdings unlösbar, weshalb die Evidenzuntergrenze (ELBO) in VAEs zum Einsatz kommt. Diese bezieht sich hierbei auf eine Art Schätzwert dafür, wie gut das Modell die echten Daten unter Bedingung der KL-Divergenz und des Rekonstruktionsverlustes nachbildet. Dabei ist das Ziel die ELBO zu maximieren, um das Modell zu verbessern, so [Bergmann24]. In ihrer Veröffentlichung beschrieben [Kingma22] auch den Reparametrisierungstrick, englisch: „reparameterization trick“. Dieser ist nach [Bergmann24] ein wichtiges Werkzeug, um die Eingabe zufällig zu

lassen, jedoch weiterhin die Möglichkeit zur Parameterisierung möglich zu machen. Eine gute Eigenschaft von VAEs ist, dass sie einfach zu trainieren sind. Jedoch liefern VAEs eher unscharfe Bilder, so [Bergmann24] und auch [Rombach22] beschreibt die Ergebnisse als schlechter im Vergleich zu GANs. Neben der Bildgenerierung werden VAEs auch in anderen Bereichen eingesetzt, zum Beispiel in der Anomalieerkennung und Arzneimittelmolekül-Erzeugung, so [Bergmann24]. VAEs sind also Bildgeneratoren die auf Deep Learning und der Bayes'schen Statistik basieren und bilden damit eine frühe Grundlage für modernere generative Modelle.

3.3.2 Generative Adversarial Networks

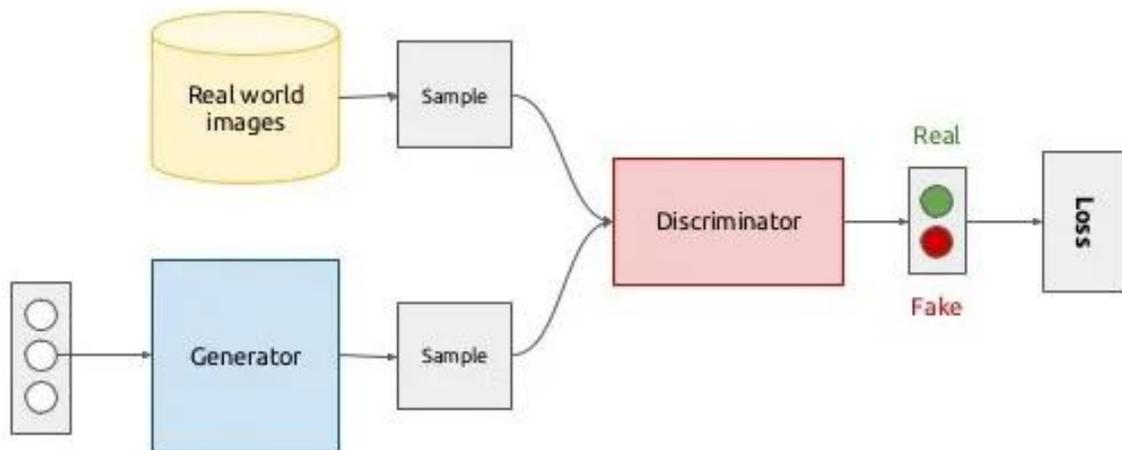


Abbildung 3 GAN-Architektur vereinfacht dargestellt. [Corpnce]

GANs waren lange Zeit die führende Methode zur Bildgenerierung. Erst vor ein paar Jahren wurden sie durch Diffusionsmodelle übertroffen. Wie von [Goodfellow14] beschrieben besteht ein GAN aus zwei Modellen, die gegeneinander trainiert werden. Das erste Modell, der Generator, soll dabei neue Bilder erzeugen, die wie ein Teil des Trainingsdatensatzes aussehen, aber noch nicht in diesem existieren. Das zweite Modell, der Diskriminator wird dann darauf trainiert, echte Bilder des Datensatzes und Fälschungen vom Generator unterscheiden zu können. Die Zielfunktion dieses kompetitiven Trainingsablauf beschreibt [Goodfellow14] in der Theorie mit folgender Funktion:

$$\min_G \max_D V(D, G) = E_{x \sim P_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

$D(x)$ beschreibt hierbei, wie gut der Diskriminator echte Daten erkennt und soll von diesem maximiert werden. $\log(1 - D(G(z)))$ soll vom Generator minimiert werden, was bedeutet der Generator soll im Fälschen besser werden. Eine vereinfachte Darstellung, dieser Funktionsweise, findet sich in Abbildung 3. Diese Optimierungsschritte wechseln sich in der Praxis beim Training ab. Also erst lernt der Diskriminator und im nächsten Schritt erst der Generator. Damit der anfangs schwache Generator in der Praxis überhaupt eine Chance zum Lernen hat, wird dieser darauf trainiert $D(G(z))$ zu maximieren statt $\log(1 - D(G(z)))$ zu minimieren. Dies soll nach [Goodfellow14] bessere

Rückmeldungen geben, auch wenn der Generator noch sehr schwach ist und so das Lernen und die Anpassung beschleunigen. Außerdem führen beide Funktionen zum gleichen Ziel. Da so ein GAN jedoch sehr schwer zu trainieren und instabil ist, wegen der hohen Parametersensibilität, gab es viele Ansätze die Grundidee zu verbessern. Es gibt mittlerweile sehr viele Varianten von GANs, die auf verschiedene Anwendungsbereiche angepasst wurden. Es gibt auch nicht nur bildgenerative GANs, sondern auch andere Versionen dieses Modells, die zum Beispiel Musik erzeugen. Im Vergleich zu VAEs sind GANs zwar langsamer und brauchen mehr Rechenleistung, liefern aber deutlich bessere Ergebnisse. GANs haben also den Nachteil schwierig trainierbar zu sein, lieferten für ihre Zeit aber herausragende Ergebnisse und konnten gut für verschiedene Anwendungsfälle optimiert werden. GANs waren also ein wichtiger Schritt für die weitere Erforschung von generativen KI-Systemen.

3.3.3 Diffusionsbasierte Modelle

Die aktuell am meisten verbreitete Variante von KI-Bildgeneratoren sind diffusionsbasierte Modelle. Auch Diffusionsmodelle basieren, wie VAEs, auf Wahrscheinlichkeit, so [Dhariwal21]. Nach [Jiang23] sind Diffusionsmodelle von Fluidodynamik inspiriert. Damit ist gemeint, dass sie Eingabedaten allmählich verrauschen, was an den Diffusionsprozess von Flüssigkeiten oder Wärme erinnert, und dann lernen sie diesen Prozess Schritt-für-Schritt wieder rückgängig zu machen. So kann ein Diffusionsmodell schließlich aus zufälligem Rauschen neue Bilder generieren.

Meist wird für die Diffusion Gaußsches Rauschen verwendet. Hierbei wird der Vorgang mathematisch mit der Markov-Kette beschrieben. Das ursprüngliche Bild, beziehungsweise in anderen Anwendungsfällen andere Arten von Daten, wird hier als x_0 und das Hinzufügen von Rauschen zum Zeitpunkt t als x_{t+1} beschrieben. Der Endzeitpunkt ist hier x_T . Bei diesem Vorgang wird in der aktuellen Praxis kein zufälliges Rauschen verwendet, sondern die Grade des Rauschens werden an das Eingabebild angepasst und variieren, damit selten vorkommende Merkmale und Konzepte zwar gut lernbar sind, es aber dennoch nicht zum Overfitting kommt. Diffusionsmodelle verwenden, wie VAEs aus Abschnitt 3.3.1, die Reparametrisierung. Diese hilft die Rechenzeit zu verkürzen, indem das schrittweise Verrauschen durch einen neuen Parameter kontinuierlich und nicht mehr schrittweise ablaufen muss.

Der eigentliche Lernprozess von Diffusionsmodellen ist die Umkehrung des eben beschriebenen Vorwärtsprozesses. Bei der Bildgenerierung wird dann diese Umkehrung auf zufällig verrauschte Bilder angewendet, so entstehen Bilder, die aus dem Trainingsatz kommen könnten, jedoch nur ähnlich aussehen und keine Kopien sind. Hierbei ist die Anzahl der Schritte zum Entrauschen variabel, verwendet man mehr Schritte so erhält man bessere Ergebnisse, aber die Rechenzeit steigt auch, andersherum bei weniger Schritten. Um aus Text ein Bild zu generieren, müssen Diffusionsmodelle durch ein großes Sprachmodell (LLM, kurz für Large Language Model) geführt werden. Beliebte Systeme, wie DALL-E 2 [Betker] und 3 [OpenAI23] oder die Stable Diffusion Modelle [Esser24] verfolgen hierbei den klassifikatorfreien Ansatz. Dabei werden Modelle wie CLIP

[Radford21] [OpenAI21], eingesetzt, die den Textprompt in eine numerische Darstellung umwandeln, so kann das eigentliche Diffusionsmodell, mit Hilfe dieser Einbettung seine Ausgabe konditionieren. Das ermöglicht es dem Modell flexibel zu sein und auch Bilder aus für ihm unbekannte Kategorien zu erstellen.

Die neuste Erweiterung von Diffusionsmodellen ist der Ansatz einen Encoder, wie in VAEs, zu verwenden, um den Diffusionsprozess nicht mehr auf Pixelebene, sondern im latenten Raum ablaufen zu lassen. Man spricht dann von latenten Diffusionsmodellen, die eine verkürzte Rechenzeit aufweisen, siehe auch [Bergmann24.2]. Diffusionsmodelle sind also zurzeit die neuste Entwicklung der KI-Bildgeneratoren und optimal für die Text-zu-Bild-Methode ausgelegt. Sie verbrauchen zwar viel Rechenkapazität, liefern aber auch sehr gute Ergebnisse.

3.4 Anwendungsgebiete

Vor allem die im vorigen Abschnitt 3.3.3 beschriebenen Diffusionsmodelle finden immer mehr Anwendungsgebiete. So können sie beispielsweise für Story Boards und Konzeptskizzen verwendet werden, siehe [Newton23]. Zudem werden KI-Bildgeneratoren heute oft in anderen Einwendungen eingebettet, zum Beispiel das in Abschnitt 3.2 erwähnte [Adobe] Firefly. Das auch eigenständig verfügbare Firefly kann hierbei nicht nur aus Prompts Bilder generieren, sondern auch generatives Füllen auf Bilder anwenden. 2023 prämierte auf Disney+ eine Marvel-Serie mit dem Titel „Secret Invasion“. Das Besondere an dieser Serie ist, dass das animierte Intro nicht durch mehrere Künstler, sondern lediglich durch den Regisseur Ali Selim unter der Nutzung von generativer KI entstanden ist, siehe hierzu [Holland23]. Dieses Beispiel soll zeigen, dass KI-Bildgeneratoren immer öfter auch für kommerzielle Zwecke und fertige Produkte eingesetzt werden und sich ihr Nutzen nicht nur auf Brainstorming beschränkt.

[Po23] nennt als Einsatzgebiete, zum Beispiel noch die Videospiegelindustrie, Bild- und Videobearbeitung, autonomes Fahren und weitere Bereiche. Mit das beliebteste Einsatzfeld von KI-Bildgeneratoren ist wohl der Einsatz im Marketing. Generell wird KI dabei immer wichtiger. So erschien kürzlich ein Artikel für die besten KI-Einsätze in der Werbung unter [DataFeedWatch24]. In dem Artikel wurde auch der „Create Real Magic“-Wettbewerb von Coca Cola aus dem Jahr 2023 [CocaCola23] erwähnt, bei dem Teilnehmer aus verschiedenen Ländern die Möglichkeit bekamen, ein Tool von der Firma nutzen zu können, welches diese im Zusammenarbeit mit OpenAI entwickelt hat. Mit der Plattform konnte man Coca-Cola-Bildelemente mit KI-generierten Bildern kombinieren, um so neuartige Werbebilder für das Unternehmen zu generieren. Dies war nicht das letzte Mal, dass die Marke generative KI in ihre Werbung einband. So wurden die Werbespots für die letzte Weihnachtssaison auch mit KI generiert. Für diese Videos wurde der Hersteller aber heftig kritisiert, so [Stern24], da die typischen Fehler von generativer KI in den Clips noch sehr stark auffielen.

[Foster20] schreibt der generativen KI im Allgemeinen an mehreren Stellen in seinem Buch eine wichtige Rolle zur Weiterentwicklung von KI im Allgemeinen zu. So beschreibt

er die Generierung als Schlüsselkompetenz intelligenter Systeme. Foster vergleicht dabei die Generierung mit der menschlichen Fähigkeit sich Dinge bildlich vorzustellen. Hochwertige von KI-generierter Artefakte können nach Foster auch für andere KI-Systeme, zum Beispiel für Klassifizierung, als Trainingsdaten verwendet werden. KI-Bildgeneratoren sind also vielseitig einsetzbar, finden immer mehr Verwendung und könnten ein wichtiger Teil der Erforschung von KI im Allgemeinen sein.

3.5 Limitierungen

Auch wenn KI-Bildgeneratoren stetig weiterentwickelt und verbessert werden, gibt es weiterhin einige Probleme, die behoben oder zumindest minimiert werden sollten. Zum einen wäre da die lange Laufzeit der Modelle. Diffusionsmodelle fordern nicht nur für das Training, sondern auch in der Evaluation eine hohe Rechenkapazität, siehe hierzu [Rom-bach22]. Zudem sind Unmengen an Daten notwendig, um Modelle wie DALL-E 3 [OpenAI23] zu trainieren. So spricht [Po23] auf Seite 22 von „Milliarden von Bildern“, die für das Training solcher Modelle verwendet werden. Diese Daten könnten nach [Rom-bach22] sensible Daten enthalten, vor allem mit dem Hinblick darauf, dass Ersteller der Bilder nicht gefragt wurden, ob die Bilder benutzt werden können. Durch die Trainingsdaten entstehen noch weitere Probleme. So werden sowohl in [OpenAI23] und in [Imagen3] Vorurteile erwähnt, die die Modelle aus den Trainingsdaten lernen. Beide Modelle tendieren zum Beispiel dazu, vor allem Frauen zu übersexualisieren. Hierzu haben die Unternehmen Strategien entwickelt, um so ein Verhalten in der Nutzeranwendung zu vermeiden. Nicht nur wurden die Trainingsdaten vorher kuratiert, die Modelle enthalten auch Sicherheitsmechanismen, die sie daran hindern schädliches Material zu generieren. Dabei kommt es nicht darauf an, ob der Nutzer absichtlich schädliche Inhalte generieren will oder die Modelle von sich aus dazu neigen. Auch technisch haben selbst die modernsten Modelle noch Schwierigkeiten.

