

Die Anwendung K.I gestützter Rotoskopie-Anwendungen im Bereich der digitalen Video- Postproduktion

Masterarbeit

Ausgeführt zum Zweck der Erlangung des akademischen Grades
Dipl.-Ing. für technisch-wissenschaftliche Berufe

am Masterstudiengang Digital Media Production an der Fachhochschule
St. Pölten, **Masterklasse Animation & Visual Effects**

von:

Matthias Bäck, BSc

mp211504

Betreuer*in: Dipl.-Ing. (FH) Mario Zeller

Zweitbetreuer*in: Mag. Franz Schubert

St. Pölten, 08.05.2023

Ehrenwörtliche Erklärung

Ich versichere, dass

- ich diese Arbeit selbständig verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt und mich auch sonst keiner unerlaubten Hilfe bedient habe.

- ich dieses Thema bisher weder im Inland noch im Ausland einem Begutachter/einer Begutachterin zur Beurteilung oder in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Diese Arbeit stimmt mit der vom Begutachter bzw. der Begutachterin beurteilten Arbeit überein.

St. Pölten, 08.05.2023

Ort, Datum



Unterschrift

Kurzfassung

Diese Masterarbeit gibt einen Überblick über die Anwendungsmöglichkeiten von Künstlicher Intelligenz (Kurzform K.I.) im Bereich der digitalen Video-Post Production und im Digital Design. Anhand einer Literaturrecherche wird zuerst der Begriff KI definiert und ein geschichtlicher Überblick gegeben. Es wird auf Anwendungsmöglichkeiten von K.I im Alltag, sowie im Kreativbereich, dem Digital Design und der digitalen Content-Produktion eingegangen. Zusätzlich werden Methoden zur Trennung von Vordergrund-Objekten und Hintergrund in Videomaterial vorgestellt.

Anschließend wird der Arbeits-Workflow der drei Anwendungen beschrieben, mit denen im Zuge dieser Arbeit Personen in Videos K.I gestützt rotoskopiert wurden.

Im empirischen Teil der Arbeit werden der Aufwand für die Bearbeitung, die Anzahl der Arbeitsschritte der Programme und die Renderzeit objektiv miteinander verglichen und die Ergebnisse dargestellt. Anhand einer quantitativen Online-Umfrage werden die optischen Qualitäten der Ergebnisse, welche mit den verwendeten Anwendungen erzeugt wurden, untersucht.

Die objektive Analyse ergab, dass sich zwei Anwendungen punkto Dauer für die Bearbeitung und Renderzeit ähnlich sind, die dritte Anwendung sich jedoch in den genannten Punkten stark unterscheidet. Ein direkter Zusammenhang zwischen längerer Renderzeit und optisch besseren Ergebnissen konnte nicht festgestellt werden.

Die Ergebnisse der quantitativen Online-Umfrage bestätigen die aufgestellte Hypothese. Während die Videos, in denen sich langsam und gleichmäßig bewegend Personen befinden, gute optische Ergebnisse lieferten, wurden in Videos mit schnellen und unregelmäßigen Bewegungen teils starke Artefakte erzeugt. Die Hintergründe haben ebenfalls einen Einfluss auf die Ergebnisse. Bei einem gleichmäßigen und ruhigen Hintergrund fielen die Ergebnisse besser aus als bei unregelmäßigen und unruhigen Hintergründen.

Die Arbeit bietet einen Mehrwert für Animatoren und VFX Artists, da der Einsatz von K.I gestützten Anwendungen zeitsparende Vorteile mit sich bringt. Überdies können K.I gestützte Anwendungen monotone, mühsame Tätigkeiten übernehmen, die ansonsten von Menschen manuell übernommen werden müssten.

Abstract

This master's thesis gives an overview of the possible applications of artificial intelligence (short form AI) in the field of digital video post production and digital design. Based on a literature research, the term AI is first defined and a historical overview is given. Possible applications of AI in everyday life, as well as in the creative area, digital design and digital content production are discussed. In addition, methods for separating foreground objects and background in video material are presented. The work steps of the three applications that were used in this thesis to rotoscope people in videos with AI support are then described.

In the empirical part of the work, the effort for editing, the number of work steps of the programs and the rendering time are objectively compared with each other and the results are presented. Using a quantitative online survey, the optical qualities of the results, generated with the applications used, are examined.

The objective analysis revealed that two applications are similar in terms of editing time and rendering time, but the third application differs greatly in the above points. A direct connection between a longer render time and optically better results could not be determined.

The results of the quantitative online survey confirm the proposed hypothesis. While the videos in which people are moving slowly and evenly delivered good optical results, strong artefacts were sometimes produced in videos with fast and irregular movements. The backgrounds also have an influence on the results. The results were better with an even and calm background than with irregular and restless backgrounds.

The work offers added value for animators and VFX artists, since the use of AI supported applications brings time-saving advantages. In addition, AI-supported applications can take over monotonous, tedious tasks that would otherwise have to be done manually by humans.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung	II
Kurzfassung	III
Abstract	IV
Inhaltsverzeichnis	V
1 Einleitung	8
1.1 Hypothese und Forschungsfragen	9
1.2 Motivation	10
1.3 Methoden	11
1.3.1 Literaturrecherche	11
1.3.2 Bearbeitung von Videos	12
1.3.3 Quantitative Umfrage	12
1.4 Gliederung der Arbeit	13
1.5 Definition wesentlicher Basisbegriffe	14
2 Künstliche Intelligenz	15
2.1 Definition	15
2.2 Geschichte	17
2.2.1 Anfänge in der K.I. Forschung	17
2.2.2 Erster K.I.- Winter	18
2.2.3 Expertensysteme	19
2.2.4 Zweiter K.I.- Winter	20
2.2.5 Intelligente Agenten	20
2.3 Funktionen der K.I.	21
2.4 Unterschied zwischen KI, Machine Learning und Deep Learning	23
2.5 Maschinelles Lernen	25
2.6 Starke und Schwache KI	29
2.6.1 Schwache KI (weak/narrow AI):	29
2.6.2 Starke künstliche Intelligenz (strong/general AI):	29
2.6.3 4 KI-Typen	30
2.6.4 Turing Test	31
2.7 Anwendungsmöglichkeiten von K.I.	32
2.7.1 K.I. Anwendungen im Alltag	33
2.7.2 Künstliche Intelligenz in der Forschung	34
2.7.3 Weitere Anwendungsmöglichkeiten von K.I.	35

2.8 Einfluss von KI auf die Arbeitswelt	37
2.9 Rechtliche Aspekte beim Einsatz von KI	38
2.9.1 Urheberrechtsschutz	38
2.9.2 Leistungsschutz	39
2.9.3 Patentrecht	41
2.9.4 Geheimnisschutz	42
2.10 Risiken und Herausforderungen	42
3 Künstliche Intelligenz im Kreativbereich, dem Digital Design und der digitalen Content-Produktion	45
3.1 Auswirkungen von Künstlicher Intelligenz in den Designprozess	47
3.2 Anwendungsmöglichkeiten von K.I in der digitalen Content-Produktion	49
3.2.1 Generative K.I	49
3.2.2 Künstliche Intelligenz in der Filmindustrie	50
4 Methoden zur Trennung von Vordergrund-Objekten und Hintergrund in Videomaterial	54
4.1 Rotoscoping	55
4.2 Luma-Key Mattes	55
4.3 Chroma-Key Mattes	56
4.4 Difference Mattes	57
5 Workflow der K.I gestützten Rotoskopie-Anwendungen	60
5.1 Adobe After Effects	60
5.2 The Foundry Nuke	63
5.3 Runway ML	66
6 Empirischer Teil	69
6.1 Aufgabenstellung an die Programme	69
6.2 Objektive Analyse der Anwendungen	71
6.2.1 Anzahl der Arbeitsschritte	71
6.2.2 Zeitaufwand	72
6.2.3 Renderzeit	74
6.3 Optische Qualitäten der Ergebnisse	75
6.3.1 Analyse der demoskopischen Fragen	75
6.3.2 Analyse der allgemeinen Fragen	76
6.3.3 Optische Analyse des ersten Videos	78
6.3.4 Optische Analyse des zweiten Videos	80
6.3.5 Optische Analyse des dritten Videos	82
6.3.6 Optische Analyse des vierten Videos	84
6.3.7 Optische Analyse des fünften Videos	86

6.3.8	Optische Analyse des sechsten Videos	88
7	Fazit	90
	Literaturverzeichnis	93
	Abbildungsverzeichnis	97
	Tabellenverzeichnis	99
	Anhang	100

1 Einleitung

Das Thema, mit dem sich diese Masterarbeit befasst, ist die Anwendung künstlicher Intelligenz (Kurzform: K.I.) in der digitalen Video-Post Production und im Digital Design. Die Verbreitung künstlicher Intelligenz im Alltag hat in den letzten Jahren stark zugenommen, sowohl im beruflichen als auch im privaten Kontext. Selbstfahrende Autos, humanoiden Robotern, AI-Assistenten wie Siri, Alexa und Cortana sowie Übersetzungstools wie DeepL sind nur einige Beispiele für Anwendungsbereiche künstlicher Intelligenz, die mittlerweile vielen Menschen vertraut sind (Schirlbauer, 2021).

Künstliche Intelligenz hat jedoch mittlerweile auch in den Bereichen Grafik und Design sowie Animation, Post Production und Visual Effects Einzug gefunden. In diesen Bereichen gibt es bereits sehr viele Tools, Programme und Anwendungen, die bestimmte Aufgabenbereiche und Arbeitsschritte übernehmen können und die Anwender bei Aufgaben unterstützen können (Schirlbauer, 2021).

Das Hauptaugenmerk dieser Arbeit ist die Anwendung von Künstlicher Intelligenz im Bereich der Rotoskopie, in Bezug auf die Trennung eines Objektes (im konkreten Fall Personen) im Vordergrund von dessen Hintergrund.

Ein wesentlicher Schritt bei der Arbeit mit visuellen Effekten und Kompositionen ist oft das Trennen eines Vordergrundobjekts von dessen Hintergrund. Durch das Erstellen einer Matte, die das Objekt isoliert, können Hintergründe ausgetauscht oder Effekte nur auf den Vordergrund angewendet werden (Adobe Inc., 2022).

Rotoscoping ist immer noch ein weitgehend manueller Prozess, der eine Reihe von Werkzeugen und einen erheblichen Zeitaufwand der Anwender fordert. Jüngste Forschungen im Bereich des maschinellen Lernens und der Bildsegmentierung haben einige Produkte hervorgebracht, die versprechen, den Prozess der digitalen Rotoskopie zu beschleunigen oder gar vollständig zu automatisieren (Bermudez, Dabby, Lin, Hilmarisdottir, Sundararajan, Kar, Dabby, Stringer, 2020, S.1).

Herkömmliches Rotoscoping erfordert die Verwendung animierter Masken, um ein Element in Videomaterial zu isolieren, was mühsam und zeitaufwändig sein kann. KI-gestützte Rotoskopie-Anwendungen bieten jedoch die Möglichkeit, Objekte im

Vordergrund vom Hintergrund zu trennen, ohne aufwendige Maskierungen durchführen zu müssen (Seymour, 2018).

Grundsätzlich ist es bereits seit längerer Zeit möglich, mittels manueller und zeitaufwendiger Methoden in Schnitt- und Compositing-Programmen Objekte vom Hintergrund zu isolieren. KI-gestützte Rotoskopie-Anwendungen versprechen jedoch eine einfachere, effizientere und schnellere Möglichkeit, auch Aufnahmen zu verwenden, die andernfalls nicht mehr verwertbar wären und neu gedreht werden müssten. Zudem sollen die KI Rotoskopie -Anwendungen insbesondere für Laien leichter anwendbar sein.

1.1 Hypothese und Forschungsfragen

Die Problemstellung, die den Ausgangspunkt dieser Arbeit darstellt, lautet: Wie, beziehungsweise mit welchen Hilfsmitteln/Anwendungen/Programmen, können Objekte effizient aus Video- und Bewegtbildmaterial rotoskopiert werden?

Daher sollen in dieser Arbeit Techniken und Programme vorgestellt werden, mit denen Objekte effizient aus Videomaterial rotoskopiert werden können. Die folgenden Programme/Anwendungen, mit denen K.I gestützte Rotoskopie möglich ist, sollen vorgestellt und miteinander verglichen werden:

- Adobe After Effects
- The Foundry Nuke, unter Anwendung der CopyCat Funktion
- Runway ML

Im Vorfeld wurden folgende Forschungsfragen definiert, die im Zuge dieser Arbeit beantwortet werden:

1. Bei welchen Tasks im Bereich der digitalen Video-Postproduktion können Artists durch den Einsatz von KI-Systemen aktuell unterstützt werden?
2. Welche Methoden eignen sich, um Vordergrund-Objekte in Videomaterial von dessen Hintergrund zu trennen?
3. Mit welchem Workflow werden Vordergrund-Objekte in den vorgestellten K.I gestützten Rotoskopie-Anwendungen von deren jeweiligen Hintergrund getrennt?
4. Worin unterscheidet sich der Aufwand (Anzahl der Arbeitsschritte, Zeitaufwand, Renderzeit) für den Anwender/die Anwenderin bei der Bearbeitung mit den vorgestellten Anwendungen/Techniken?

5. Wie unterscheiden sich die Ergebnisse der vorgestellten Anwendungen/Techniken hinsichtlich optischer Qualität?

Die sich aus den Forschungsfragen ergebende **Hypothese**, die in dieser Arbeit überprüft wird, lautet:

- Je nach Inhalt des verwendeten Videomaterials unterscheiden sich die Ergebnisse deutlich voneinander. Sollen beispielsweise Personen rotoskopiert werden, die schnelle oder unregelmäßige Bewegungen ausführen, werden vermutlich Artefakte, vor allem an den Objekt Kanten, ersichtlich werden. Überdies könnten dabei bestimmte Teile der Objekte (der Personen) nicht vollständig rotoskopiert werden und somit im Endergebnis fehlen. Hingegen bei langsamen und gleichmäßigen Bewegungen werden die Ergebnisse vermutlich wesentlich besser ausfallen.
- Einen weiteren Einfluss kann der Hintergrund des verwendeten Videomaterials haben. Bei einem gleichmäßigen und ruhigen Hintergrund fallen die Ergebnisse besser aus, als bei unregelmäßigen und unruhigen Hintergründen, in denen beispielsweise viele andere bewegte Objekte oder unregelmäßige Muster zu sehen sind.

1.2 Motivation

Rotoscoping ist oft eine der ersten Aufgaben, die ein angehender „Digital Composer“ erhält, wenn er/sie eine Stelle in einem Unternehmen für visuelle Effekte antritt. Obwohl Rotoscoping nicht als die glamouröseste Arbeit angesehen wird, ist es eine Tatsache, dass es sich um einen äußerst wichtigen Teil der Produktionspipeline handelt, der oft der entscheidende Faktor für den Erfolg oder Misserfolg bei der Bearbeitung einer bestimmten Aufnahme sein kann (Brinkmann, 2008, S.192).

Künstliche Intelligenz ist eine der in den letzten Jahren am schnellsten voranschreitende wissenschaftliche Technik, die für praktische Zwecke eingesetzt werden kann. Mit voranschreitender Rechenleistung und Speicherkapazität können immer mehr Tätigkeiten von künstlicher Intelligenz übernommen werden. Diese reichen von einfachen Aufgaben des täglichen Lebens bis hin zu hochspezifischen Anwendungen im Bereich der Wirtschaft und Technik. Diese Potenziale kommen jedoch nicht ohne Bedenken hinsichtlich sozialer Auswirkungen von K.I Technologien, die beim Entwurf und Einsatz dieser

Anwendungen berücksichtigt werden müssen (Anantrasirichai, Bull, 2021, S. 1 – 2).

Es gibt viele Situationen, in denen eine schnelle Hold-Out-Matte erforderlich ist. Oft wird dies durch eine schnelle manuelle Bearbeitung oder durch Fremdfirmen (häufig aus dem Ausland, v.a. in Entwicklungsländern) realisiert. Jedoch auch, wenn auf spezialisierte Firmen in Entwicklungsländern zurückgegriffen wird, in denen die Künstler nur geringe Löhne erhalten, ist dies mit einem erheblichen Kosten- und Zeitaufwand verbunden (Seymour, 2018).