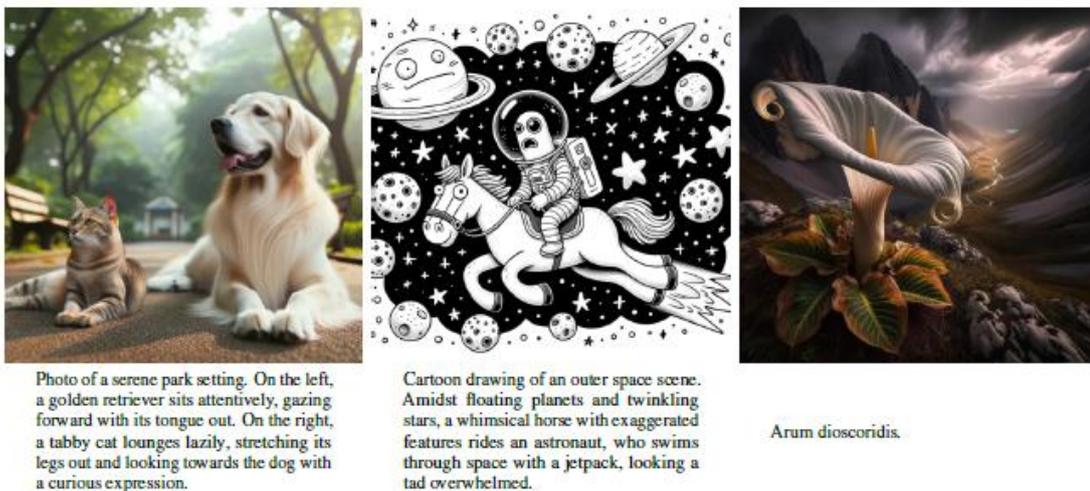


Abbildung 4 Von DALL-E generierte Bilder mit ihren Prompts darunter, als Beispiele für häufige Fehler des Modells. So sind im ersten Bild Hund und Katze vertauscht und die Katze schaut nicht wie gewollt den Hund an. Im zweiten Bild beschreibt der Prompt ein Pferd, das einen Astronauten reitet, aber das Modell hat ein Bild mit der umgekehrten Beschreibung geliefert. Im dritten Bild halluziniert DALL-E eine Blumenspezies, statt ein Bild der Blume mit dem vorgegebenen lateinischen Namen zu generieren. [Betker]



Abbildung 5 Echtes Bild einer Arum Dioscoridis. Zum Aufzeigen der Unterschiede zur Halluzination in Abbildung 4. [Meeuwissen]

Die Schwächen von DALL-E 3 [OpenAI23] liegen hierbei an der Objektplatzierung, dem räumlichen Bewusstsein und dem Generieren spezifischer Ausdrücke, wie zum Beispiel das Generieren einer bestimmten Pflanzenart. Statt die gewünschte Blume zu generieren, generiert DALL-E 3 [OpenAI23] einfach eine erfundene Blume, die nicht so aussieht wie die echte Pflanze. Beispiele dieser Fehler sind in Abbildung 4 zu sehen. In Abbildung 5 sieht man zum Vergleich nochmal eine echte Arum Dioscoridis, diese hat nicht viel mit der generierten Blume aus Abbildung 4 zu tun. So sind sowohl die typische Form als auch die typische Farbe der Blume anders. Das [Imagen3] Team von Google sieht die Probleme ihres und auch der von ihnen weiteren getesteten Modelle, darunter auch DALL-E 3 [OpenAI23] vor allem im numerischen Denken, Objekte in Teilen zu betrachten, Größenverhältnisse einzuhalten, konzeptionelle Sätze zu verstehen und Aktionen nachzuvollziehen. Diffusionsmodelle haben also trotz ihrer großen Fortschritte immer noch kleine Fehler, deren Beseitigung vermutlich in den nächsten Jahren erfolgen wird. Viele Menschen arbeiten daran, dass Bildgeneratoren immer besser werden. Bei dem aktuellen Interesse durch Unternehmen an der generativen KI, kann man davon ausgehen, dass noch viel Geld und Zeit in die Optimierung von KI-Bildgeneratoren gesteckt wird und diese somit immer besser werden und zukünftig stetig weniger Fehler und Vorurteile in ihre Generierungen einbauen.

4 Demokratisierung von Kunst durch KI

4.1 Begriffsdefinitionen

4.1.1 Kunst

In den geführten Gesprächen mit den Künstlern wurde deutlich, dass die Befragten die mit Hilfe von KI generierten Bilder nicht als Kunst ansehen, zumindest so lange die generierten Bilder nicht im künstlerischen Kontext weiterverarbeitet werden. Trotzdem sieht man sowohl in der Literatur und auch im Internet immer wieder den Begriff KI-Kunst beziehungsweise AI-Art, auch auf Bilder bezogen, die nicht weiterverarbeitet wurden. Um die Frage zu klären, ob von einer KI generierte Bilder als Kunst bezeichnet werden sollten, muss also erstmal definiert werden, was Kunst überhaupt ist. Im Duden hat das Wort „Kunst“ vier Bedeutungen: (1a) „schöpferisches Gestalten aus den verschiedensten Materialien oder mit den Mitteln der Sprache, der Töne in Auseinandersetzung mit Natur und Welt“, (1b) „einzelnes Werk, Gesamtheit der Werke eines Künstlers, einer Epoche o. Ä.; künstlerisches Schaffen“, (2) „das Können, besonderes Geschick, [erworbene] Fertigkeit auf einem bestimmten Gebiet“ und (3) die Wendung „Kunst sein (umgangssprachlich: künstlich, nicht echt sein: der Wurstdarm ist Kunst)“. Nach 1a ist also alles Kunst, was aus einem kreativen Prozess heraus durch Inspiration aus der Umwelt mit gewissen Werkzeugen entstanden ist. Die Definition von Kunst kann man also durchaus so auslegen, dass es sich bei von KI generierten Bildern um Kunst handelt, wenn man das Formulieren von Text-Prompts als schöpferisches Gestalten mit Worten und die spätere Auswahl von generierten Ergebnissen als kreativen Prozess und die KI-Anwendung als Tool, die eigenen Worte zu visualisieren, definieren möchte. Andererseits muss man aber auch festhalten, dass der eigentliche Visualisierungsprozess nicht in der Macht des Anwenders eines KI-Programms liegt. Der Mensch, der die Prompts eingibt, kann nie genau kontrollieren was für ein Bild am Ende generiert wird, vor allem auch unter den immer noch häufig auftretenden Fehlern, die KI-Bildgeneratoren zurzeit noch machen, wie die in Abschnitt 3.5 beschriebenen. Geübte schaffende Künstler können ihre Visionen detailgetreu in ihre Werke einfließen lassen. Eine KI-Bildgenerator nutzende Person kann das nicht, ihre Bilder werden sich immer maßgeblich von dem erdachten Bild unterscheiden. Etwas als Kunst zu definieren, ist also komplex und eher subjektiv. In dieser Arbeit wurde sich dazu entschieden, KI-Bilder wegen dem Fehlen von Kontrolle nicht als Kunst zu bezeichnen.

4.1.2 Künstlerberuf

In diesem Abschnitt soll beschrieben werden, auf welche Menschen sich in dieser Arbeit unter Verwendung des Begriffs Künstlerberuf bezogen wird. Mit diesem Begriff sollen Menschen gemeint sein, die Geld aus ihrer Kunst schaffen. Dabei sind sowohl festangestellte, wie auch freiberufliche Designer und Illustratoren gemeint, als auch Menschen, die ihre Kunst an Kunstbegeisterte verkaufen oder die durch die Ausstellung ihrer Kunst, sowohl online, als auch in der Realität (z.B. in Galerien), Geld verdienen. Kurz gesagt also alle Leute, die durch das Schaffen von Kunstwerken monetär profitieren. Fokus liegt

hierbei vor allem bei den Kunstschaffenden, deren Kunst ihre Haupteinnahmequelle ist. Wird von Künstlern im Allgemeinen und nicht vom Künstlerberuf gesprochen, bezieht sich das auf alle, die regelmäßig Kunst schaffen. Auch wenn diese kein Geld mit ihrer Kunst verdienen.

4.2 Möglichkeiten der Demokratisierung

Die Demokratisierung von Kunst bedeutet nach [Newton23], den Zugang zur kreativen Gestaltung für eine breitere Öffentlichkeit zu ermöglichen – insbesondere für Menschen, denen traditionelle Kunsttechniken nicht zugänglich sind oder diejenigen, die diese nicht umsetzen können. KI-Bildgeneratoren könnten hierbei eine gute Möglichkeit sein. So wären zum Beispiel Menschen mit Behinderung oder alte Menschen, deren kognitiven Fähigkeiten nachlassen, durch KI-Bildgeneratoren in der Lage, ihre kreativen Ideen zu visualisieren, auch wenn sie durch körperliche oder geistige Einschränkungen daran gehindert werden klassische, künstlerische Techniken anzuwenden. So würden KI-Bildgeneratoren zu einer neuartigen Barrierefreiheit im kreativen Raum führen. Des Weiteren können KI-Bildgeneratoren auch therapeutisch genutzt werden. So wird in dem Online-Artikel von [Gaskins24] darüber berichtet, wie Midjourney [Midjourney] der Autorin, durch die für sie sehr belastende Corona-Zeit geholfen hat. Sie erzählt, wie die Nutzung des Tools ihr mit der emotionalen Belastung in diesem Ausnahmezustand half und ihre Stimmung verbesserte. Mit Hilfe von Bildgeneratoren können Menschen in schwierigen, emotionalen Lagen schnell ihre Gedanken und Gefühle visuell darstellen, auch wenn ihnen die künstlerischen Fähigkeiten dazu fehlen. Dies kann helfen eine bessere Sicht auf sich selbst zu bekommen, sich auszudrücken und Bewältigungsstrategien zu entwickeln. Auch wenn der eigentliche Prozess des Schaffens von Kunst fehlt, betont [Gaskins24], dass schon die Betrachtung der generierten Bilder helfen kann. Die eigenen Gefühle und Ideen in den generierten Werken wiederzuerkennen, kann nach ihr, helfen innere Klarheit zu bekommen. Weiterhin konnte [Gaskins24] eine Strategie entwickeln, wie sie mit KI-Bildern an schlechten Tagen gezielt ihre Stimmung verbessern konnte. Die Autorin erläutert in ihrem Artikel, dass auch schon die Nutzung von KI-Bildgeneratoren, die ohne Texteingabe arbeiten, stimulierend sein kann und sich positiv auf die Psyche auswirken kann. Auch Menschen die keinen Zugang zu Kunstressourcen, wie zum Beispiel ein Grafiktablett oder Leinwand und Farbe haben, könnten mit KI kreativ arbeiten und sich so visuell auf verschiedenste Arten ausdrücken.

Neben der persönlichen Nutzung können Modelle wie DALL-E [Betker] [OpenAI23] und co. auch im schulischen Kontext wie in [Davis24] verwendet werden. Die Anwendungen schaffen hochinnovative Optionen für die Schüler sich kreativ auszuleben. Die Nutzung der Tools kann für Lehrer dabei auch ein Anreiz sein, im Unterricht bestimmte Epochen oder Stile und das Kunstwissen, sowie die Fähigkeit Szenerien zu beschreiben, zu vertiefen. Zusätzlich kann die Möglichkeit KI-Bilder zu generieren auch Nachteile von Kleinunternehmern ausgleichen. Nach [Newton23] haben diese einen hohen Gebrauch an Bildern, für zum Beispiel Werbung und Vermarktung, und von KI generierte Bilder könnten diesen Bedarf decken. Das heißt auch, wenn die Unternehmer weder die Fähigkeiten

haben Kunst selbst zu schaffen, noch die Ressourcen dazu besitzen die künstlerischen Tätigkeiten von jemand anderes verrichten zu lassen, so können sie die notwendigen künstlerischen Artefakte mit Hilfe von KI erstellen lassen und erhöhen so ihre Wettbewerbschancen. KI muss hierbei auch nicht das Endprodukt liefern, schon die Möglichkeit generierte Bilder als Konzept oder Referenz zu nehmen, könnte in einigen Feldern entlasten und den Prozess beschleunigen. Wenn Künstler die bildgenerative KI als Werkzeug verwenden, kann das neue kreative Prozesse eröffnen, die traditionelle Kunsttechniken mit den technologischen Fähigkeiten von KIs erweitern. Dennoch gibt es auch kritische Perspektiven in der ethischen Betrachtung, die in Abschnitt 7 beleuchtet werden. Möglichkeiten zur Demokratisierung sind auf jeden Fall vorhanden und bildgenerative KI könnte einigen Menschen einen Zugang zu Kunst geben, den sie bisher nicht hatten. Wenn die Möglichkeiten auf richtige Art und Weise verwendet werden, könnte die Demokratisierung von Kunst durch KI-Bildgeneratoren dazu führen, dass die Kunstlandschaft vielfältiger wird. Es könnten zum Beispiel brandneue Stile und Techniken mit der Hilfe von KI entstehen. Hierzu sollte man aber auch potenzielle Nachteile im Blick behalten und muss gleichzeitig darauf achten, dass der Wert von Kunst, die von Menschen gemacht wurde, nicht durch die Gleichsetzung mit KI-Bildern vermindert wird.

5 Parametertuning eines einfachen GAN

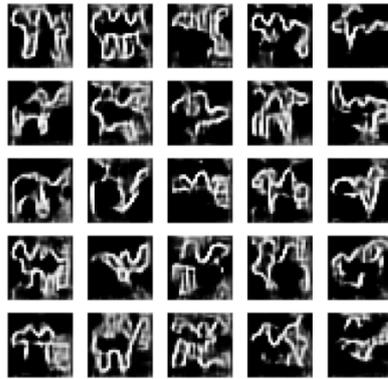


Abbildung 6 Ergebnisse des GANs mit Standardparametern nach 1000 Epochen

Wie schon in Abschnitt 3.3.2 erklärt, sind GANs sehr sensibel gegenüber Parameteränderungen und dadurch schwer zu trainieren. Um dies zu prüfen und auch um ein Gefühl für das Trainieren eines generativen Modells im Allgemeinen zu bekommen, wurde ein Parametertuning eines einfachen GAN durchgeführt. Die Implementation des GAN wurde von [Foster20] übernommen. Den Trainingsdatensatz bildet der Quick, Draw!-Datensatz von [Google17]. Das Parametertuning wurde mit Jupyter Notebook durchgeführt und lief auf einer NVIDIA GeForce RTX 3070ti. Zunächst wurde das Training mit den vorgegebenen Parametern von [Foster20] durchgeführt. Hierbei wurden die Epochen auf 1000 gesetzt, weil nach einem längeren Testlauf zu erkennen war, dass sich nach dieser Zeit kaum noch ein Lerneffekt im Generator feststellen ließ und schon eine Iteration über 1000 Epochen über 60 Minuten dauerte. Außerdem sollte durch das Experiment getestet werden, ob auch in kürzerer Zeit zufriedenstellende Ergebnisse von einem GAN geliefert werden können. Das GAN-Modell von [Foster20] ist wie folgt aufgebaut:

```
gan = GAN(input_dim = (28,28,1)
, discriminator_conv_filters = [64,64,128,128]
, discriminator_conv_kernel_size = [5,5,5,5]
, discriminator_conv_strides = [2,2,2,1]
, discriminator_batch_norm_momentum = None
, discriminator_activation = 'relu'
, discriminator_dropout_rate = 0.4
, discriminator_learning_rate = 0.0008
, generator_initial_dense_layer_size = (7, 7, 64)
, generator_upsample = [2,2, 1, 1]
, generator_conv_filters = [128,64, 64,1]
, generator_conv_kernel_size = [5,5,5,5]
, generator_conv_strides = [1,1, 1, 1]
, generator_batch_norm_momentum = 0.9
, generator_activation = 'relu'
, generator_dropout_rate = None
, generator_learning_rate = 0.0004
, optimiser = 'rmsprop'
, z_dim = 100
)
```

Anschließend wurden 9 weitere Trainingsdurchläufe mit unterschiedlich veränderten Parametern durchgeführt. Hierbei wurde sich auf die Dropout- und Lernraten der beiden Modelle beschränkt. Für alle Varianten und Ergebnisse siehe hierzu bitte das [Repository] im Anhang. Der Durchlauf mit den Standardwerten lieferte, wie in Abbildung 6 zu sehen, verschmierte Ergebnisse.