In dieser Masterarbeit sollen mehrere Methoden der K.I unterstützten Rotoskopie vorgestellt und miteinander verglichen werden. Überdies soll ein Überblick über Anwendungsmöglichkeiten von künstlicher Intelligenz in der digitalen Video-Post Production, im Digital Design und im Bereich der Creative Industries gegeben werden. Anhand einer Literaturrecherche soll gezeigt werden, in welchen Bereichen der digitalen Video-Post Production, der Creative Industries und des Digital Designs künstliche Intelligenz eingesetzt werden kann und welche Vorteile oder Nachteile sich dadurch ergeben.

Diese Masterarbeit soll einen Mehrwert für Animatoren und VFX Artists bieten, da der Einsatz von K.I gestützten Anwendungen einerseits zeitsparende Vorteile mit sich bringt. Überdies können K.I gestützte Anwendungen monotone, mühsame Tätigkeiten übernehmen, die ansonsten von Menschen manuell übernommen werden müssten. Des Weiteren können K.I Anwendungen genutzt werden, um schnelle, erste Entwürfe zu generieren, die bei Entscheidungen über weitere Arbeitsverläufe helfen können.

Zusätzlich profitieren Laien oder Anfänger im VFX-Bereich davon, da sie mithilfe der K.I Anwendungen mit wenig Aufwand und Kenntnissen zu schnellen Ergebnissen gelangen können.

1.3 Methoden

1.3.1 Literaturrecherche

Um die theoretischen Fragen zu beantworten, wird analysiert, mittels welchen Tasks im Bereich der digitalen Video-Postproduktion Artists durch den Einsatz von KI-Systemen aktuell unterstützt werden können und welche Methoden sich eignen, um Vordergrund-Objekte in Videomaterial von dessen Hintergrund zu trennen. Dazu dient eine ausführliche Literaturrecherche. Als Literatur dienen sowohl Bücher als auch Fachartikel (Papers) und relevante Online-Artikel.

1.3.2 Bearbeitung von Videos

Für den praktischen und empirischen Teil der Arbeit, sowie zur Beantwortung der aufgestellten Hypothese, wurde Videomaterial verwendet, in dem sich Personen (reale Menschen) in unterschiedlichen Situationen, vor unterschiedlichen Hintergründen und unterschiedlichen Kameraeinstellungen befinden. Es wurden 6 Videos verwendet, in denen eine oder mehrere Personen aus unterschiedlichen Winkeln und Entfernungen aufgenommen wurden. Die Videos stammen von einer Online-Videoplattform und sind lizenztechnisch kostenlos und ohne Namensnennung zu verwenden.

Das gleiche Videomaterial wurde jeweils mit 3 unterschiedlichen Programmen beziehungsweise Anwendungen (Adobe After Effects, The Foundry Nuke-CopyCat Funktion, Runway ML) bearbeitet, wobei die beinhalteten Personen, K.I gestützt, rotoskopiert wurden. Eine kurze Videosequenz (ca. 5 bis 10 Sek.) mit dem bearbeiteten Material wurde exportiert.

Während der Bearbeitung wurde gemessen und analysiert, mit welchem Workflow in den jeweiligen Anwendungen die K.I gestützte Rotoskopie durchgeführt wurde. Der Aufwand für die Bearbeitung (Anzahl der Arbeitsschritte, Zeitaufwand, Renderzeit) wurde gemessen und in Diagrammen und Tabellen gegenübergestellt.

1.3.3 Quantitative Umfrage

Um den empirischen Teil zu vervollständigen wurde eine quantitative Online-Umfrage durchgeführt, in der die optischen Qualitäten der Ergebnisse miteinander verglichen wurden. Dadurch wird die 5. Forschungsfrage beantwortet.

Zuerst wurden hierfür allgemeine und demoskopische Fragen von den Teilnehmern beantwortet. Anschließend sollten jeweils die bearbeiteten Videosequenzen angesehen, und dazu gestellte Fragen beantwortet werden.

Die Videos wurden im Fragebogen nebeneinander platziert, damit diese besser optisch miteinander verglichen werden konnten. Der Hintergrund wurde neutral gehalten und auf das jeweilige Fototage angepasst. In den Vergleich-Videos der Umfrage wurde der Fokus auf die rotoskopierten Personen gelegt.

Der Fragebogen wurde anschließend über mehrere Verteiler an Studienkollegen und Bekannte übermittelt und in Online Foren für empirische Umfragen gestellt. So wurde gewährleistet, dass 81 Personen für die Befragung gefunden wurden. Anschließend erfolgten eine Auswertung und Analyse der Befragung.

Die Ergebnisse der Umfrage wurden zusammengefasst und analysiert. Dazu wurden Diagramme erstellt, um die Ergebnisse der Fragen optisch darzustellen.

1.4 Gliederung der Arbeit

Das erste Kapitel dieser Arbeit führt an das Thema heran und stellt die Hypothese und Forschungsfragen vor, die im Laufe der Arbeit beantwortet werden. Die verwendeten Methoden für den theoretischen und empirischen/praktischen Teil der Arbeit werden vorgestellt und die Gliederung der Arbeit, sowie die Definition wesentlicher Basisbegriffe wird beschrieben.

Das zweite Kapitel befasst sich umfangreich mit dem Thema Künstliche Intelligenz (K.I). Anfänglich wird der Begriff definiert und ein geschichtlicher Überblick geboten. Es werden die Funktionen der K.I erläutert und auf die Unterschiede zwischen KI, Machine Learning und Deep Learning eingegangen. Die Unterschiede zwischen einer „starken K.I.“ und einer „schwachen K.I.“ werden erläutert, sowie der bekannte „Turing Test“ vorgestellt. Nach den allgemeinen Definitionen wird auf Anwendungsmöglichkeiten von K.I eingegangen, sowohl bezüglich Anwendungen im Alltag als auch in der Forschung und weiteren Bereichen. Anschließend wird erläutert, welchen Einfluss K.I auf die Arbeitswelt hat und welche rechtlichen Aspekte beim Einsatz von K.I zu beachten sind. Abschließend werden Chancen, Risiken und Herausforderungen gegenübergestellt.

Im dritten Kapitel wird gezeigt, wie K.I im Kreativbereich, dem Digital Design und der digitalen Content-Produktion eingesetzt werden kann und welche Auswirkungen K.I in den Designprozess bringt.

Das vierte Kapitel beschäftigt sich mit allgemeinen und klassischen Methoden zur Trennung von Vordergrund-Objekten und Hintergrund in Videomaterial. Hierbei wird auf den Rotoskopie Prozess allgemein eingegangen und verschiedene Arten von Mattes vorgestellt.

Der Workflow der K.I gestützten Rotoskopie-Anwendungen, die in dieser Arbeit zum Einsatz kommen, wird im 5. Kapitel erklärt. Dafür werden die einzelnen, notwendigen Arbeitsschritte allgemein beschrieben.

Das sechste Kapitel ist dem empirischen Teil der Arbeit gewidmet. Hierfür wird zuerst die Aufgabenstellung an die Programme/Anwendungen erläutert und eine objektive Analyse (Anzahl der Arbeitsschritte, Zeitaufwand, Renderzeit) durchgeführt. Mittels einer quantitativen Umfrage wurden anschließend die optischen Qualitäten der Ergebnisse überprüft und veranschaulicht.

Das Fazit der Arbeit (7. Kapitel) fasst die Erkenntnisse aus der Literaturrecherche sowie die Ergebnisse der empirischen Studie zusammen.

1.5 Definition wesentlicher Basisbegriffe

Künstliche Intelligenz (K.I) beschreibt die Fähigkeit von Maschinen, basierend auf Algorithmen, Aufgaben autonom auszuführen und dabei anpassungsfähig auf unbekannte Situationen zu reagieren. Ihr Verhalten ähnelt damit dem menschlichen: Sie führen nicht nur repetitive Aufgaben aus, sondern lernen aus Erfolg und Misserfolg und passen ihr Verhalten entsprechend an. Zukünftig sollen Künstliche-Intelligenz-Maschinen (KIM) auch in der Lage sein, wie Menschen zu denken und zu kommunizieren (Stadler, 2022, S.1).

Machine Learning ist ein Teilgebiet der KI, das sich mit der Entwicklung von Algorithmen und Modellen befasst, die es Computern ermöglichen, aus historischen Daten zu lernen (Stadler, 2019, S.8).

Deep Learning ist ein Teilgebiet von Machine Learning, bei dem die Lernmethoden der KI am Aufbau des menschlichen Gehirns orientiert sind (Stadler, 2019, S.8).

Generative KI ist eine Technologie, die maschinelles Lernen und künstliche Intelligenz einsetzt, um neue Inhalte aus vorhandenen Daten zu generieren. Im Gegensatz zu anderen Ansätzen des maschinellen Lernens, die darauf abzielen, vorhandene Muster zu erkennen und Vorhersagen zu treffen, geht es bei der generativen KI darum, kreative Prozesse zu automatisieren und neue, bisher unbekannte Inhalte zu generieren (Industry of Things, 2022).

Der Begriff **Rotoscoping** wurde ursprünglich verwendet, um einen Arbeitsablauf zu beschreiben, der auf jedem einzelnen Frame einer Sequenz durchgeführt werden musste. Seitdem hat er sich zu einer allgemeineren Beschreibung eines Prozesses entwickelt, der wo immer möglich, prozedurale oder algorithmische Unterstützung einschließt. Konkret bezieht es sich in der Regel auf den Prozess der Erstellung einer **Matte** für ein bestimmtes Objekt in einer Szene, indem man eine Spline-Form an bestimmten Schlüsselbildern der Bildsequenz definiert und eine Software die Form für die Zwischenbilder interpolieren lässt (Brinkmann, 2008, S. 192).

2 Künstliche Intelligenz

2.1 Definition

Künstliche Intelligenz (KI) ist ein Begriff, der seit Jahrzehnten Menschen auf der ganzen Welt beschäftigt. Bis zum heutigen Tag gibt es jedoch keine allgemein akzeptierte oder eindeutige Definition dafür. Cambridge Advanced Learner's Dictionary definiert den Begriff Intelligenz als die Fähigkeit zu lernen, zu verstehen und auf Vernunft basierend Entscheidungen zu treffen. („Intelligence is the ability to learn, understand and make judgements based on reason“). Im Duden wird Intelligenz als die Fähigkeit [des Menschen], abstrakt und vernünftig zu denken und daraus zweckvolles Handeln abzuleiten, definiert (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.12).

Eine der ersten Definitionen für Künstliche Intelligenz aus dem Jahr 1955 bezieht sich auf John McCarthy und die sogenannte "Dartmouth-Konferenz". Anfang September 1955 reichte der damals 28-jährige Assistenzprofessor für Mathematik am Dartmouth College in Hanover, New Hampshire, gemeinsam mit drei Kollegen einen Antrag zur Förderung eines ambitionierten Projekts bei der Rockefeller-Stiftung ein. Im Sommer 1956 wurde eine Gruppe von zehn ausgewählten Forschern auserwählt, um an der ersten Konferenz über Künstliche Intelligenz, auch als "Dartmouth-Konferenz" bekannt, teilzunehmen. Die Forscher hatten zwei Monate, um zu versuchen herauszufinden, "wie Maschinen dazu gebracht werden können, Sprache zu benutzen, Abstraktionen und Begriffe zu bilden, Probleme zu lösen, die bislang nur dem Menschen vorbehalten waren, und sich selbst zu verbessern". Dieses Projekt, das auch als „Artificial Intelligence“ bezeichnet wurde, hatte das Ziel, Maschinen zu entwickeln, die ein vergleichbares Verständnis für komplexe Probleme wie Menschen aufweisen. Die „Dartmouth-Konferenz“ wird heute als Startschuss der KI-Forschung definiert (Buchkremer, Heupel, Koch, 2020, S.4).

Elaine Rich definiert KI als die Studie darüber, wie man Computer dazu bringen kann, Dinge zu tun, bei denen Menschen derzeit besser sind. („Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.“) Ein Beispiel hierfür ist die Fähigkeit, einen unbekannten Raum schnell zu erkennen und Entscheidungen zu treffen, die auf dieser Erkennung basieren. Obwohl Menschen diese Aufgabe mühelos bewältigen können, sind autonome Roboter damit noch überfordert. Rich sieht jedoch gerade in der Forschung an autonomen Robotern eines der wichtigsten Felder der KI. Hier geht es darum, die Fähigkeiten von Robotern und anderen Maschinen so zu verbessern, dass sie komplexe Aufgaben eigenständig ausführen können, ohne

dass sie ständig von Menschen gesteuert werden müssen. Insgesamt beschäftigt sich die KI also mit der Entwicklung von Technologien, die es Computern ermöglichen, menschenähnliche Fähigkeiten zu erlangen und komplexe Probleme zu lösen (Ertel, 2016, S. 3).

Die Versuche, Intelligenz zu beschreiben und nachzubilden, lassen sich grob in vier Ansätze einteilen, die sich mit menschlichem Denken, menschlichem Handeln, rationalem Denken und rationalem Handeln befassen. Ein bekanntes Beispiel für menschliches Handeln ist der sogenannte Turing-Test, bei dem eine KI menschliches Verhalten perfekt reproduzieren muss. Programme zur Bilderkennung und Entscheidungsfindung werden dagegen beispielsweise eher im Bereich des rationalen Handelns verortet. Die philosophische Debatte zur KI beschäftigt sich neben den Schwierigkeiten in der Definition auch mit den Unterschieden und Konsequenzen zwischen:

- einer schwachen oder eingeschränkten KI, die spezifische Probleme intelligent lösen kann,
- einer starken oder generellen KI, die allgemeine Probleme ebenso gut wie Menschen lösen kann, und
- einer künstlichen Superintelligenz, die die Fähigkeiten von Menschen bei weitem übertrifft.

(Wittpahl, 2019, S.21)

Eine weitere Definition, um Künstliche Intelligenz zu beschreiben lautet folgendermaßen:

„Künstliche Intelligenz beschreibt die Fähigkeit von Maschinen, basierend auf Algorithmen Aufgaben autonom auszuführen und dabei anpassungsfähig auf unbekannte Situationen zu reagieren. Ihr Verhalten ähnelt damit dem menschlichen: Sie führen nicht nur repetitive Aufgaben aus, sondern lernen aus Erfolg und Misserfolg und passen ihr Verhalten entsprechend an. Zukünftig sollen Künstliche-Intelligenz-Maschinen (KIM) auch in der Lage sein, wie Menschen zu denken und zu kommunizieren.“ (Stadler, 2022, S.1)

Obwohl es eine Vielzahl von Ansätzen und Definitionen gibt, lässt sich ein zentraler Aspekt der KI benennen, den alle als KI bezeichneten Systeme aufweisen: Es handelt sich um den Versuch, ein System zu entwickeln, das in der Lage ist, eigenständig komplexe Probleme zu bearbeiten. Das bedeutet, dass die KI-Systeme in der Lage sein müssen, Informationen zu sammeln, zu verarbeiten, Entscheidungen zu treffen und Handlungen auszuführen, um ein bestimmtes Ziel zu erreichen. Es gibt viele verschiedene Wege, das vielschichtige Forschungsgebiet der KI und seiner zahlreichen Unterkategorien zu beschreiben. Einige Ansätze befassen sich mit den Herausforderungen, die auf dem Weg zur

Entwicklung intelligenter Computersysteme auftreten, andere mit den Lösungsansätzen für diese Probleme und wieder andere mit dem Vergleich zur menschlichen Intelligenz (Wittpahl, 2019, S.21).

2.2 Geschichte

In den vergangenen Jahren waren Digitale Transformation, Industrie 4.0 und künstliche Intelligenz (KI) zentrale Themen in der Diskussion über die Entwicklung von Unternehmen und ganzen Geschäftsfeldern. Die Möglichkeiten von KI wurden sowohl positiv als auch negativ bewertet. Manche sahen in KI eine Heilsbringerin, während andere befürchteten, dass sie unheilbringende Eigenschaften besitzt. Die Interpretation ihrer Auswirkungen auf die Gesellschaft und den Unternehmensalltag variierte von der Vorstellung vollautonomer Roboter, die menschliche Arbeitskräfte komplett ersetzen, bis hin zu der Annahme, dass es sich lediglich um eine Blase handelt, die bald platzen wird und KI keinen nennenswerten, nachhaltigen wirtschaftlichen Effekt haben wird (Wennker, 2020, S.1).