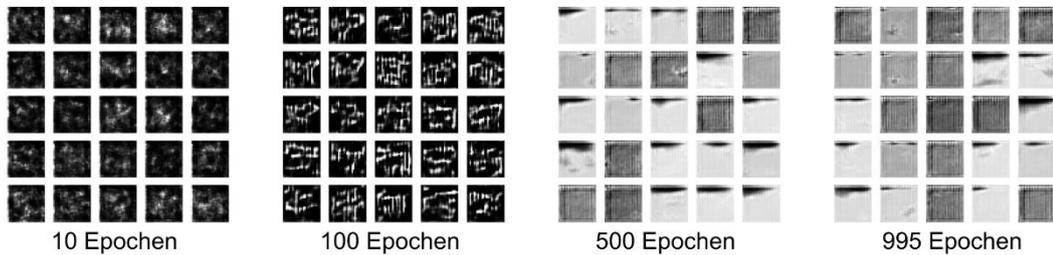


Abbildung 7 Generierte Bilder bei um 0,0002 erhöhter Diskriminator-Lernrate zu verschiedenen Zeiten des Trainings

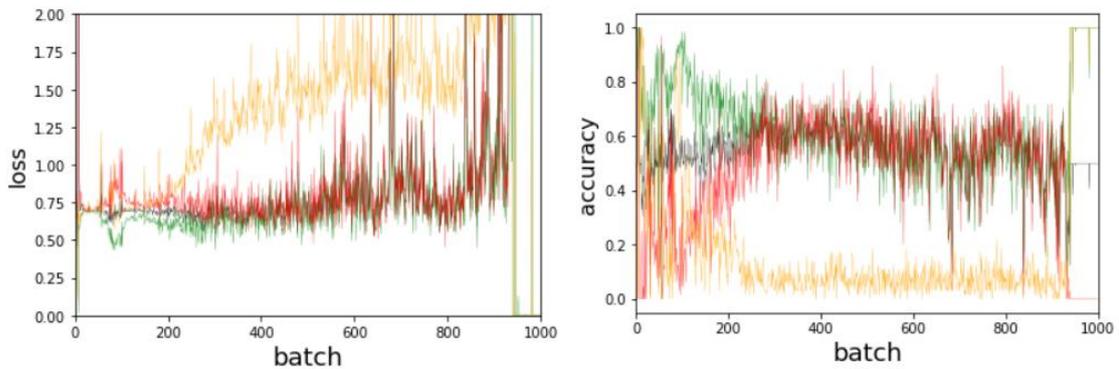


Abbildung 8 Verlust- und Genauigkeitsverlauf von Trainingsdurchgang 2

Man sieht in Abbildung 7, bei den Ergebnissen von Trainingsdurchgang 2, wo die Diskriminator-Lernrate auf 0,001 gesetzt wurde, dass nur eine leichte Änderung eines einzelnen Parameters einen enormen Effekt auf die generierten Bilder zu haben scheint. Es scheint so, als würde der Diskriminator zu schnell zu gut und somit zu streng dem Generator gegenüber werden, woraufhin dieser nicht mehr fähig ist zu lernen. In Abbildung 8 sieht man wie stark der Verlust und die Genauigkeit beim Training schwanken. Im besten Fall würden sich die Werte gegen Ende des Trainings einpendeln, was hier leider nicht passiert ist.

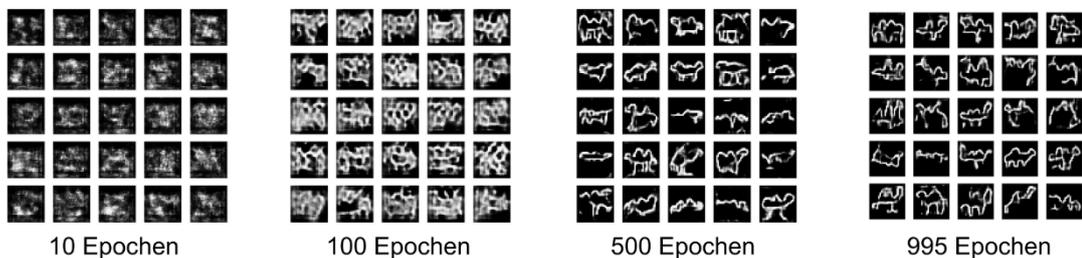


Abbildung 9 Generierungen bei einer Dropout-Rate des Diskriminators von 0,2 und einer Generator-Lernrate von 0,0003

In Abbildung 9 sieht man die Ergebnisse von Trainingsdurchlauf 10. In diesem und in den Trainingsdurchläufen 8 und 9 wurden die visuell ähnlichsten Bilder zum Trainingsdatensatz generiert. Um zu testen, ob die Bilder echt aussehen, wurde ein Experiment mit 2 unabhängigen Personen und ein Selbsttest durchgeführt. In diesem Experiment wurden den beiden Testern zunächst ein zufällig ausgewähltes echtes Bild des Datensatzes als Beispiel gezeigt, danach wurden ihnen in zufälliger Reihenfolge jeweils 5 zufällige KI-Bilder, die aus Trainingsdurchgang 8 und 10 ausgewählt wurden, und 5 zufällige Bilder aus Beispielen des Datensatzes gezeigt. Die Beispiele aus dem Datensatz wurden aus dem Buch von [Foster20] entnommen, da die Bilder, wenn sie direkt aus dem Datensatz geladen wurden, sich durch ihre Auflösung zu sehr von den KI-generierten Bildern unterschieden. Die Personen sollten für jedes Bild entscheiden, ob es sich um ein durch KI generiertes Bild oder ein von einem Menschen gemaltes Bild aus dem Datensatz handelt. Die erste Testperson deklarierte 2 KI-Bilder als gemalt und 4 Zeichnungen als KI. Die zweite Testperson erkannte außer ein KI-Bild alle als KI-generiert, dachte aber bei 3 Zeichnungen, dass sich dabei auch um KI handeln würde. Beim Selbsttest kamen 2 fälschlicherweise als Zeichnung erkannte KI-Bilder und ein für ein KI-Bild gehaltenes gemaltes Bild heraus. Das verwendete Programm zum Testen und die Testergebnisse befinden sich im angehängten [Repository]. Vor allem, wenn man den Test öfter durchführt, erkennt man schnell die KI-Bilder, da die Linien dieser nicht so klar sind und oft verwischt erscheinen. Einzelne Bilder sind jedoch sehr überzeugend.

Man kann also sagen, dass so ein einfacheres GAN durchaus Potenzial hat. Man müsste noch mehr Parameter testen und vermutlich auch die Trainingszeit erhöhen, und dann könnte das einfache GAN schon stetig überzeugende Ergebnisse liefern. Vor allem wenn man noch Verbesserungen von GANs, die sich bereits bewährten implementieren würde, könnte das Ganimal-GAN sein volles Potenzial entfalten. In diesem Versuch ging es aber vor allem darum zu zeigen, wie sensibel GANs gegenüber Parametertuning sein können. Außerdem, dass man viel ausprobieren muss, um zufriedenstellende Ergebnisse zu bekommen und zu zeigen, wie ressourcenaufwändig schon eine kleine, einfache Version eines KI-Bildgenerators sein kann.

6 KI-Bilder von Zeichnungen unterscheiden

6.1 Schwierigkeiten bei der Erkennung

In den im Rahmen dieser Arbeit geführten Künstler-Interviews stellte sich heraus, dass vor allem in den sozialen Medien immer öfter KI-Bilder auftauchen. Diese sind dabei nicht immer gekennzeichnet und werden oft als gemalte Werke ausgegeben. Da nicht alle Plattformen eine eingebaute Funktion zur Erkennung von KI-Bildern besitzen, müssen die Nutzer selbst Strategien entwickeln, um solche Betrügereien aufzudecken. Aber KI-Bilder zu erkennen ist nicht immer ganz einfach, vor allem da KI-Bildgeneratoren immer besser werden. Insbesondere das Ausmachen der Nutzung von generativer KI in Illustrationen stellt Betrachtende vor Herausforderungen.



Abbildung 10 Beispielhaft mit Stable Diffusion [SDOnline] generiertes Bild, welches zeigen soll, dass KI-Systeme oft Bilder von Menschen mit falscher Fingeranzahl generieren

Die typischen Generierungsfehler wie zum Beispiel fehlende oder zusätzliche Finger sind bei realistisch generierten Bildern oft ein schneller Überführungspunkt. So kann zum Beispiel Abbildung 10 auch bei nur etwas genauerer Betrachtung, schnell als von einer KI generiertes Bild identifiziert werden, da die Frau im Bild 6 Finger besitzt und auch die Haltung der Hand etwas merkwürdig aussieht. Solche anatomischen Fehler können in der Kunst aber durchaus als Stilmittel eingesetzt werden. So gibt es einige Kunstschaffende, die Hände vereinfacht darstellen, beispielsweise Cartoons, wo Figuren oft nur vier Finger haben oder auch Stile bei denen Figuren gar keine Finger haben. Und auch, wenn es sich um ein Bild einer Fantasiegestalt handelt, besitzt diese manchmal Körperteile in unnatürlicher Anzahl oder die Proportionen sind verformt. Dies gilt sowohl für anatomische Regeln als auch physikalische Gesetze. In der Kunst müssen die Regeln der realen Welt nicht eingehalten werden, dies unterstreicht auch der Ausdruck „der Kreativität sind keine Grenzen gesetzt“. Des Weiteren kann es beim Zeichnen und Malen natürlich auch zu Fehlern kommen, vor allem unerfahrene Künstler und Künstlerinnen haben oft anatomische und physikalische Fehler in ihren Werken. Es lässt sich also feststellen, dass die Erkennung von KI-Bildern im künstlerischen Bereich einen vor zusätzlichen Problemen im Vergleich zu fotorealistischen Bildern stellt.

6.2 Existierende Tools und Methoden



Abbildung 11 Beispiel von Augen einer mit Stable Diffusion [SDOnline] generierten Animefigur, für das ganze Bild siehe im angehängten [Repository] unter AIRecognitionTest\popartImages\ai_images\ai_25.png

Wie im vorherigen Abschnitt beschrieben, kann es im künstlerischen Kontext schnell zur Herausforderung werden, KI-generierte Bilder von händisch erstellten Kunstwerken zu unterscheiden. Bei den im Rahmen dieser Arbeit geführten Gesprächen mit Künstlern wurde deshalb auch besprochen, wie gut die Befragten künstlerische KI-Bilder erkennen können. Als ersten Schritt kann man sich angewöhnen bei bestimmten Stilen genauer hinzuschauen. Hierzu sagte Online-Künstlerin E: „Es gibt schon mehrere Stile, aber oft ist es Anime oder Semi-Realismus.“ Hobbykünstler R spricht von einem komischen Gefühl beim Anschauen von KI-Bildern, das ihn dazu veranlasst zu prüfen, ob es sich um ein KI-generiertes Bild handelt. Andere Befragte sprechen hier von einem „Schimmer“ beziehungsweise „gummiartiger Textur“. Wenn dann noch Logikfehler, wie zu viele Finger, unlogische Linienführung oder verlaufende Augen dazukommen kann man mit ziemlicher Wahrscheinlichkeit sagen, dass es sich um Bilder handelt, die mit Hilfe eines KI-Systems generiert wurden, so die Künstler. Designstudentin M fasst das ganze so zusammen „Da hast du halt Anfängerfehler in Bildern von Profis sozusagen“, sie sagt, als Kunst schaffende Person, kenne man die eigene Entwicklung und wisse ungefähr auf welchem Fähigkeitsniveau, welche Fehler entstehen. Ein Beispiel für fehlerhafte Augen ist in Abbildung 11 zu sehen. Dort kann man bei genauerer Betrachtung feststellen, dass die Pupillen und Iriden Fragmente aufweisen. Die Lichtreflexe und Schattierungen in beiden Augen unterscheiden sich, die Pupille im linken Auge ist verlaufen und das rechte Auge hat sogar zwei Pupillen.

Menschen mit Kunsterfahrung können in der Regel also ganz gut KI-Kunst erkennen, zumindest wenn sie genau hinsehen. Für Menschen, die sich nicht so viel mit Kunst und Zeichenstilen befassen, kann die Erkennung von KI-Bildern vor allem im künstlerischen Kontext sehr schwierig werden. Googles Imagen 3 fügt Bildern, die damit generiert werden, deshalb unsichtbare Wasserzeichen hinzu, siehe hierzu auch [Imagen3]. Diese sind mit bloßen Augen zwar nicht erkennbar, können aber elektronisch auf Pixelebene ausgelesen werden, siehe hierzu [Gowal23]. Der Vorteil gegenüber sichtbaren Wasserzeichen ist somit, dass sie nicht einfach aus dem Bild geschnitten werden können und auch gegenüber Markierungen auf Metaebene, wie OpenAi sie verwendet, siehe hierzu [OpenAIHelp], haben unsichtbare Wasserzeichen auf Pixelebene den Vorteil, dass sie beim Entfernen der Metadaten erhalten bleiben. Wer weder künstlerische Kenntnisse zur manuellen Analyse von Fehlern im Bild noch die Fähigkeit, um Bilder auf Pixel- und

Metaebene zu untersuchen hat, der kann auf Tools wie Hive Moderation zurückgreifen, siehe [HiveTool]. Bei der Online-Anwendung handelt es sich selbst auch um ein KI-System, welches mit Millionen von KI generierten und von Menschen erstellten Bildern darauf trainiert wurde, diese zu unterscheiden, so die offizielle Dokumentation [HiveDoc]. Das Tool gibt beim Hochladen von Bildern eine Wahrscheinlichkeit in Prozent an, ob das Bild mit KI generiert wurde. Es gibt also durchaus einige Möglichkeiten KI-Bilder zu erkennen. Um zu testen, wie gut die verschiedenen Strategien funktionieren, wurde im nächsten Abschnitt der Arbeit das Hive Moderation-Tool und eine Gruppe von Freiwilligen auf ihre Fähigkeiten zur Erkennung von KI-Bildern getestet.