Allerdings gibt es Studien, die zeigen, dass knapp 40% der als KI-Start-ups gehandelten Unternehmen gar keine künstliche Intelligenz nutzen. Laut einer Studie des deutschen Bundeswirtschaftsministeriums nutzen gerade einmal 6% aller deutschen Unternehmen künstliche Intelligenz. Dies könnte zum Teil auf überzogene Prognosen in der Frühzeit der künstlichen Intelligenz zurückzuführen sein, kombiniert mit popkulturellen Darstellungen von autonomen, tödlichen Maschinen oder wahnsinnig gewordenen Computern, die bei Unternehmen zu unrealistischen Erwartungen führen könnten (Wennker, 2020, S.1).

Ein weiteres Problem besteht darin, dass Begriffe wie Machine Learning, Deep Learning und schwache und starke künstliche Intelligenz nicht klar definiert sind und somit zu Verwirrung führen können. Wie bei fast allen neuen Technologien durchlief die Entwicklung der künstlichen Intelligenz Zyklen von großer Begeisterung und Aufmerksamkeit, gefolgt von Jahren nur marginaler Entwicklung und Förderung. Künstlicher Intelligenz ist es hier ähnlich ergangen wie der Elektrizität, den Eisenbahnen oder der Dotcom-Blase, allgemein im Englischen beschrieben als Hype Cycle. Auf überbordende Versprechen folgten Phasen tiefer Ernüchterung, gefolgt von wieder ansteigendem Interesse und zunehmend praktischeren Anwendungen (Wennker, 2020, S.1-2).

2.2.1 Anfänge in der K.I Forschung

Die Idee von künstlichen Lebewesen, die in der Lage sind, formelles Denken anzuwenden, beschäftigt Philosophen und Naturwissenschaftler seit vielen

Jahrhunderten und ist in fast allen Kulturen und Gesellschaften zu finden. Die Geburtsstunde der akademischen Auseinandersetzung mit künstlicher Intelligenz wird gemeinhin auf die Dartmouth Conference im Sommer 1956 zurückgeführt, die auch als Dartmouth Summer Research Project on Artificial Intelligence bekannt ist. Die Forschungsziele der Konferenz waren darauf ausgerichtet zu untersuchen, ob jeder Aspekt des Lernens oder anderer Merkmale der Intelligenz so präzise beschrieben werden kann, dass eine Maschine es simulieren kann. („The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.“) (Wennker, 2020, S.2)

In den Jahren nach der Dartmouth Conference wurden bahnbrechende Fortschritte im Bereich der künstlichen Intelligenz erzielt, die zuvor für unmöglich gehalten worden waren. Zum Beispiel wurden Programme wie ELIZA und STRIPS entwickelt. Diese Fortschritte führten dazu, dass Forscher optimistische Prognosen für zukünftige Entwicklungen abgaben, wie H.A. Simon, der sagte: "Maschinen werden in der Lage sein, innerhalb von zwanzig Jahren jede Arbeit zu erledigen, die ein Mensch tun kann." (Wennker, 2020, S.2-3)

2.2.2 Erster K.I- Winter

Infolge der Versprechen, die mit der Erforschung künstlicher Intelligenz verbunden waren, investierten Regierungsorganisationen Millionen in diese Forschungsrichtung, um unter anderem einen Vorteil im Kalten Krieg zu erlangen. Ein Beispiel hierfür ist die Erforschung der maschinellen Übersetzung, die die Herausforderungen und technischen Entwicklungen der folgenden Jahre verdeutlicht. Die US-amerikanische Defense Advanced Research Projects Agency (DARPA) investierte bis in die 1960er Jahre 20 Millionen US-Dollar in die Entwicklung von Systemen, die eine Übersetzung vom Russischen ins Englische ermöglichen sollten. Die Forschungen in diesem Bereich des Natural Language Processing stießen schnell an Grenzen, bedingt zum einen durch die fehlende Rechenleistung der damaligen Computer und zum anderen durch ein Problem, das auch heute noch viele Bereiche der künstlichen Intelligenz herausfordert: das fehlende Hintergrundwissen. Um eine Sprache mit Umschreibungen, Idiomen, Disambiguationen und ähnlichen Sprachphänomenen zu verstehen, muss ein System auf eine große Datenbank mit Informationen zur beschriebenen Welt zugreifen können, was in den 1960er Jahren noch unmöglich war. Durch diesen Umstand waren die Geldgeber mit den langsamen Fortschritten unzufrieden und setzten eine Kommission zur Überprüfung der Ergebnisse ein. Der im Jahr 1966 veröffentlichte Bericht des Automatic Language Processing Advisory Committee (ALPAC) kam zu dem Schluss, dass maschinelle Übersetzungen langsamer, weniger genau und deutlich teurer waren als menschliche Übersetzer. In denselben Zeitraum fiel der Lighthil-Bericht (1973) im Vereinigten Königreich, der

zur fast vollständigen Beendigung von Forschungen im Bereich künstlicher Intelligenz in Großbritannien führte. Obwohl damit die erste Epoche der Künstlichen-Intelligenz-Forschung annähernd abrupt zu Ende ging, wurden in den Anfangsjahren einige wichtige Entdeckungen gemacht, vom Perceptron, das Jahre später im Bereich der neuronalen Netze einen Aufschwung erhielt, bis hin zu Sprachassistenten (Speech Understanding Research an der Carnegie Mellon University) (Wennker, 2020, S.3-4).

2.2.3 Expertensysteme

Die zweite Welle der künstlichen Intelligenz wurde durch die Entwicklung von Expertensystemen begründet, die speziell für ein enges Themenfeld entwickelt wurden, um regelbasierte Antworten zu geben. Eines der ersten Expertensysteme war MYCIN, das 1972 an der Stanford University entwickelt wurde. MYCIN zeigte bereits alle Attribute späterer Expertensysteme, indem es anhand von Ja- oder Nein-Fragen, die von einem Arzt zu einem Patienten gestellt wurden, auf eine Bakterieninfektion schloss und ein Antibiotikum inklusive Dosierung empfahl. Aus ethischen und rechtlichen Bedenken wurde MYCIN jedoch nie in der Praxis eingesetzt, lieferte aber die Vorlage für folgende Expertensysteme.

Das erste produktive Expertensystem wurde vom Computerhersteller DEC mit dem R1/XCON eingeführt, um fehleranfällige Bestellungen von Computerhardware zu reduzieren. Basierend auf 2500 Regeln gab der R1 den Verkäufern Fragen vor, um eine sinnvolle Bestellung zusammenzustellen. Die geschätzten Einsparungen durch den Einsatz von R1 beliefen sich auf 25 Mio. US-Dollar pro Jahr. Dem Beispiel folgend investierten Unternehmen 1985 eine Milliarde US-Dollar in die Entwicklung eigener Expertensysteme. Eine eigene Industrie entwickelte sich schnell um diese Expertensysteme, die Hard- und Software bereitstellte.

1982 investierte das japanische Ministerium für Handel und Industrie das Äquivalent von 850 Mio. US-Dollar in das Fifth Generation Computer Project, mit dem Ziel, einen Computer oder eine Software zu entwickeln, die Bilder erkennen, Übersetzungen liefern und argumentieren können sollte wie ein menschliches Wesen. Andere Nationen, allen voran die USA, sahen in der Fifth Generation Computer Project Initiative die eigene technologische Vormachtstellung bedroht und starteten eigene Initiativen. Die von der US-Regierung finanzierte Strategic Computing Initiative verschlang mehr als eine Milliarde US-Dollar, um innerhalb von zehn Jahren eine komplette Maschinenintelligenz zu schaffen. Ein ähnliches privatwirtschaftliches Engagement ging in den USA von der Microelectronics and Computer Technology Corporation aus. In Großbritannien wurde von der Regierung analog das Programm Alvey gefördert (Wennker, 2020, S.4-5).

2.2.4 Zweiter K.I.- Winter

In den späten 1980er Jahren begann der Niedergang der Expertensysteme und der sie umgebenden Industrie. IBM und Apple stellten Computer vor, die die Leistung spezialisierter Systeme wie die Lisp-Maschine, auf der Expertensysteme basierten, in den Schatten stellten. Für Unternehmen wurde es immer unwirtschaftlicher, in dedizierte Hardware und schwer zu wartende und aktualisierbare Systeme zu investieren. Gleichzeitig wurden die Ergebnisse des Fifth Generation Computer Projects, das das Ziel hatte, einen Computer oder eine Software zu entwickeln, die in der Lage war, wie ein menschliches Wesen zu argumentieren, Bilder zu erkennen und Übersetzungen zu liefern, evaluiert und zeigten, dass viele der hochgesteckten Ziele nicht annähernd erreicht wurden. Ähnlich erging es der Strategic Computing Initiative und Alvey, die keine weitere Finanzierung erhielten. Trotz des Niedergangs der Expertensysteme wurden in den 1980er Jahren wichtige Errungenschaften in der künstlichen Intelligenz entwickelt, die in den folgenden Jahren die Entwicklung maßgeblich beeinflussten. Hierzu gehören das Hopfield-Netz und die Backpropagation, die bis heute wichtige Elemente in neuronalen Netzen darstellen. Obwohl die Expertensysteme und die sie umgebende Industrie an Bedeutung verloren haben, trugen ihre Entwicklungen und Errungenschaften wesentlich zur Entwicklung der künstlichen Intelligenz bei und legten den Grundstein für weitere Entwicklungen in diesem Bereich (Wennker, 2020, S.5).

2.2.5 Intelligente Agenten

In den 1990er und frühen 2000er Jahren erhielten intelligente Agenten Aufmerksamkeit, die durch mediale Auftritte wie die Schachmatches zwischen Gary Kasparov und IBMs Deep Blue bekannt wurden. Intelligente Agenten sind Systeme, die ihre Umgebung wahrnehmen und Handlungen durchführen, um ihre Chancen auf Erfolg zu erhöhen. Dieser recht umfassenden Definition folgend, sind einfachste Computerprogramme bis hin zu hochkomplexen Lebewesen, wie Menschen intelligente Agenten. Die Definition von künstlicher Intelligenz ergibt sich daraus als "das Studium von intelligenten Agenten". Ende der 1990er Jahre entstand ein System, das auf einer einheitlichen mathematischen Notation basierte und es ermöglichte, viele komplexe Probleme schrittweise zu lösen. Die Idee einer umfassenden, allgemeinen künstlichen Intelligenz trat in den Hintergrund. Auch der Begriff künstliche Intelligenz selbst wurde vermieden, um nicht als unrealistischer Traum abgetan zu werden.

In den 2000er Jahren änderte sich dies durch die fortschreitende Digitalisierung der Arbeits- und Privatwelt sowie die zunehmende Vernetzung. Immer größere Datenmengen fielen an, die mit herkömmlichen Methoden nicht mehr sinnvoll analysiert werden konnten. Große Datenanalysen wurden als wichtige Grundlage

des Wettbewerbs betrachtet und als Möglichkeit, Produktivitätswachstum, Innovation und Konsumentenrente zu fördern. Technikzentrierte US-Konzerne wie Google, Apple, Facebook und Amazon treiben die Entwicklung voran.

Künstliche Intelligenz durchdringt mittlerweile viele Bereiche des Alltags und erledigt eng gefasste Aufgaben auf einem Niveau, das dem des Menschen entspricht oder es teilweise sogar übertrifft. Die in den Medien dargestellte künstliche Intelligenz, die alle Aufgaben mit gleicher Effizienz erledigt und jede Aufgabe erlernen kann, wird als künstliche generelle Intelligenz bezeichnet. Es ist jedoch unklar, ob und wann diese erreicht werden kann. Es gibt Zweifel, ob die bisherigen Algorithmen ausreichen, um künstliche generelle Intelligenz zu erreichen. Es ist überdies fraglich, ob genügend Daten vorhanden sind und ob die Technologie ausreichend entwickelt ist. Experten schätzen, dass künstliche generelle Intelligenz zwischen 2030 und 2060 erreicht werden kann. Dennoch entwickelt sich die Forschung und damit auch die Befähigung von neuronalen Netzen und künstlicher Intelligenz sehr schnell und dringt in kognitive Bereiche vor, die bisher dem Menschen vorbehalten waren.

(Wennker, 2020, S.5-8)

2.3 Funktionen der K.I

KI-Systeme können für unterschiedlichste Anwendungen und Funktionen eingesetzt werden. In den meisten Fällen wird KI dafür genutzt, um zumindest eine der nachfolgenden sieben allgemeinen Funktionen auszuführen:

Im Bereich des Monitorings können KI-Systeme große Mengen an Daten in kürzester Zeit analysieren und Abweichungen sowie Muster feststellen. Aufgrund ihrer Fähigkeit, dies in Echtzeit und genauer als Menschen zu tun, eignen sich KI-Systeme besonders gut für Anwendungen im Bereich der Cyber-Security und der Umweltüberwachung.

Durch das Auffinden von Mustern und Trends können KI-Systeme wertvolle Erkenntnisse aus großen Datenmengen extrahieren (Data Mining) und neue Lösungen durch Simulationen erarbeiten. Da KI-Systeme dynamische Modelle verwenden, die aus den Daten lernen und sich anpassen, können sie sehr effektiv beim Entdecken von abstrakten Mustern und neuen Erkenntnissen sein, welche von traditionellen Computerprogrammen nicht erkannt werden würden.

KI-Systeme können auch Vorhersagen treffen und Modelle darstellen, wie sich Trends in der Zukunft entwickeln werden. Auf diese Weise können Systeme Antworten vorhersagen, empfehlen und personalisieren.

Durch die Fähigkeit, unstrukturierte Daten wie Bilder, Videos, Audiodateien und Textdateien zu interpretieren, also solche, die nur schwer einzuordnen sind, können KI-Systeme eine vielfach größere Datenmenge analysieren als traditionellen Datenanalysemethoden.

KI-Systeme ermöglichen eine Vielzahl von Interaktionen zwischen Maschinen und ihrer Umgebung. Beispiele hierfür sind Robotersysteme, die durch ihre Umgebung navigieren und diese manipulieren, sowie selbstfahrende Autos, die eine Unmenge an Echtzeitdaten analysieren und eine sichere und effiziente Fahrtrichtung bestimmen.

KI-Systeme erleichtern auch die Interaktion zwischen Menschen und Computersystemen. Im Normalfall passen sich die Menschen an die Anforderungen der Computer an (beispielsweise die Eingabe über Tastatur und Maus). Durch KI können Menschen mit Computersystemen auf eine Art und Weise interagieren, wie sie es auch mit anderen Menschen machen würden, etwa über Gesten, Sprache und Gesichtsausdrücke.

Schließlich können KI-Systeme automatisiert komplizierte Maschine-zu-Maschine-Interaktionen koordinieren und automatisch Anpassungen an diesen durchführen. So können sich auch mehrere KI-Systeme untereinander und miteinander koordinieren.

(Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.13-14)

2.4 Unterschied zwischen KI, Machine Learning und Deep Learning

Wenn man sich mit dem Themengebiet der künstlichen Intelligenz beschäftigt, ist es unvermeidlich, auf die Begriffe Machine Learning und Deep Learning zu stoßen. Obwohl sie häufig synonym verwendet werden, unterscheiden sich diese Begriffe in ihrem Wesen und ihrer Anwendung (Stadler, 2019, S.8).

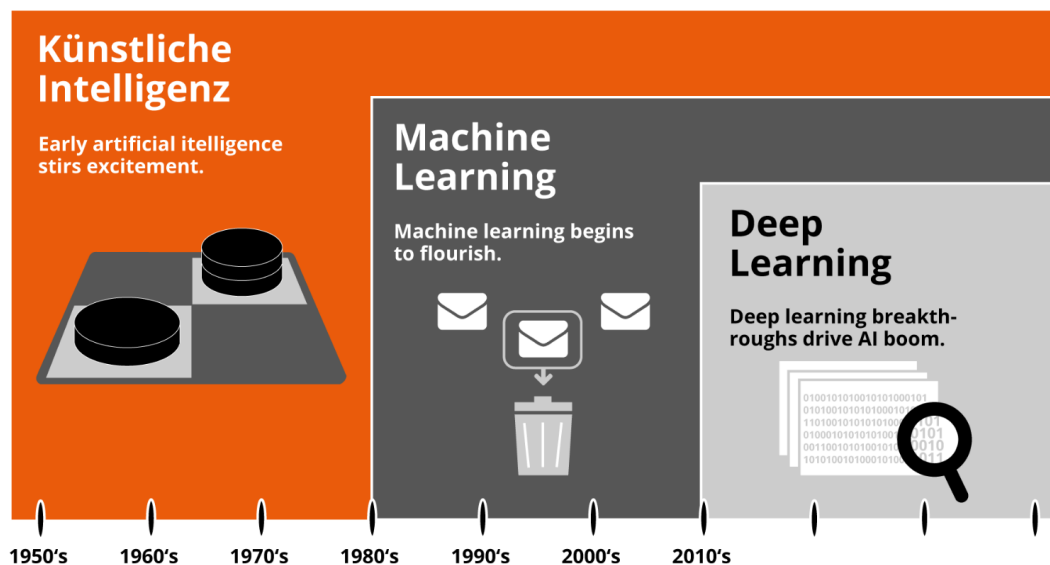


Abbildung 1. KI, Machine Learning und Deep Learning in der Entwicklung (Stadler, 2019, S.8).