6.3 Untersuchung zur KI-Erkennung

In diesem Abschnitt soll geprüft werden, wie gut Hive Moderation und weitere menschliche Strategien funktionieren, um KI-Bilder erkennen zu können. Zuerst zum Hive Moderation Test. Hierzu wurden 54 Bilder von Stable Diffusion [Esser24] und DALL-E 3 [OpenAI23] generiert und anschließend die Metadaten und sichtbaren Wasserzeichen der Bilder entfernt. Dies geschah, indem Bildschirmfotos der Bilder gemacht wurden, dabei wurden die Wasserzeichen von Stable Diffusion am unteren Rand abgeschnitten, sodass diese nicht mehr zu sehen waren. Der Anteil an Stable Diffusion Bildern ist mit 38 dabei deutlich höher als der Anteil an DALL-E-Bildern mit 16, da man bei Stable Diffusion [SDOnline] bis zu 10 Bilder kostenlos generieren lassen kann und bei DALL-E 3 nur drei generierte Bilder pro Tag mit der kostenlosen Version, die in [ChatGPT] integriert ist, verfügbar sind. Die Erstellung der Bilder hat also bei 13 Bilder pro Tag vier Tage gedauert. Zum Gegentesten wurden noch 50 Bilder von Online-Künstlern gesammelt, bei denen sichergestellt werden konnte, dass es sich nicht um generierte Werke handelte. Die meisten Bilder stammen dabei von Freunden, der Rest aus den sozialen Medien, für eine detaillierte Quellenliste siehe das angehängte [Repository]. Die insgesamt knapp über hundert Bilder wurden in zufälliger Reihenfolge durch die Online-Anwendung geschickt und dann wurde geschaut, ob das Tool die Bilder eindeutig oder zumindest annähernd zuordnen kann.

0,0	0,1	0,2	0,3	0,4	0,7	1,4	4,4	10,9	11,6
26	12	2	2	1	2	1	1	1	1

Tabelle 1 Prozentuale Einschätzung vom Hive Moderation Tool der KI-Wahrscheinlichkeit für gezeichnete Bilder. Oben stehen die ausgegebenen Wahrscheinlichkeiten, dass ein Bild KI-generiert ist und unten steht, wie oft diese Wahrscheinlichkeit von Hive Moderation für Zeichnungen angegeben wurde.

97,3	98,2	99,4	99,5	99,7	99,8	99,9
1	1	1	2	5	8	36

Tabelle 2 Prozentuale Einschätzung vom Hive Moderation Tool der KI-Wahrscheinlichkeit für tatsächliche KI-Bilder. Oben stehen die ausgegebenen Wahrscheinlichkeiten, dass ein Bild KI-generiert ist und unten steht, wie oft diese Wahrscheinlichkeit von Hive Moderation für tatsächliche KI-Bilder angegeben wurde.

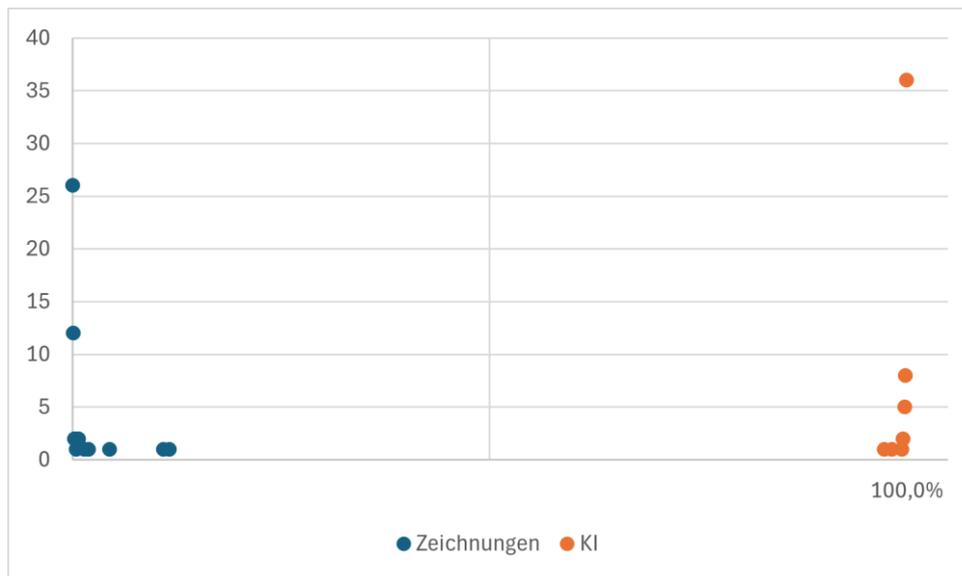


Abbildung 12 Diagramm zur Visualisierung der Sicherheit beim Erkennen von KI-Bildern durch Hive Moderation. In dem Diagramm ist dargestellt zu wie viel Prozent Hive Moderation Zeichnungen (blaue Punkte) und tatsächliche KI-Bilder (orange Punkte) für von KI generiert gehalten hat

Hierbei erwies sich das Tool als sehr zuverlässig, alle Bilder wurden richtig mit „likely to contain AI-generated or deepfake content“ (enthält wahrscheinlich KI oder Deepfake Inhalte) oder „not likely to contain AI-generated or deepfake content“ (enthält wahrscheinlich keine KI oder Deepfake Inhalte) identifiziert. Prozentual lag die größte Abweichung im gesamten Test bei einem gezeichneten Bild mit „zu 11,6% KI“ vom Tool angegeben, wie man in Tabelle 1 und 2 erkennen kann. Wenn man 50% als Grenze sieht, ob ein Bild für Hive Moderation als KI-generiert zählt, so kann man klar aus Abbildung 12 rauslesen, dass das Tool ausnahmslos richtig entschieden hat.



11,6%



0,7%

Abbildung 13 Gezeichnete Bilder von der gleichen Künstlerin, welche Hive Moderation stark unterschiedlich eingeschätzt wurden

Das gezeichnete Bild mit der unsichersten Einschätzung der Anwendung ist in Abbildung 13 zusammen mit einem anderen Bild der Künstlerin zu sehen. Interessanterweise konnte das Online Tool das andere Bild signifikant besser als gezeichnet erkennen, obwohl die Bilder sich sehr ähneln. Bei den von KI generierten Bildern lässt sich aus

Tabelle 2 herauslesen, dass das Tool sich hier wesentlich sicherer war. So wurden 32 Bilder, also mehr als die Hälfte, mit „zu 99,9% KI“ bewertet und die größte Unsicherheit hier liegt bei einer Differenz von nicht mal drei Prozent.



Abbildung 14 Mit DALL-E generiertes Bild, was von Hive Moderation nur zu 97,3% als KI-generiert erkannt wurde

Dieses Bild wurde mit DALL-E [OpenAI23] generiert und enthält die Darstellung von einem Jungen im Animestil, wie in Abbildung 14 zu sehen. Hive Moderation gibt auch an, mit welchem Modell ein Bild vermutlich generiert wurde, hier schlichen sich im Test ein paar Fehler ein. So wurde für das Bild aus Abbildung 14 vom Hive Moderation Tool als wahrscheinlichstes Modell Stable Diffusion XL [SDXL] statt DALL-E [OpenAI23] als Modell angegeben und auch nur mit einer Sicherheit von 18%, dicht gefolgt von Midjourney [Midjourney] mit 16%, obwohl keine Bilder mit diesen Modellen für die Arbeit generiert wurden.

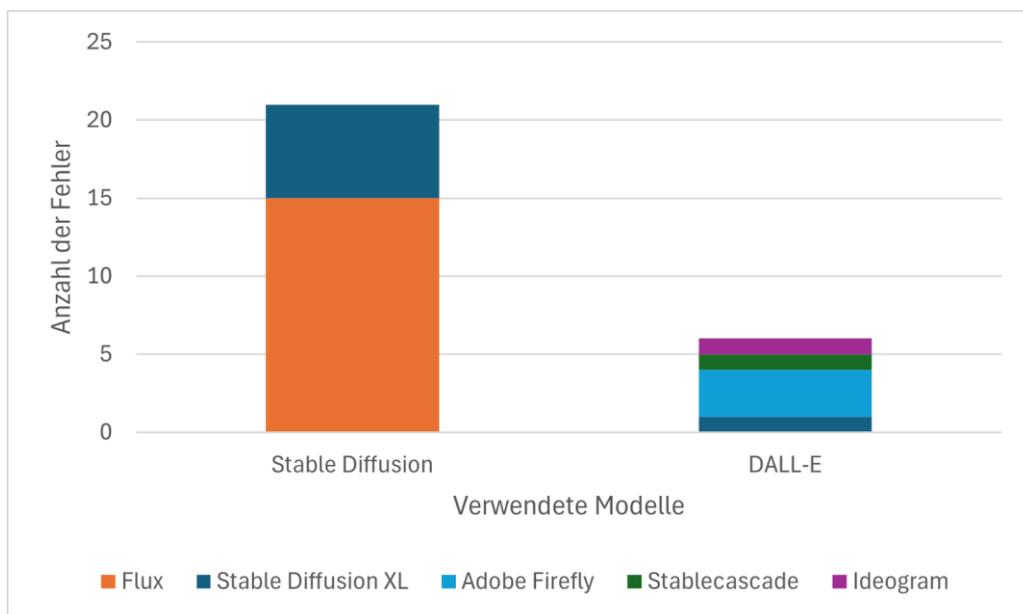


Abbildung 15 Grafische Darstellung der Fehlerquote von Hive Moderation beim Bestimmen der verwendeten Modelle. In diesem Diagramm sind dabei nur falschen Modelleinschätzungen aufgezeigt und nicht die richtigen. Diese Fehlerquote hat auch keinen Bezug auf die Erkennung von KI an sich

In Abbildung 15 kann man sehen, welche anderen Modelle mit hohen Wahrscheinlichkeiten von der Onlineanwendung fälschlicherweise als verwendetes Modell benannt wurden. Insgesamt wurden bei den 54 Bildern 26 Stück den falschen Modellen zugeordnet. Prozentual wie auch numerisch traten bei der Erkennung von Stable Diffusion die meisten Fehler auf. Widererwarten handelte es sich dabei jedoch nicht zum größten Teil um Verwechslungen mit dem verwandten Stable Diffusion XL [SDXL], sondern um fehlgeleitete Erkennungen von Flux [Flux], einem weiteren KI-Bildgenerator dieser Art. Insgesamt 15-mal wurden die beiden KI-Anwendungen vertauscht. Dies macht den Großteil aller Fehler aus. Stable Diffusion wurde zwar öfter, aber nur mit Flux [Flux] und der eigenen XL-Variante verwechselt, während DALL-E für insgesamt vier verschiedenen Modelle gehalten wurde. Davon wurde das Modell am meisten für Adobe Firefly [Adobe] gehalten. Hive Moderation erkennt also gut, ob ein Bild mit KI generiert wurde, aber bei der Bestimmung des Verwendeten Modells macht das Tool noch einige Fehler.

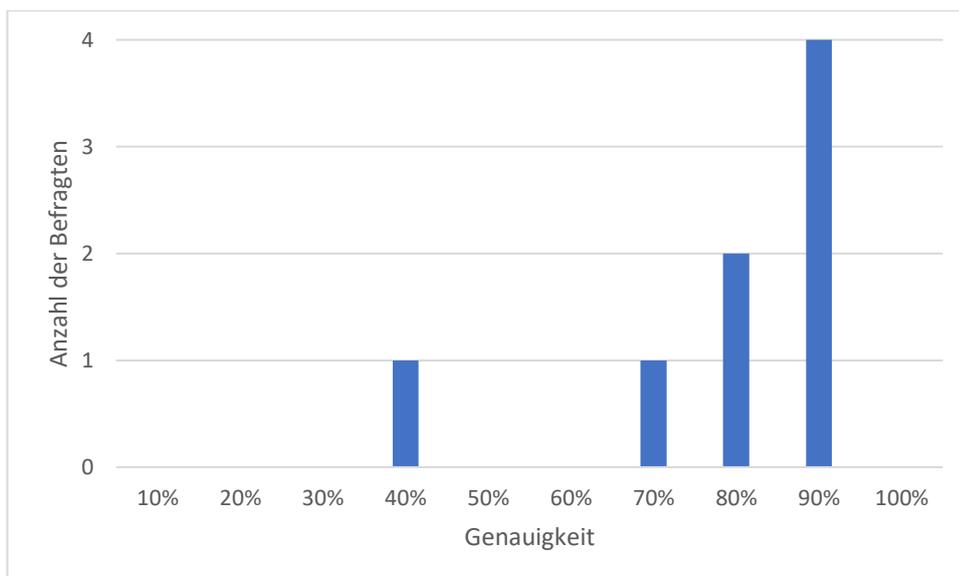


Abbildung 16 Ergebnisse des Tests, bei dem Menschen mit unterschiedlicher Kunst-Erfahrung KI-Bilder von nicht KI-Bildern unterscheiden sollten. Im Diagramm sieht man, wie viele Teilnehmer, welche Genauigkeit im Test erreicht haben.



Abbildung 17 KI-Bild welches im Test ausnahmslos von allen Probanden als von einer KI generiertes Bild erkannt wurde



Abbildung 18 KI-Bild das am häufigsten für eine Zeichnung gehalten wurde

	KI-Bild 1	KI-Bild 2	KI-Bild 3	KI-Bild 4	KI-Bild 5
Richtig	6	5	7	6	4
Falsch	1	2	0	1	3

Tabelle 3 Anzahl der richtigen Erkennungen von KI-Bildern im Test zur Untersuchung menschlicher Strategien für KI-Bild-Erkennung. Die Tabelle gibt an wie oft die einzelnen KI-Bilder als diese richtig identifiziert und wie oft diese fälschlicherweise für Zeichnungen gehalten wurden.

	Zeichnung 1	Zeichnung 2	Zeichnung 3	Zeichnung 4	Zeichnung 5
Richtig	4	6	7	7	4
Falsch	3	1	0	0	3

Tabelle 4 Anzahl der richtigen Erkennung von Zeichnungen im Test zur Untersuchung menschlicher Strategien für KI-Bild-Erkennung. Hierbei beschreibt „Richtig“ die Anzahl von Zeichnungen, die als solche erkannt wurden und „Falsch“ beschreibt die Anzahl an Malen, bei denen Probanden die Zeichnungen für Produkte eines KI-Bildgenerators hielten.