Wie aus Abbildung 1 ersichtlich ist, stellt die künstliche Intelligenz (KI) den Überbegriff dar, unter den die Teilgebiete Machine Learning und Deep Learning fallen. KI umfasst dabei den gesamten Bereich des menschlichen Denkens. Machine Learning ist ein Teilgebiet der KI, das sich mit der Entwicklung von Algorithmen und Modellen befasst, die es Computern ermöglichen, aus historischen Daten zu lernen. Deep Learning ist wiederum ein Teilgebiet von Machine Learning, bei dem die Lernmethoden der KI am Aufbau des menschlichen Gehirns orientiert sind (Stadler, 2019, S.8).

K.I

Künstliche Intelligenz kann nicht nur aus Daten lernen, sondern durch geschicktes Programmieren durch Menschen. KI umfasst allerdings nicht nur die Simulation des menschlichen Denkens und Algorithmen, sondern findet sich beispielsweise auch in der Robotik und somit bei der physischen Imitation des Menschen wieder (Stadler, 2019, S.9).

Machine Learning

Machine Learning (ML) ist ein Teilgebiet der Künstlichen Intelligenz (KI), das auf der Idee beruht, dass Maschinen aus historischen Daten lernen und auf dieser Basis Entscheidungen treffen können. Hierbei wird die Funktionsweise des menschlichen Lernens imitiert, indem die Maschine aus Erfahrungen und Daten lernt, um zukünftige Entscheidungen treffen zu können. In der Praxis können verschiedene Methoden zum Einsatz kommen, um ML-Modelle zu entwickeln. Beispielsweise können statistische Modelle wie der Naive-Bayes-Klassifizierer oder Support Vector Machines (SVM) verwendet werden. Die Wahl des optimalen Modells hängt jedoch immer von dem konkreten Anwendungsfall ab, da unterschiedliche Modelle unterschiedliche Stärken und Schwächen aufweisen und unterschiedliche Mengen an historischen Daten benötigen.

Es gibt jedoch keine allgemeingültige Aussage darüber, welches Modell das "beste" ist, da es immer auf den spezifischen Anwendungsfall ankommt. Der Werkzeugkasten an verfügbaren Modellen bietet die Möglichkeit, je nach Bedarf die geeignete Lösung zu wählen und anzupassen. Somit ist Machine Learning eine vielseitige Methode, die es ermöglicht, aus historischen Daten Entscheidungen zu treffen und diese Entscheidungen stetig zu verbessern (Stadler, 2019, S.9).

Deep Learning

Deep Learning ist ein spezifisches Teilgebiet des Machine Learnings, bei dem das Lernverfahren ebenfalls auf historischen Daten basiert. Im Gegensatz zum klassischen Machine Learning baut Deep Learning jedoch auf einem neuronalen Netzwerk auf, das dem menschlichen Gehirn ähnelt. Dieses Netzwerk besteht aus Neuronen, welche als einfache Einheiten fungieren und Eingangssignale empfangen, um daraus ein Ausgangssignal zu generieren. Dabei sind die Eingangssignale gewichtet, sodass einige Kanäle einen stärkeren Einfluss auf das Ausgangssignal haben als andere.

Allerdings ist ein einzelnes Neuron nicht in der Lage, komplexe Probleme zu lösen. Stattdessen besteht das neuronale Netzwerk aus einer großen Anzahl an Neuronen, die in Schichten angeordnet sind. Die erste Schicht des Netzwerks ist mit den Eingangskanälen verbunden, während die letzte Schicht die Ausgangsschicht darstellt. Dazwischen können beliebig viele versteckte Schichten angeordnet sein, die miteinander verbunden sind. Die Ausgangssignale der letzten Schicht bilden die Ausgabe des Modells (Stadler, 2019, S.10). Abbildung 2 veranschaulicht den Aufbau von Neuronen-Schichten im Deep Learning.

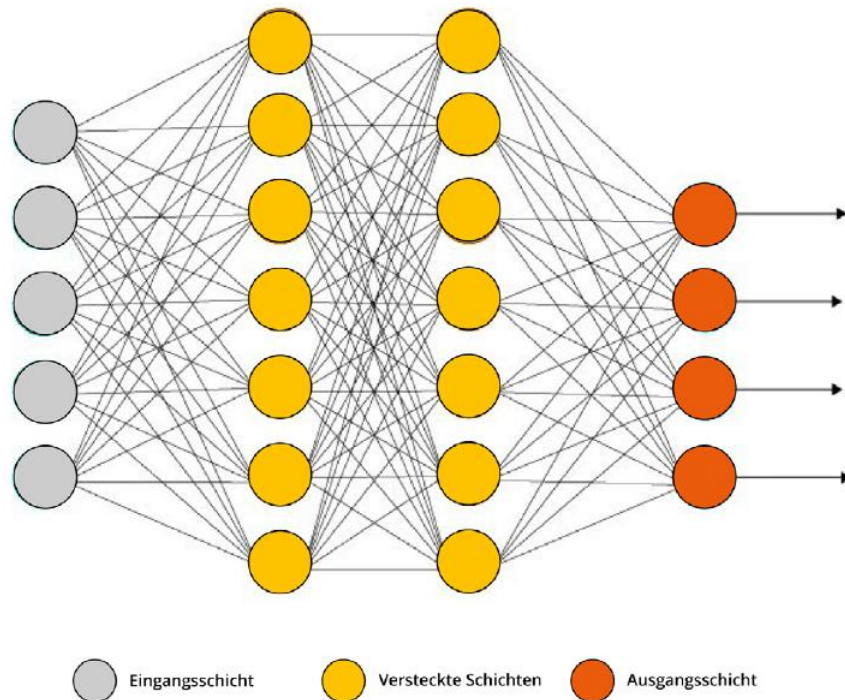


Abbildung 2. Der Aufbau von Neuronen-Schichten im Deep Learning (Stadler, 2019, S.10).

Deep Learning ist aufgrund der Anzahl der Neuronen und Schichten sehr aufwendig. Die Größe des Netzwerks hängt dabei immer vom Umfang des jeweiligen Problems ab. Trotzdem liefert Deep Learning bemerkenswert gute Ergebnisse bei der Lösung komplexer Probleme, wie beispielsweise der Bilderkennung (Stadler, 2019, S.10).

2.5 Maschinelles Lernen

Neben dem oben gegebenen Überblick, soll in diesem Teilkapitel näher auf das Thema Maschinelles Lernen (Machine Learning) eingegangen werden. Bei maschinellem Lernen wird künstliches Wissen aus Erfahrungen generiert, jedoch ist menschliches Handeln im Vorfeld notwendig, damit die Software eigenständig lernen und Lösungen finden kann. Hierfür müssen Menschen die Systeme mit den erforderlichen Daten und Algorithmen versorgen. Vorab müssen die Regeln für die Analyse des Datenbestands und das Erkennen von Mustern definiert werden. Sobald dies erfolgt ist, können die Systeme die folgenden Tätigkeiten ausführen:

- Optimierung von Prozessen aufgrund erkannter Muster
- eigenständige Anpassung an Entwicklungen
- Vorhersagen, basierend auf analysierten Daten, treffen

- Berechnung von Wahrscheinlichkeiten bestimmter Ereignisse
- relevante Daten finden, zusammenfassen und extrahieren

(Stadler, 2021)

Algorithmen spielen beim Machine Learning eine zentrale Rolle, da sie für das Erkennen von Mustern und das Generieren von passenden Lösungen verantwortlich sind. Allerdings ist ein wesentlicher Aspekt des maschinellen Lernens die Wiederholung, da Modelle sich durch kontinuierliches „Füttern“ mit neuen Daten selbstständig anpassen können (Stadler, 2021).