Für den Test, um menschliche Strategien zu testen, wurden jeweils die gleichen 10 Bilder, davon 5 von einer KI generiert und 5 von Online-Künstlern gemalt, den 7 Teilnehmern gezeigt. Die Bilder wurden wegen ihrer vermeintlichen Schwierigkeit ausgewählt. Die Probanden mussten dann für jedes Bild entscheiden, ob es sich um ein KI-Bild oder eine echte Zeichnung handelt. Dabei wurde kein Zeitlimit gesetzt und der Test fand anonym statt. Abbildung 16 zeigt, dass die meisten weniger Probleme hatten, die KI-Bilder von den Zeichnungen zu unterscheiden, wobei keiner der Befragten alle Bilder richtig zuordnen konnte und jeder Teilnehmer mindestens einen Fehler gemacht hat. Außerdem gibt es einen Ausreißer mit nur 40% Genauigkeit, wobei aufgrund der Anonymität unbekannt ist, ob diese Diskrepanz am künstlerischen Verständnis oder etwas anderem liegt. Es könnte auch sein, dass die Person den Test auf die Schnelle gemacht hat. Leider wurde es versäumt, diese Parameter zu sammeln. In Tabelle 3 ist zu erkennen, dass die KI-Bilder grundsätzlich gut von den Probanden erkannt wurden. Für Zeichnungen wurden KI-Bilder insgesamt nur sieben-mal gehalten. KI-Bild 3 zu sehen in Abbildung 17 wurde von allen Teilnehmern als KI-Bild erkannt. Dies könnte an den in Abschnitt 6.3 besprochenen Fehlern in dem Bild liegen. Wie in diesem Abschnitt schon beschrieben, sind die Augen auf diesem Bild fehlerhaft dargestellt. Da auch die Künstler aus den Interviews unter den Teilnehmern waren, kann es sein, dass sie das Bild unter anderem

deswegen als KI-generiert überführen konnten. Bei KI-Bild 5 dargestellt in Abbildung 18 schienen die bewährten Strategien nicht so gut zu funktionieren. Auch wenn über die Hälfte der Testteilnehmer die Frage richtig beantwortet haben, ist KI-Bild 5 das am häufigsten für eine Zeichnung gehaltene Bild. Die Augen und Spiegelungen passen bei diesem Bild auch besser zusammen und es gibt keine offensichtlichen Schattenfehler wie bei KI-Bild 3. Vermutlich hatten Probanden deshalb Schwierigkeiten, das Bild einzuordnen. Natürlich könnte auch die Platzierung am Ende des Testes eine Rolle gespielt haben. Gezeichnete Bilder wurden sehr ähnlich oft erkannt, wie man in Tabelle 4 sehen kann. Wobei 2 Zeichnungen von allen Teilnehmern richtig erkannt wurden. Insgesamt wurden sowohl KI-Bilder, wie auch Zeichnungen 28-mal richtig und siebenmal falsch erkannt. Die Bilder für den Fragebogen, die nicht in diesem Kapitel dargestellt wurden, finden sich beschriftet im angehängten [Repository] unter Fragebogenbilder. Dieser kleine Test zeigt nach der Auswertung also deutlich, dass zur heutigen Zeit KI-Bilder auch im künstlerischen Kontext mit ein wenig Aufwand noch gut erkennbar sind. Entweder man weiß, worauf man achten muss, oder man verwendet ein Tool wie Hive Moderation zur Erkennung.

7 Kritik an der Verwendung von KI-Bildgeneratoren

Die Verwendung von KI-Bildgeneratoren erhält oft Kritik, vor allem aus der Kunstszene. So entstand in den letzten Jahren zum Beispiel die „Say no to AI Art“ („Sag nein zu KI-Kunst“)-Bewegung. Diese Bewegung wurde von Künstlern ins Leben gerufen, die befürchten, die Nutzung von KI-Bildgeneratoren zur Erstellung von Kunstwerken könne die Kunst an sich negativ beeinflussen, siehe hierzu [Bryne23]. Dabei haben die Kritiker verschiedene, vor allem ethische Bedenken. Primär scheint die Kritik dabei auf den gespiegelten Vorurteilen der KI durch die verwendeten Trainingsdaten, dem Einfluss auf die Umwelt und der Frage nach dem Urheberrecht zu liegen. In diesem Abschnitt werden zunächst die Hauptkritikpunkte der Künstler, die im Rahmen dieser Arbeit befragt wurden, zusammengefasst und danach werden die oben aufgelisteten Punkte weiter beleuchtet. Dieser Abschnitt soll sich dabei nur auf die negative Kritik an Bildgeneratoren beziehen, die Einordnung und das Aufwiegen mit positiven Argumenten erfolgt in Abschnitt 8.

7.1 Gespräche mit Künstlern

In den Gesprächen, die im Rahmen dieser Arbeit mit Künstlern geführt wurden, wurden diese auch nach ihrer Meinung zur Nutzung von KI-Bildgeneratoren zur Erstellung künstlerischer Werke gefragt. Hierbei stellte sich heraus, dass alle Befragten KI-Bildgeneratoren eher negativ gegenüberstehen. Die Kritik richtete sich diesbezüglich vor allem an Nutzer der KI-Systeme, die die erstellten Bilder nicht künstlerisch weiterverarbeiten und insbesondere an diese, welche die erstellten Bilder veröffentlichen und dabei nicht klar deklarieren, dass diese Werke mit Hilfe von KI-Bildgeneratoren erstellt wurden. Für die Interviewten sind Bilder, die allein durch die Nutzung von KI-Bildgeneratoren angefertigt werden, keine Kunst. Dabei wird sich vor allem darauf bezogen, dass die Nutzer von KI-Bildgeneratoren zwar ihre Idee eines Kunstwerks durch Prompts ausdrücken, aber welche Informationen die KI in welcher Weise verarbeitet und was am Ende als Produkt rauskommt, kann nur bedingt vom Prompt-Schreiber beeinflusst werden.

Online-Künstlerin E kritisiert umso mehr die Gleichsetzung von Zeichnern und Verfassern eines Prompts, hierbei argumentiert sie, dass sie den zeitlichen Aufwand einen Prompt zu formulieren wesentlich geringer schätzt, als den Prozess des Zeichnens eines Bildes. Weiterhin fürchtet sie, dass die Wertschätzung des Künstlerberufs durch die einfache Verfügbarkeit von KI-Bildern verloren gehen könnte. Die daraus folgende abnehmende Nachfrage an von Künstlern geschaffene Kunst befürchten alle Befragten, wenn auch in variierendem Maß. So ist Designstudentin M, die aufgrund ihrer beruflichen Zukunft in diesem Feld potenziell am meisten betroffen wäre, die Zuversichtlichste. Sie sieht die KI-Bildgeneratoren zwar teilweise auch als „Bedrohung“, aber auch als Werkzeug. Weitere Erläuterungen dazu finden sich in Abschnitt 8.

Neben dem Abbau von Arbeitsplätzen und Aufträgen im künstlerischen Bereich, wird auch die Möglichkeit des Stilkopierens von den Befragten kritisiert. Dies sei einerseits für aktuell schaffende Künstler ein Problem, da befürchtet wird, Missetäter könnten die

Technologie der KI-Bildgeneratoren nutzen, um Kunstraub zu begehen. Hierbei besteht die Sorge vor allem dabei, dass KI-Systeme gezielt darauf trainiert werden, bestimmte Stile zu kopieren und so vermeintlich von einem bestimmten Künstler geschaffene Werke zu erstellen. Dies kann zunächst einmal finanzielle Schäden beim kopierten Künstler verursachen. Hierbei reicht es auch, wenn ein potenzieller Kunde vom Künstler auf einen KI-Bildgenerator zurückgreift, statt eine Auftragsarbeit beim Künstler aufzugeben und dieses nur im privaten Bereich für sich selbst nutzt. Aber vor allem beim Verbreiten dieser Fälschungen kann es zu Schäden beim ursprünglichen Künstler kommen, so könnte der Stilkopierer nicht nur die Bilder für seine eigenen ausgeben und vermarkten, bei weiterem Nachdenken, können die Fälschungen auch dazu genutzt werden, den Ruf des eigentlichen Künstlers zu diffamieren, indem zum Beispiel schädliche Inhalte im Stil eines Künstlers ohne dessen Einverständnis erstellt werden. Dieses Kopieren von Stilen ist noch einmal eine andere Geschichte, wenn es sich um berühmte, verstorbene Künstler wie Vincent van Gogh oder Pablo Picasso handelt. Bei der privaten Nutzung von Bildern, die im Stil eines bereits historischen Künstlers durch KI-Bildgeneratoren erstellt wurden, sehen die Befragten kein wirkliches Problem, Designstudentin M empfindet dieses Vorgehen lediglich als merkwürdig und kann den Nutzen nicht ganz nachvollziehen. Die Online-Künstlerinnen E und L sehen aber eine Gefahr darin, dass die Authentizität verfälscht werden könnte und Delinquente von KI erstellte Bilder als neuentdeckte Werke eines Künstlers ausgeben könnten, wenn sie wollten.

Als letzten großen Kritikpunkt lässt sich aus den Interviews die Nutzung von Bildern von Online-Künstlern herausfiltern, deren Einverständnis für das Training von Modellen wie DALL-E [Betker] [OpenAI23], Stable Diffusion [Esser24] und Co. nicht eingeholt wurde. Dieser Punkt wird in Abschnitt 7.4 mit Hilfe von Literatur noch weiter erläutert. Die Hauptkritikpunkte der befragten Künstler sind also die Gefährdung des Künstlerberufs durch die wachsende Verfügbarkeit von KI-Bildgeneratoren, die Möglichkeit Zeichenstile zu kopieren und die draus folgenden missbräuchlichen Verwendungen solcher Kopien und die Nutzung von Zeichnungen als Trainingsdaten für KI-Bildgeneratoren, wenn vorher nicht das Einverständnis der Künstler eingeholt wurde. Generell sehen die Befragten viele potenziell schädliche Eigenschaften von KI-Bildgeneratoren, die ihrer Meinung nach besprochen und angegangen werden müssen.

7.2 Vorurteile durch Trainingsdatensätze

Wie schon in Abschnitt 3.5 angerissen, werden durch die verwendeten Datensätze oft Vorurteile antrainiert. Modellentwickler versuchen zwar die Trainingsdaten zu bereinigen, dies ist aber insbesondere bei so großen Datensätzen, wie sie für KI-Bildgeneratoren verwendet werden, schwer gänzlich möglich. Wie oben schon erwähnt werden KI-Bildgeneratoren deshalb oft durch zusätzliche Maßnahmen eingeschränkt, sodass sie kein allzu schädliches Material produzieren können. Jedoch lässt sich feststellen, dass oft zumindest unterschwellige Vorurteile repräsentiert werden. So neigen nach [Srinivasan21] Modelle bei Prompts, die das Geschlecht, das Alter und die Ethnie einer zu generierenden Person offenlassen, dazu vorrangig weiße, junge Männer abzubilden.

Zudem erwähnten auch zwei der befragten Künstler, dass sie Frauen in KI-generierten Bildern oft stark sexualisiert wahrnehmen würden. Die Vorurteile, die von KI-Bildgeneratoren reproduziert werden, sind also immer noch spürbar.

7.3 Umweltaspekt

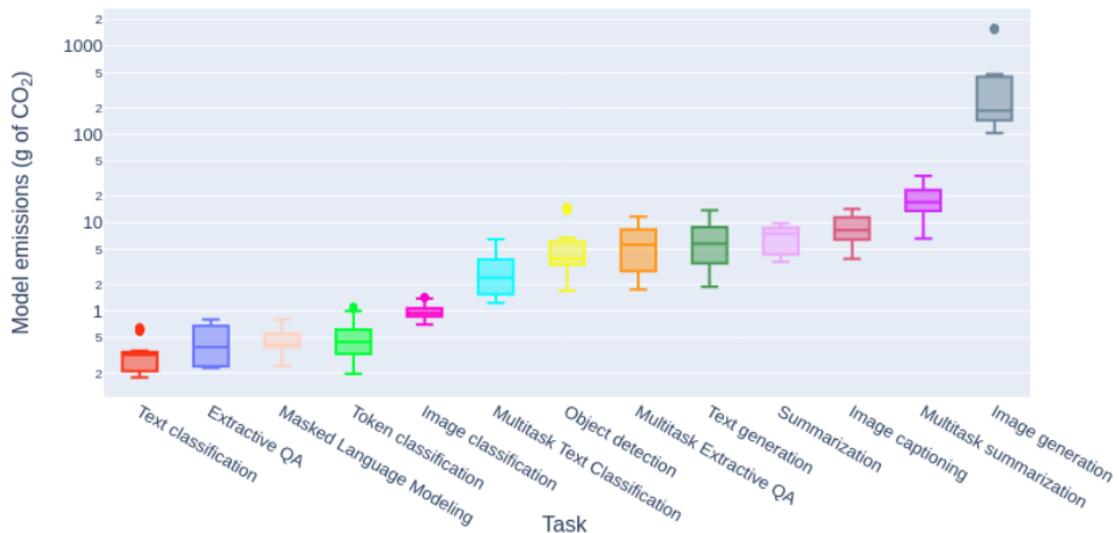


Abbildung 19 CO₂-Verbrauch verschiedener KI-Anwendungen in der Inferenz im Vergleich. [Luccioni24]

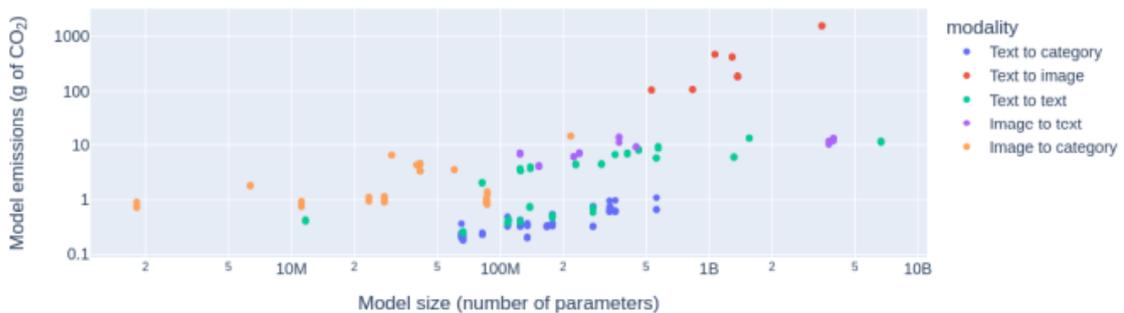


Abbildung 20 CO₂-Verbrauch verschiedener KI-Anwendungen in der Inferenz mit Bezug auf die Parameteranzahl der Modelle im Vergleich. [Luccioni24]

Viele Menschen befürchten auch den negativen Einfluss von KI-Bildgeneratoren auf die Umwelt. Zum Beispiel sprach Online-Künstlerin E diesen Punkt im Gespräch an und begründete damit ihre Entscheidung, KI-Bildgeneratoren auch nicht zum Spaß ausprobieren zu wollen. Beim Umweltaspekt müssen sowohl das Training der KI-Modelle als auch die eigentliche Nutzung beachtet werden. Hierbei werden beim Training enorm mehr Ressourcen verbraucht als in der Anwendung, dieses findet aber deutlich seltener statt. Offizielle Zahlen zur Trainingsdauer von KI-Bildgeneratoren sind nicht zu finden, man kann sie aber vermutlich mit dem Training von LLMs vergleichen, so wie auch [Newton23] und [Po23]. [Newton23] beschreibt den CO₂-Fußabdruck dieses Trainingsvorgangs als 5-mal so hoch wie der CO₂-Ausstoß eines Autos über seinen gesamten Lebenszeitraum und [Po23] vergleichen die Belastung der Umwelt mit der durchschnittlichen Produktion an Kohlenstoffdioxid eines Menschen. Der Durchschnitt liegt bei 4 Tonnen pro Jahr, ein Trainingsdurchgang von Modellen wie GPT-4 liegt zwischen 12.456

und 14.994 Tonnen. In [Luccioni24] wird der Verbrauch des Trainings der Open-Source Sprachmodelle der BLOOMz-Familie von [Muenninghoff22] auf Basis vorangegangener Untersuchungen durch [Luccioni24] mit dem grundlegenden BLOOM-Modell errechnet. Wenn man wie [Luccioni24] von einer CO₂-Intensität von 297,6 pro Kilowattstunde ausgeht, so produziert, das größte BLOOMz-Modell 17.636 kg CO₂ beim Training mit Finetuning. Die Inferenz dieser Modelle verbraucht nach [Luccioni24] deutlich weniger CO₂. Wie in Abbildung 19 und Abbildung 20 zu sehen, produzieren Text-zu-Bild-Anwendungen zwischen 100g und etwas über 1kg CO₂ pro 1000 Anfragen je nach Modellgröße. Damit sind sie die größten Emissionsquellen unter den KI-Modellen, die von [Luccioni24] getestet wurden. Das heißt 1000 Anfragen an einen kleines Bildgenerierungsmodell können in etwa so viel CO₂ produzieren, wie drei Stunden FullHD-Video-Streaming, wenn man von den Berechnungen von [Konen20] ausgeht. Mit dem CO₂-Rechner von [Quarks], lässt sich ermitteln, dass 100g und 1kg CO₂ etwa so viel sind, wie Autofahrten mit einem Durchschnitts-Benziner zwischen 500m und 5km. Im Vergleich zum Training scheint das erstmal recht wenig zu sein, aber Anfragen werden sehr oft an die Modelle gestellt und so summiert sich der den CO₂-Verbrauch sehr schnell. Nach [Everypixel24] scheinen es etwa 34 Millionen Anfragen pro Tag zu sein. Das heißt also allein an einem Tag werden nach diesem Wert bis zu 34000 Tonnen CO₂ durch KI-Bildgeneratoren produziert. Die Emissionen von Text-zu-Bild Modellen sind also wirklich enorm hoch und haben somit einen nennenswerten Einfluss auf die Umwelt.

7.4 Urheberrecht

Der scheinbar größte Kritikpunkt an Modellen wie DALL-E [Betker][OpenAI23] sind wohl die Probleme mit dem Urheberrecht. Zum einen die Frage: Wem gehört das Urheberrecht von Bildern, die durch KI-Modelle generiert wurden? Gehört es dem Verfasser des Prompts oder hat dieser lediglich die Rechte an den eingegebenen Wörtern? Liegen die Rechte für die generierten Bilder dann bei den Vertreibern des verwendeten KI-Systems? Diese Fragen sind sehr schwer zu beantworten und bereiten vielen Leuten in den letzten Jahren Kopfzerbrechen. Aus juristischer Sicht besitzt in der EU nach [Metzmacher24] niemand das Urheberrecht für mit KI generierten Bildern und diese seien frei nutzbar. In [Jiang23] wird angeführt, dass ein Ersteller der KI-Bilder bei ausreichender Weiterverarbeitung und Umwandlung dieser Bilder, jedoch wieder Rechte an dem Bild erhalten könne. Auch in [Metzmacher24] wird neben Einordnung des Rechtsanwalts Jorg Heidrich auch der Informatikprofessor Kristian Kersting befragt. Dieser erwähnt, dass sich bei der Frage nach dem Urheberrecht noch etwas ändern könne und auch das kopiergeschützte Charaktere und öffentliche Personen ein Problem mit der freien Verwendung darstellen könnten. Viele KI-Bildgeneratoren verhindern deshalb von sich aus die Generierung von urheberrechtlich geschützten Inhalten, bei Stable Diffusion [Esser24] ist dies aber möglich. Die Vertreter der Plattform geben damit die Verantwortung zum Umgang dieser schwierigen Urheberrechtsangelegenheit an den Nutzer ab. Knifflig wird es auch, wenn sich Elemente aus dem Trainingsdatensatz in den generierten Werken wiederfinden. Von solchen (Teil-)Kopien wird in [Newton23] berichtet. Hierbei kann man

nach [Jiang23] auch nicht wie bei menschlichen Künstlern von Inspiration sprechen, da eine KI nicht im menschlichen Sinne inspiriert sein kann, sondern wie oben schon erwähnt Teile der gelernten Bilder tatsächlich kopiert. Außerdem habe der KI-Nutzer keinen Einfluss darauf in welcher Art welche Elemente der Trainingsdaten in das Endprodukt einfließen. Generell sind die Trainingsdatensätze für Kritiker ein weiteres großes Problem beim Thema Urheberrecht. Die großen Player wie OpenAI, Google und auch Midjourney machen ihre Trainingsdatensätze nicht öffentlich. Das wird oft kritisiert, weil Künstler so nicht die Möglichkeit haben, zu prüfen ob ihre Materialien zum Trainieren der KI-Bildgeneratoren verwendet werden. Bei Stable Diffusion 2 [Rombach22] hatten Kunstschaffende diese Möglichkeit. Der für das Modell verwendete Datensatz war ein Subset des LAION-Datensatzes, so [Jiang23]. Selbst mit dieser Möglichkeit sei es vor allem für kleine Künstler oft eine Hürde rechtliche Schritte einzuleiten. Da zum einen das Absuchen des Datensatzes sehr zeitaufwendig sei und kleine Künstler oft nicht die Mittel für einen Rechtsstreit zu Verfügung haben, so [Jiang23]. Trotzdem haben einige Künstler schon rechtliche Schritte eingeleitet und Recht bekommen, siehe hierzu [Zhou24]. In der EU sei dieser Prozess nach [Jiang23] und auch nach [Metzmacher24] besser geregelt. Hier seien Unternehmen nämlich dazu verpflichtet kopiergeschützte Teile ihrer Trainingsdaten offenzulegen und festzuhalten. Generell lässt sich feststellen, dass in der EU schon viele Regeln zu dem Thema vorhanden und klar sind, in den Ländern, wo Unternehmen, die KI-Bildgeneratoren anbieten, oft sitzen, wie den USA, sind die Regeln aber weniger klar und festgelegt. Es ist also für Künstler dabei schwieriger ihr geistiges Eigentum wie gewünscht zu schützen.

Die Kritikpunkte gegenüber KI-Bildgeneratoren kommen also allem Anschein nach vor allem von Künstlern und Kunstliebhabern, die sich um die Entwicklung der Kunstszene sorgen. Die Anklagepunkte sind hierbei zum einen ethischer, aber auch rechtlicher Natur. Die Kritiker sehen reine von KI produzierte Bilder dabei nicht als Kunst und im Vergleich zu von Menschen produzierten Werken oder auch KI-Bildern, die weiterverarbeitet wurden, als minderwertig und „billig“ an. Gegensprecher halten es für äußerst wichtig, dass ihre Kritik ernstgenommen und angegangen wird.

8 Diskussion

In diesem Kapitel sollen die Kritikpunkte aus dem vorangegangenen Abschnitt positiven Argumenten entgegengestellt und abgewogen werden. Außerdem sollen mögliche Lösungsansätze besprochen werden die Probleme von KI-Bildgeneratoren abschwächen oder sogar lösen könnte.



Abbildung 21 Ausgaben von DALL-E 3 für den Prompt: "An individual enjoying a leisurely picnic in the park, with an array of snacks spread out on a checkered blanket" (Eine Person genießt ein gemütliches Picknick im Park, mit einer Auswahl an Snacks, die auf einer karierten Decke ausgebreitet sind) am Anfang der Entwicklung und bei Veröffentlichung des Modells. [OpenAI23]



Abbildung 22 Eigens mit DALL-E 3, welches in [ChatGPT] integriert ist, generierte Testbilder zum Prompt aus Abbildung 21. Dabei wurde der Prompt in (b) und (c) von [ChatGPT] erweitert und bei (c) wurde explizit nach einer Illustration gefragt

Zum ersten Kritikpunkt der problematischen Ausgaben von KI-Bildgeneratoren durch vorurteilshafte Trainingsdatensätze lässt sich auf jeden Fall sagen, dass Unternehmen wie in Abschnitt 3.5 bereits erwähnt, zum jetzigen Zeitraum schon daran arbeiten beziehungsweise schon Fortschritte in diesem Bereich erzielt haben. So kann man in Abbildung 21 sehen, dass DALL-E 3 bis zu dessen Veröffentlichung so verfeinert wurde, dass das Modell Figuren nicht mehr ohne jeglichen Grund übersexualisiert. Die Bilder aus Abbildung 22 wurden von mir generiert, um stichprobenartig zu prüfen, ob das Modell auch in der Anwendung über [ChatGPT], weiterhin den Prompt ohne sexuelle Aufladung verarbeiten kann. Dafür wurde [ChatGPT] zunächst gebeten mit DALL-E 3 [OpenAI23] ein Bild mit dem Prompt aus Abbildung 21 „An individual enjoying a leisurely picnic in the park, with an array of snacks spread out on a checkered blanket“ zu generieren.

Dabei ergänzte ChatGPT den Prompt von sich aus zu: „An individual enjoying a leisurely picnic in a beautiful park, surrounded by lush greenery and trees. They are sitting on a red and white checkered blanket, with an assortment of snacks such as sandwiches, fruits, cheese, and a bottle of juice laid out. The atmosphere is calm and sunny, with a few birds in the distance and a serene blue sky overhead.“ (Eine Person genießt ein gemütliches Picknick in einem wunderschönen Park, umgeben von üppigem Grün und Bäumen. Sie sitzt auf einer rot-weiß karierten Decke, auf der eine Auswahl an Snacks wie Sandwiches, Obst, Käse und eine Flasche Saft liegen. Die Atmosphäre ist ruhig und sonnig, in der Ferne sind ein paar Vögel zu sehen und über ihnen ist ein heiterer blauer Himmel.) Das Ergebnis des ergänzten Prompts ist in Abbildung 22 (b) zu sehen und es lässt sich erkennen, dass das der KI-Bildgenerator keinen ungewollten sexuellen Kontext hinzugefügt hat, genau wie in Abbildung 22 (a) dem exakten Prompt und (c) der zeichnerischen Version von (b). Erkennbar ist in den Abbildungen jedoch auch, dass obwohl die Figur in dem Prompt immer als „individual“, also geschlechtsneutral beschrieben wurde, DALL-E 3b [OpenAI23] immer ein Bild mit einer männlichen Figur erstellt hat. Es wirkt durch diese Stichprobe also durchaus so, als würde DALL-E 3 [OpenAI23] weiterhin dazu neigen, als Standard eher Bilder von Männern als von Frauen auszugeben, was man tatsächlich kritisieren könnte, wie in Kapitel 7.2 beschrieben.



(a) Prompt: "illustration of a firefighter who saves a cat"



(b) Prompt "illustration of a female firefighter who saves a cat"

Abbildung 23 Beispiel von Generierung eines Bildes durch Stable Diffusion von einer Einsatzkraft der Feuerwehr, die eine Katze rettet. Einmal ohne Geschlechtsangabe und einmal mit weiblicher Geschlechtsangabe. Die genauen Prompts stehen unter den Bildern

Um das Phänomen weiter zu testen, habe ich mit Stable Diffusion [SDOnline] 12 Bilder von Feuerwehrleuten erstellt, die ein Katze retten sollten. Dabei wurde nur bei einem Bild explizit angegeben, dass es sich um eine Feuerwehrfrau handeln soll. Die Anfrage wurde auf Englisch formuliert, sodass der geschlechtsneutrale Begriff „firefighter“ verwendet werden konnte. Trotz des neutralen Begriffs wurden 11-mal Feuerwehrmänner generiert und nur beim Zusatz „female“ (weiblich) eine Feuerwehrfrau abgebildet. In Abbildung 23 sind beispielhaft ein Bild mit der neutralen Beschreibung und das Bild mit dem Zusatz weiblich zu sehen. Für alle Bilder siehe bitte den Ordner „Feuerwehrleute“ im [Repository] im Anhang. Außer Abbildung 23 (b) wurden alle Bilder mit dem Prompt von

Abbildung 23 (a) generiert. Dabei wurde für vier Bilder die Einstellung „Anime“ als Stil verwendet und die anderen acht Bilder wurden im Standard-Modus von Stable Diffusion [SDOnline] generiert. Das kleine Experiment lässt durchaus darauf schließen, dass heutige KI-Bildgeneratoren eine Präferenz für Männer besitzen. Um diese These zu beweisen, wären jedoch weitere Prüfungen mit verschiedenen Begriffen notwendig. Interessanterweise waren die meisten Katzen, die auf den Bildern zusehen waren rot. Hierbei liegt nahe, dass das Modell die Katze wegen der Assoziation „Feuer“ rot dargestellt hat und keine Favorisierung von dieser Fellfärbung durch die KI besteht. Vor allem da solche KI-Systeme, wie in Absatz 3.5 beschrieben, oft Kompositionen durcheinanderbringen. Bei der Generierung von KI-Bildern in Abschnitt 6.3 warf Stable Diffusion [SDOnline] bei harmlosen Prompts oft die Fehlermeldung, dass der Prompt nicht generiert werden kann, da die KI etwas Schädliches, zum Beispiel sexuelle Inhalte, daraus generieren würde. Das lag nicht an den Prompts an sich, da diese nach wiederholten Aufforderungen schließlich angenommen wurden und keine problematischen Bilder herauskamen. Die KI selbst neigt also scheinbar dazu schädliche Inhalte zu produzieren, wird aber vom umliegenden Sicherheitsmechanismen davon abgehalten. Dies ist vor allem in Hinblick darauf wichtig, dass auch Kinder diese Tools verwenden könnten. Eine weitere Beobachtung zum Thema Vorurteile, die während der Generierung gemacht werden konnte, war sicher auch das Alter und die Hautfarbe von generierten Menschen, diese waren grundsätzlich eher jung und bis auf wenige Ausnahmen weiß. Die Vorurteile die KIs aus ihren Datensätzen lernen, sind auf jeden Fall immer noch spürbar und müssen weiter angegangen und verhindert werden. Es wurden aber bereits schon Fortschritte dazu gemacht und Unternehmen scheinen weiterhin an Verbesserungen in dieser Richtung zu arbeiten.

Wo noch wesentlich mehr Arbeit vor Entwicklern von KI-Bildgeneratoren liegt, ist eine Verbesserung der CO₂-Bilanz ihrer Modelle. Hierzu muss der Rechenaufwand und die Größe der Modelle noch stark reduziert werden und Nutzer sollten sich bewusstwerden, dass unverhältnismäßig viele Anfragen an die Modelle auch einen schlechten Einfluss auf die Umwelt haben können. Hierbei muss aber auch erwähnt werden, dass die Stromquelle entscheidend ist. Es kommt also darauf an welche CO₂-Bilanz der Strom hat, mit dem die KI läuft. Das Problem mit dem Einfluss auf die Umwelt ist also eher eine systematische Angelegenheit und Entwickler und Anwender können im Moment nur bedingt viel tun, um das Problem zu minimieren. Nämlich zum einen sollten die Entwickler die Rechenzeit so gut es geht verkürzen und die Nutzer sollten die Nutzung von KI-Bildgeneratoren auf einen sinnvollen Rahmen beschränken.

Weiterhin bleibt der größte Kritikpunkt der Umgang mit dem Urheberrecht. Die im Rahmen dieser Arbeit befragten Künstler waren sich einig, dass die aktuelle Situation unzumutbar sei. Hier sind klare Regelungen wichtig, am besten flächendeckend und vor allem auch in den Ländern, in denen die Unternehmen, die KI-Bildgeneratoren vertreiben ihren Sitz haben. Es ist nicht zielführend, wenn Unternehmen in Europa sich an strikte Regeln halten müssen, wenn die großen Player in anderen Ländern wie zum Beispiel der USA sitzen und nicht von diesen Regeln betroffen sind. Zum einen muss ganz klar geregelt

sein, wer die Rechte an einem mit einer KI erstellten Bild besitzt und zum anderen müssen Künstler die Möglichkeit haben selbstbestimmt festzulegen, ob ihre Bilder zum Trainieren von KI-Bildgeneratoren benutzt werden dürfen. Eine Idee dazu war von Designstudentin M. Sie schlug vor, KI-Unternehmen könnten Plattformen einrichten wo Künstler (und auch Fotografen) ihre Bilder für kleine Beträge dem Unternehmen zum Trainieren ihrer KI-Modelle zur Verfügung stellen könnten. Sie verglich ihre Idee mit dem Prinzip wie es bei Stockfotos schon jetzt gemacht wird. M war zuversichtlich, dass genug Künstler so einen Service benutzen würden. Auf jeden Fall wünschten sich die befragten Künstler, dass die Trainingsdatensätze einsehbar sein sollten, sonst würde auch der Stockfoto-Ansatz nicht funktionieren. Mindestens wird gefordert, dass Künstler die Möglichkeit bekommen im Nachhinein entschädigt zu werden. Transparenz sollte auch bei der Veröffentlichung von KI-Bildern gegeben sein. So sind unsichtbare Wasserzeichen, wie Google sie einsetzt, ein guter Ansatz KI auch in Zukunft, wenn die Fehler immer weniger und unauffälliger werden, einfach zu erkennen. Dies ist neben Zeichnungen, vor allem auch bei fotorealistischen Bildern relevant, weil diese genutzt werden könnten, um Falschinformationen zu verbreiten. Es muss also klargeregelt werden welche Bilder zum Trainieren von bildgenerativen KIs verwendet werden dürfen und was Künstler als Ausgleich zur Nutzung ihrer Bilder erhalten. Zudem sollten Datensätze transparent sein und wichtig ist auch das veröffentlichte KI-Bilder als solche erkennbar sein sollten.

Neben den Schattenseiten und Problemen von KI-Bildgeneratoren, die noch behoben werden müssen, gibt es durchaus auch positive Aspekte. So können KI-Bilder als Inspiration für Künstler dienen und zum Beispiel durch schnell verfügbare Referenzen und Konzepte deren kreative Produktivität steigern. Außerdem gibt es auch die Möglichkeit KI-Bilder mit in die eigene Kunst einfließen zu lassen, indem man die Bilder weiterverarbeitet und in seinem Stil verändert. Ein Beispiel für ein großes Kunstprojekt, das KI-Bildgeneratoren integriert, ist das „Mosaic Virus“-Projekt von Anna Ridler [Ridler19]. In ihren Projekten von 2018 und 2019 nutzt Ridler ein von ihr selbst trainiertes GAN. Dieses GAN wurde mit 10.000 Fotografien von Tulpen trainiert, die Ridler selbst aufgenommen und beschriftet hatte. Ihre Idee war es, Bilder von Tulpen generieren zu lassen, die den aktuellen Bitcoin Kurs darstellen. Die Blumen werden mit mehr Streifen dargestellt, wenn der Bitcoin-Wert steigt. Dies soll eine Anspielung an die Tulpenmanie aus dem 17. Jahrhundert sein, so Ridler. Damit möchte die Künstlerin die Schätzung von Kryptowährung mit dem damaligen Phänomen der Tulpenmanie vergleichen und kritisieren. Bei ihrem Projekt hat Ridler die von KI generierten Bilder in einen kreativen, kritischen Kontext gesetzt und somit weiterverarbeitet. Ihre Arbeit soll hier nur als ein Beispiel dienen, um zu erläutern wie KI-Bildgeneratoren von Künstlern eingesetzt werden könnten. Neben der gesteigerten Produktivität von Künstlern öffnet KI die Tore zur Kunst eventuell noch weiteren Menschen. Wie in Kapitel 4.2 schon beschrieben, können so Menschen, die aus verschiedenen Gründen nicht dazu in der Lage wären, mit Hilfe von KI ihre Ideen visualisieren. Dabei stellt sich aber die Frage, ob man die Nutzung von KI zum sichtbar machen seiner Gedanken mit künstlerischen Schaffen vergleichen kann. In den geführten Interviews waren die Teilnehmer einstimmig der Meinung, dass KI-Bilder keine Kunst

sind. Eine der Befragten beschreibt Kunst als „Ventil“. Ihr hilft Kunst sich zu entspannen und ob die Eingabe von Wörtern und auf die Ergebnisse warten einen ähnlichen Effekt hat, ist fraglich. KI scheint eine großartige Möglichkeit zu sein, Gedanken optisch darzustellen, aber der Prozess des Malens kann damit nicht simuliert werden, und für viele Künstler ist genau dieser Prozess das Wichtigste an ihrem Hobby, beziehungsweise Beruf. Auch wenn sich jeder über ein schönes Endergebnis freut. KI kann Kunst also, wenn überhaupt, nur bedingt demokratisieren. Trotzdem ermöglichen Text-zu-Bild-Modelle, Start-Ups und Kleinunternehmen mit wenig Budget, künstlerische Artefakte nutzen zu können, die sie sich sonst nicht leisten können. Dies bedeutet auf der anderen Seite aber natürlich weniger Arbeitsplätze oder Aufträge für Künstler. [Newton23] vergleicht die aktuelle Situation dabei mit der industriellen Revolution. Auch im Gespräch mit Designstudentin M fiel dieser Vergleich. Ihrer Meinung nach kann der Aufstieg von KI-Bildern auch zu einer gesteigerten Wertschätzung für von Menschen gemalten Bildern führen. Sie formuliert ihre These folgenderweise: „Es ist so ähnlich, wie bei der maschinellen Revolution. Die Maschine übernimmt einen Teil der Arbeit, dadurch bekommt aber auch das, was der Mensch macht mehr Wert.“ Ihre These wird auch durch [Newton23] gestützt. Dort wird beschrieben, dass Menschen durchaus bereit seien, mehr für handgefertigte Produkte, die in gewisser Weise einzigartig sind, zu bezahlen, als für Masseware. M könnte mit ihrer These also durchaus recht behalten, dass Menschen mehr für von Menschen gemachte Kunst bezahlen, wenn diese sich von KI-Bildern abhebt. KI-Bildgeneratoren schaffen also viele neue Möglichkeiten, bedeuten aber auch einen Wandel für Künstler und ihr Berufsleben. Außerdem ist es wichtig die richtigen Grenzen für diese noch recht neue Technologie zu setzen und dafür zu sorgen, dass die Technik sinnvoll und sicher eingesetzt werden kann. Vor allem Faktoren wie der Umwelteinfluss und das Urheberrecht müssen von außen geregelt werden. Außerdem müssen Unternehmen dafür sorgen, dass die Ergebnisse ihrer Bildgeneratoren sicher und inklusiv werden, und Künstler müssen sich überlegen, inwiefern sie mit KI-Bildgeneratoren interagieren und diese eventuell als Tool in ihren Arbeitsprozess einfließen lassen wollen.

9 Fazit

Durch die Untersuchungen, die im Rahmen dieser Arbeit durchgeführt wurden, konnten einige Erkenntnisse gesammelt werden. So führte das Parametertuning eines einfachen GAN in Kapitel 5 zu einer besseren Vorstellung des Aufwands vom Training eines KI-Bildgenerators. Auch wenn das Parametertuning nur beispielhaft war, so lässt sich daraus schließen, dass größere Modelle dieser Art, nur noch aufwendiger zu trainieren sind und somit eher von ressourcenstarken Unternehmen entwickelt werden können. Des Weiteren konnte mit dem Test vom Hive Moderation Tool gezeigt werden, dass Detektoren dieser Art zuverlässig in der Lage sind menschengemachte Zeichnungen von KI-generierten, künstlerischen Werken zu unterscheiden, da das Modell im durchgeführten Test alle Unterscheidungen richtig gemacht hat und lediglich bei der Benennung der zur Generierung verwendeten Modelle, Fehler aufgetreten sind. KI-Bilder von echten Zeichnungen zu unterscheiden ist händisch zwar etwas schwerer, doch kann man auch lernen Illustrationen mit bloßem Auge als KI-generiert oder menschengemacht zu identifizieren. Dies zeigte der für diese Arbeit durchgeführte Test, bei dem die Teilnehmer eine durchschnittliche Genauigkeit von 78,75% im Erkennen von KI-Bildern erreichten.

Auch die Forschungsfrage, inwieweit KI-Bildgeneratoren Einfluss auf Kunst und Künstler nehmen, kann durch die Untersuchungen und Literaturarbeit dieser Arbeit, zumindest ansatzweise beantwortet werden. Dass Künstler schon jetzt von der fortschreitenden Entwicklung von KI-Bildgeneratoren beeinflusst werden, ist nicht abzustreiten. Im Moment wirkt die Technologie auf die meisten Kunstschaffenden wie eine Bedrohung, die ihnen die Arbeitsplätze streitig macht. Immer mehr Unternehmen scheinen KI generierte Bilder in ihren Projekten zu verwenden. Nicht verwunderlich, dass diese Entwicklung manchen Künstlern, die von Auftragsarbeiten leben, erstmal Angst macht. Doch die Frage ist, ist diese Angst berechtigt und wird KI klassische Künstler aus ihrem Milieu verdrängen oder sind KI-Bildgeneratoren eine Unterstützung und Verbesserung des Künstlerberufs? Auch wenn KI den Eindruck macht immer bessere Ergebnisse zu erzielen, und die rapide Entwicklung in den letzten Jahren und die monetären Einsatzmöglichkeiten darauf schließen lässt, dass in den nächsten Jahren noch einige weitere Fortschritte folgen werden, können die KI-Bilder auch zum jetzigen Zeitpunkt zuverlässig von Kunstkennern als solche erkannt werden und Mechaniken, wie unsichtbare Wasserzeichen und Tools wie Hive Moderation tragen dazu bei, dass dies auch in Zukunft, wenn KI-Bilder keine fehlerhaften Artefakte mehr aufweisen, die mit dem bloßen Auge sichtbar sind, möglich sein wird. So könnte die Nutzung von KI-Bildern einige Konsumenten abschrecken, wobei die Kritik die ein oder andere Firma dazu bewegen könnte, weiterhin auf Kunst von Designern zurückzugreifen. Für viele Konsumenten scheint von Menschen geschaffene Kunst auch weiterhin einen höheren Wert als KI-Bilder zu haben. Jedoch gibt es auch Menschen, für die nur ein schönes Motiv zählt. Hierbei muss auch beachtet werden, dass KI-Bilder von Stable Diffusion [Esser24] und co. oft sehr generisch aussehen und spezielle Ziele damit nur schwer generiert werden können. Um ein Bild in einem spezielleren Stil zu erhalten, ist meistens ein auf diesen Stil trainiertes

Modell notwendig, auch weil viele Anwendungen, wie zum Beispiel DALL-E [Betker], [OpenAI23], die Generierung von Bildern, wie von einem speziellen Künstler gemalt, verhindern. Der Otto-Normalverbraucher wird nicht so weit gehen und muss so mit den generischeren Bildern Vorlieb nehmen oder einen menschlichen Künstler finden, der in dem gewünschten Stil malt.

Es scheint also so zu sein, dass KI-Bildgeneratoren vor allem eher Designer- und Illustratoren-Jobs übernehmen könnten, als dass sie die klassische Kunst ersetzen könnten. Vor allem da KI-Bilder für viele Leute keine Kunst in dem Sinne ist, da der menschliche Einfluss auf das Endergebnis und dessen Kompositionen fehlt. Man kann also durchaus davon ausgehen, dass die aktuelle Situation mit KI-Bildgeneratoren einen ähnlichen Effekt auf die Kunst haben wird wie die industrielle Revolution auf das Handwerk. Praktische Lösungen für Marketing und co. würden automatisiert von KI-Systemen erstellt werden und künstlerische Werke würden weiterhin bei Menschen in Auftrag gegeben werden, deren Aufwand dabei auch mehr geschätzt werden würde.

Diese Entwicklung scheint unvermeidbar zu sein und deswegen ist es so wichtig, dass Probleme im Zusammenhang mit KI-Bildgeneratoren angegangen und gemindert werden, damit die Entwicklung fair bleibt. Wichtig ist zum einen, dass diese Technologie nicht in schädlicher Art und Weise verwendet werden darf. Dieser Punkt gilt zwar vor allem bei fotorealistischen Bildern, ist aber auch bei Kunst sinnvoll, um Kinder oder auch traumatisierte Menschen, die gewisse Inhalte verstören und verletzen könnten, zu schützen. Des Weiteren sollte das Urheberrecht von KI-Bildern klar geregelt werden und das von Kunst in Trainingssätzen gewahrt werden. Künstler könnten so KI-Bildgeneratoren, sogar als Tool statt als Bedrohung sehen. So könnten sie zum Beispiel ihre Kunst als Trainingsmaterial für die KI-Modelle für kleine Beträge verkaufen oder sie könnten die Anwendungen selbst nutzen und in ihren Arbeitsablauf an verschiedenen Stellen einsetzen. Dies könnte sogar zu gesteigerter Produktivität ihres Schaffensprozesses führen. Die aktuelle Entwicklung fordert feste Regeln, damit Künstler nicht einfach von KI-Bildgeneratoren überrannt werden. Dann können sich Künstler auf KI-Bildgeneratoren und die mit ihnen eintretende Entwicklung anpassen. Denn klar ist, der Beruf von Künstlern wird sich durch diese Technik ändern, auch wenn Kunst selbst wohl noch etwas menschengemachtes bleiben wird.

Literaturverzeichnis

[Meeuwissen] H. M. Meeuwissen Voorhout: „Arum Discordis“ <https://meeuwissen-voorhout.nl/en/product/arum-dioscoridis/>

[Adobe] Adobe: „Nutze die Power der generativen KI Adobe Firefly“ <https://www.adobe.com/de/products/firefly.html>, 11.12.2024

[Bergmann24] Bergmann, Dave; Stryker, Cole: „Was ist ein Variational Autoencoder? IBM: 2024. <https://www.ibm.com/de-de/think/topics/variational-autoencoder>, 11.12.2024

[Bergmann24.2] Bergmann, Dave; Stryker, Cole: Was sind Diffusionsmodelle? IBM: 2024, <https://www.ibm.com/de-de/think/topics/diffusion-models#:~:text=Diffusionsmodelle%20geh%C3%B6ren%20zu%20den%20neuronalen,und%20Imagen%20von%20Google%20repr%C3%A4sentiert>, 13.12.2024

[Betker] Betker, James; Goh, Gabriel; Jing, Li; Brooks, Tim; Wang, Jianfeng; Ouyang, Long; Zhuang, Juntang; Lee, Joyce; Guo, Yufei; Manassra, Wesam; Dhariwal, Prafulla; Chu, Casey; Jiao, Yunxin; Ramesh, Aditya: Improving Image Generation with Better Captions. 2023, <https://cdn.openai.com/papers/dall-e-3.pdf> (10.12.2024)

[Bryne23] Bryne, Liam: “Say no to AI Art”, Yourstuffmade: 2023. <https://yourstuff-made.com/de-de/a/blog/say-no-to-ai-art#:~:text=AI%20Art%20Movement%E2%80%9D%3F-.The%20%22Say%20No%20to%20..AI%20on%20the%20art%20world>, 30.12.2024

[Canva] Canva: „Dream Lab. Welche Bilder möchtest du heute erstellen?“, <https://www.canva.com/dream-lab>, 08.01.2025

[Cetinic22] Cetinic, Eva; She, James: Understanding and Creating Art with AI: Review and Outlook. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 18.2 (2022), 1-22. <https://arxiv.org/abs/2102.09109>, 22.10.2024

[ChatGPT] “ChatGPT” <https://chatgpt.com/> 03.02.2025

[Christie’s] Christie’s: “Obvious and the interface between art and artificial intelligence”, 2018, <https://www.christies.com/en/stories/a-collaboration-between-two-artists-one-human-one-a-machine-0cd01f4e232f4279a525a446d60d4cd1> (24.11.2024)

[CocaCola23] The Coca Cola Company: “Coca-Cola Invites Digital Artists to ‘Create Real Magic’ Using New AI Platform”, 2023 <https://www.coca-colacompany.com/media-center/coca-cola-invites-digital-artists-to-create-real-magic-using-new-ai-platform>, 09.01.2025

[Corpnce] Corpnce: “Generative Adversarial Networks (GANS)” <https://www.corpnce.com/generative-adversarial-networks-gans/>, 03.02.2025

[DataFeedWatch24] DataFeedWatch blog: “11 Best AI Advertising Examples of 2024 cart”, 2024, <https://www.datafeedwatch.com/blog/best-ai-advertising-examples>, 06.01.2025

[Davis24] Davis, Vicki: "ChatGPT & Generative AI. Incorporating AI Art Generation in the Classroom", Edutopia: 2024, <https://www.edutopia.org/article/generative-ai-art-school/>, 10.01.2025

[Dhariwal21] Dhariwal, Prafulla; Nichol, Alex: Diffusion Models Beat GANs on Image Synthesis. Advances in neural information processing systems, 34 (2021), 8780-8794. <https://arxiv.org/pdf/2105.05233>, 27.01.2025

[Duden] Duden: "Kunst, die", <https://www.duden.de/rechtschreibung/Kunst>, 03.02.2025

[Edwards22] Edwards, Benj: "Artists stage mass protest against AI-generated artwork on ArtStation. Users of popular portfolio site seek to castigate and disrupt AI-generated art", arstechnica: 2022, <https://arstechnica.com/information-technology/2022/12/artstation-artists-stage-mass-protest-against-ai-generated-artwork/>, 07.01.2025

[Elgammal19] Elgammal, Ahmed: "AI Is Blurring the Definition of Artist. Advanced algorithms are using machine learning to create art autonomously", American Scientist: 2019, <https://www.americanscientist.org/article/ai-is-blurring-the-definition-of-artist>, 10.12.2024

[Esser24] Esser, Patrick; Kulal Sumith; Blattmann, Andreas; Entezari, Rahim; Müller, Jonas; Saini, Harry; Levi, Yam; Lorenz, Dominik; Sauer, Axel; Boesel, Frederic; Podell, Dustin; Dockhorn, Tim; English, Zion; Lacey, Kyle; Goodwin, Alex; Marek, Yannik; Rombach, Robin: Scaling Rectified Flow Transformers for High-Resolution Image Synthesis. Forty-first International Conference on Machine Learning: 2024. <https://arxiv.org/abs/2403.03206>, 22.10.2024

[Eugenio21] EugenioTL: „Variational Autoencoder structure“ Wikipedia: 2021, https://en.wikipedia.org/wiki/Variational_autoencoder#/media/File:VAE_Basic.png, 03.02.2025

[Everypixel24] Everypixel Journal: "People Are Creating an Average of 34 Million Images Per Day. Statistics for 2024", 2024, <https://journal.everypixel.com/ai-image-statistics>, 03.01.2025

[Flux] Flux.1 AI: "Flux AI Image Generator By Flux.1 AI. Create Flux AI images from Flux.1 AI for free", <https://flux1.ai/>, 11.01.2025

[Foster20] Foster, David: Generatives Deep Learning. Maschinen das Malen, Schreiben und Komponieren beibringen. Übersetzt von Marcus Fraaß und Konstantin Mack. O'Reilly: 2020

[Gaskins24] Gaskins, Nettrice: "Generative AI Art as a Therapeutic Tool", Medium: 2024, <https://nettricegaskins.medium.com/generative-ai-art-as-a-therapeutic-tool-ba72a57675e1>, 10.01.2024

[Goodfellow14] Goodfellow, Ian J., Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua: Generative Adversarial Nets. Advances in neural information processing systems 27 (2014), <https://arxiv.org/abs/1406.2661>, 10.12.2024

- [Google17] Google: „quickdraw_dataset“, 2017, <http://bit.ly/30HyNqg>, 03.02.2025
- [Gowal23] Gowal, Sven; Kohli, Pushmeet: “Identifying AI-generated images with SynthID”, Google DeepMind: 2023, <https://deepmind.google/discover/blog/identifying-ai-generated-images-with-synthid/>, 22.01.2025
- [Grba22] Grba, Dejan: Deep Else: A Critical Framework for AI Art. Digital 2.1 (2022), 1-32, <https://www.mdpi.com/2673-6470/2/1/1>, 22.10.2024
- [HiveDoc] “Image and Video Detection”, <https://docs.thehive.ai/docs/ai-image-and-video-detection>, 17.12.2024
- [HiveTool] “Ai-Generated Content Detection. Machine learning models to detect AI-generated content”, <https://hivemoderation.com/ai-generated-content-detection>, 17.12.2024
- [Holland23] Holland, Martin: „Secret Invasion‘ auf Disney+: Künstler kritisieren KI-Vorspann“, Heise online: 2023, <https://www.heise.de/news/Secret-Invasion-auf-Disney-Viel-Kritik-an-Intro-von-KI-Bildgenerator-9194847.html>, 13.12.2024
- [Imagen3] Imagen 3 Team; Google: „Imagen 3“, arXiv: 2408.07009 (2024), https://storage.googleapis.com/deepmind-media/imagen/imagen_3_report.pdf, 24.11.2024
- [Jiang23] Jiang, Harry; Brown, Lauren; Cheng, Jessica; anonymer Künstler; Khan, Mehtab; Gupta, Abhishek; Workman, Deja; Hanna, Alex; Flowers, Jonathan; Gebu Timnit: “AI Art and its Impact on Artists”, Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (2023), 363-374, <https://dl.acm.org/doi/10.1145/3600211.3604681>, 22.10.2024
- [Kingma22] Kingma, Diederik P.; Welling, Max: Auto-Encoding Variational Bayes. arXiv: 1312.6114 (2013), <https://arxiv.org/abs/1312.6114>, 09.11.24
- [Konen20] Konen, Wolfgang: „Die CO2-Kosten des Video-Streaming“, Die Neue Hochschule, Heft 3/2020 (2020), https://www.th-koeln.de/mam/downloads/deutsch/hochschule/aktuell/nachrichten/dnh_2020-3_vorveroeffentlichung_beitrag_konen.pdf, 22.01.2025
- [Luccioni24] Luccioni, Alexandra Sasha; Jernite, Yacine; Strubell, Emma: „Power Hungry Processing: Watts Driving the Cost of AI Deployment?“ The 2024 ACM Conference on Fairness, Accountability, and Transparency (2024), 85-99, <https://arxiv.org/pdf/2311.16863>, 03.01.2025
- [Metzmacher24] Metzmacher, David: „Urheberrecht und Trainingsdaten. Was muss man bei KI-Bildern beachten?“ zdfHeute: 2024 <https://www.zdf.de/nachrichten/wirtschaft/urheberrecht-kuenstliche-intelligenz-ki-internet-100.html>, 02.01.2025
- [Midjourney] „Midjourney“ <https://www.midjourney.com/home/>, 10.12.2024
- [Muenninghoff22] Muennighoff, Niklas; Wang, Thomas; Sutawika, Lintang; Roberts, Adam; Biderman, Stella; Scao, Teven Le; Bari, M Saiful; Shen, Sheng; Yong, Zheng-Xin; Schoelkopf, Hailey; Tang, Xiangru; Radev, Dragomir; Aji, Alham Fikri; Khalid Almu-barak, Khalid; Albanie, Samuel; Alyafeai, Zaid; Webson, Albert; Raff, Edward; Raffel,

Colin: „Crosslingual generalization through multitask finetuning“, arXiv: 2211.01786 (2022) <https://arxiv.org/pdf/2211.01786>, 22.01.2025

[Newton23] Newton, Alexis; Dhole, Kaustubh: „Is AI Art Another Industrial Revolution in the Making?“, arXiv: 2301.05133 (2023) <https://arxiv.org/abs/2301.05133>, 22.10.2024

[OpenAI23] OpenAI: „DALL·E 3 System Card“, 2023, <https://openai.com/index/dall-e-3-system-card/>, 10.12.2024

[OpenAI21] OpenAI: „CLIP: Connecting text and images“, 2021, <https://openai.com/index/clip/>, 05.12.2024

[OpenAIHelp] OpenAI: „C2PA in DALL·E 3. C2PA standard, OpenAI's implementation, and C2PA metadata“, <https://help.openai.com/en/articles/8912793-c2pa-in-dall-e-3>, 17.12.2024

[Po23] Po, Ryan; Wang, Yifan; Golyanik, Vladislav; Aberman, Kfir; Barron, Jonathan T.; Bermano, Amit H.; Chan, Eric Ryan; Dekel, Tali; Holynski, Aleksander; Kanazawa, Angjoo; Liu, C. Karen; Liu, Lingjie; Mildenhall, Ben; Nießner, Matthias; Ommer, Björn; Theobalt, Christian; Wonka, Peter; Wetzstein, Gordon: „State of the Art on Diffusion Models for Visual Computing“, arXiv: 2310.07204 (2023), <https://arxiv.org/abs/2310.07204>, 13.12.2024

[Quarks] Quarks: „Wie viel CO2 setzt du frei? CO2-Rechner für Auto, Flugzeug und Co“, <https://www.quarks.de/umwelt/klimawandel/co2-rechner-fuer-auto-flugzeug-und-co/>, 22.01.2025

[Radford21] Radford, Alec; Kim, Jong Wook; Hallacy, Chris; Ramesh, Aditya; Goh, Gabriel; Agarwal, Sandhini; Sastry, Girish; Askell, Amanda; Mishkin, Pamela; Clark, Jack; Krueger, Gretchen; Sutskever, Ilya: „Learning Transferable Visual Models From Natural Language Supervision“ International conference on machine learning, PMLR (2021), 8748-8763, <https://arxiv.org/abs/2103.00020>, 22.10.2024

[Ridler19] Ridler, Anna: „Mosaic Virus. 3-screen GAN video installation“, 2019. <https://annaridler.com/mosaic-virus>, 06.01.2025

[Rombach22] Rombach, Robin; Blattmann, Andreas; Lorenz, Dominik; Esser, Patrick; Ommer, Björn: High-Resolution Image Synthesis with Latent Diffusion Models. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (2022), 10684-10695, <https://arxiv.org/abs/2112.10752>, 22.10.2024

[Schuhmann21] Schuhmann, Christoph: „LAION-400-MILLION OPEN DATASET“, Laion: 2021, <https://laion.ai/blog/laion-400-open-dataset/>, 03.02.2025

[SDOnline] „Stable Diffusion Online“: <https://stablediffusionweb.com/>, 03.02.2025

[SDXL] „Stable Diffusion XL“, <https://stablediffusionxl.com/>, 03.02.2025

[Srinivasan21] Srinivasan, Ramya; Uchino, Kanji: „Biases in Generative Art. A Causal Look from the Lens of Art History“, Proceedings of the 2021 ACM Conference on

Fairness, Accountability, and Transparency (2021), 41-51
<https://arxiv.org/pdf/2010.13266>, 30.12.2024

[Stern24] Stern: „Weihnachts-Werbespots. Coca-Colas KI-generierte Videos ernten heftigen Spott“, 2024, <https://www.stern.de/wirtschaft/coca-colas-ki-generierte-videos-ernten-heftigen-spott-35253060.html>, 09.01.2025

[Vass24] Vass, Kate: „Harold Cohen: ‘Once upon a time there was an entity named Aaron’“ 2024, <https://www.katevassgalerie.com/blog/harold-cohen-aaron-computer-art>, 10.12.2024

[Zhou24] Zhou, Eric; Lee, Dokyun: „Generative artificial intelligence, human creativity, and art“ PNAS nexus 3.3 (2024), 52, <https://doi.org/10.1093/pnasnexus/pgae052>, 04.11.2024

[X1] @NoAiArt: „NO to AI Generated Images“, X <https://x.com/noaiart?mx=2>, 07.01.2025

Anhang

[Repository]: <https://github.com/nessa1107/bachelorarbeit-anhang>

Erklärung

Ich versichere, die von mir vorgelegte Arbeit selbstständig verfasst zu haben. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer oder der Verfasserin/des Verfassers selbst entnommen sind, habe ich als entnommen kenntlich gemacht. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit benutzt habe, sind angegeben. Die Arbeit hat mit gleichem Inhalt bzw. in wesentlichen Teilen noch keiner anderen Prüfungsbehörde vorgelegen.

Gummersbach, 07.02.2025

Ort, Datum

V. Ommers

Unterschrift