Im Unterschied zu statischen/traditionellen Programmen, passen sich die Entscheidungsregeln bei Algorithmen über eine Rückkopplung an das Erlernte an (Wittpahl, 2019, S.24).

Abbildung 3 veranschaulicht den Unterschied zwischen statischen/traditionellen Programmen und maschinellem Lernen (ML).

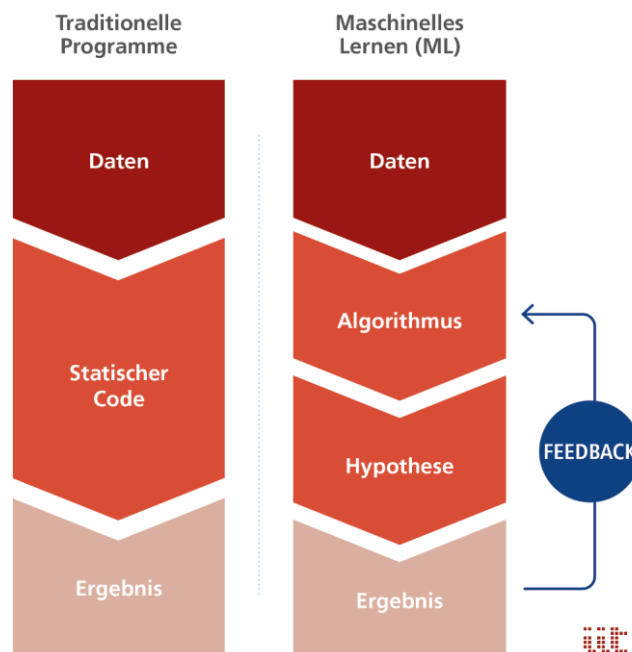


Abbildung 3. Traditionelle Programme versus ML (Wittpahl, 2019, S.25).

Algorithmen lassen sich grundsätzlich in fünf unterschiedliche Kategorien einteilen:

Überwachtes Lernen: Hier werden bereits im Vorfeld Beispielmuster definiert und spezifiziert, um den Algorithmen passende Informationen zuzuordnen. Die Überwachung ist somit nur auf die Trainingsdaten im Voraus bezogen. Das Ziel ist, dem System im Rahmen von aufeinanderfolgenden Rechengängen mit unterschiedlichen Eingaben und Ausgaben anzutrainieren Zusammenhänge herzustellen (Stadler, 2021).

Unüberwachtes Lernen: Beim unüberwachten Lernen werden Modellgruppen anhand eigenständiger Muster erkannt und anschließend automatisiert. Der Algorithmus kann hierbei, im Gegensatz zum überwachten Lernen, nicht trainiert werden. Die Maschine versucht, die eingegebenen Daten nach gewissen Merkmalen zu strukturieren und zu sortieren (Stadler, 2021).

Teilüberwachtes Lernen: Eine Mischung aus dem überwachten und dem unüberwachten Lernen stellt das sogenannte teilüberwachte Lernen dar (Stadler, 2021).

Bestärkendes Lernen: Diese Art des Lernens basiert auf Belohnungen und Bestrafungen. Durch diese Interaktion wird dem Algorithmus mitgeteilt, wie er auf verschiedene Situationen reagieren muss, was der Lernweise der Menschen sehr ähnelt (Stadler, 2021).

Aktives Lernen: Aktives Lernen ermöglicht es einem Algorithmus, gezielte Anfragen für bestimmte Eingangsdaten zu stellen, um die gewünschten Ergebnisse zu erzielen. Um die Anzahl der Anfragen zu minimieren, wählt der Algorithmus selbst relevante Fragen mit hoher Ergebnisrelevanz aus. Die Datenbasis kann je nach System entweder offline oder online vorliegen (Stadler, 2021).

Machine Learning bietet viele Vorteile. Es ermöglicht Menschen einfacher und effizienter zu arbeiten. Es erleichtert und optimiert die Arbeit von Menschen, indem es monotone und aufwendige Aufgaben an den Computer delegiert. Hierzu zählen beispielsweise das Scannen, Speichern und Archivieren von Papierdokumenten sowie das Organisieren und Bearbeiten von Bildern. Machine Learning kann aber auch bei komplexeren Aufgaben wie der Erkennung von Fehlermustern eingesetzt werden. Die Vorteile von Machine Learning sind vielfältig und es gibt zahlreiche Anwendungsmöglichkeiten in verschiedenen Branchen, von denen Unternehmen profitieren können vorliegen (Stadler, 2021).

Des Weiteren wird Machine Learning in einer Vielzahl von Anwendungsbereichen eingesetzt, insbesondere im Internetumfeld. Hierzu zählen die automatische Erkennung von Spam-Mails und die Entwicklung von Spam-Filtern, die Sprach- und Texterkennung für digitale Assistenten und Text-To-Speech-Systeme sowie die Bestimmung der Relevanz von Websites für Suchbegriffe (Stadler, 2021). Neben den bereits genannten Anwendungsbereichen wird Machine Learning in einer Vielzahl anderer Bereiche eingesetzt. So ist es beispielsweise im Bereich der vorausschauenden Wartung von großer Bedeutung, um Wartungsinformationen auf der Basis von großen Datenquellen zu analysieren. Dadurch werden die Ausfallzeit oder Fehler von Maschinen und Anlagen auf ein Minimum reduziert (Stadler, 2021).

Im Supply Chain Management wird Machine Learning eingesetzt, um präzisere Prognosen über Lagerbestände und -auffüllungen treffen zu können. Wodurch die Möglichkeit besteht, fehlerhafte Bedarfsschätzungen zwischen 20 und 50 % zu reduzieren. Auch im Marketing kann von maschinellem Lernen profitiert werden. Durch Machine Learning können beispielsweise Kaufprognosen für einzelne Kunden oder Kundensegmente erstellt werden. Wenn Kunden beispielsweise vermehrt Sportkleidung zu Jahresbeginn kaufen, können passende Angebote genau zu diesem Zeitpunkt unterbreitet werden. Auf diese Weise kann das Marketing gezielter und effektiver auf die individuellen Bedürfnisse und Vorlieben der Kunden eingehen (Stadler, 2021).

Auch im Bereich des Energiemanagements wird Machine Learning eingesetzt. Der Klimawandel und die Energiewende stellen gegenwärtig eine der größten Herausforderungen für die Gesellschaft, Politik und Industrie dar. Data-Science-Methoden wie maschinelles Lernen bieten die Möglichkeit, den zunehmend komplexen Energiemarkt zu beherrschen. Sie vereinfachen die Abstimmung von Angebot und Nachfrage, indem sie einen voraussichtlichen Bedarf ermitteln. Eine intelligente Steuerung in Echtzeit ermöglicht eine preisoptimierte Strategie für Stromerzeuger (Stadler, 2021).

Die genannten Beispiele machen deutlich, welchen Einfluss Machine Learning in der heutigen Zeit bereits hat. Möglich wurde dies durch die Entwicklung der Programmiersprache Python. Python wurde 1991 entwickelt und ist heute die am häufigsten verwendete Programmiersprache für die Entwicklung von Machine-Learning-Programmen. Seit ihrer Entstehung zeichnet sich Python durch einen übersichtlichen und einfachen Quellcode aus, der sowohl funktionale als auch objektorientierte Programmierung ermöglicht. Python ist eine freie Programmiersprache, was bedeutet, dass es viele verschiedene frei nutzbare Pakete gibt, die von einer großen Entwicklergemeinschaft entwickelt und unterstützt werden (Stadler, 2021).

2.6 Starke und Schwache KI

Grundsätzlich werden im Bereich der KI-Forschung zwei verschiedene Arten unterschieden: Die schwache und die starke Künstliche Intelligenz (Stadler,2022, S. 2).

2.6.1 Schwache KI (weak/narrow AI):

Diese erste Form der künstlichen Intelligenz erreicht bereits heute in begrenzten Teilbereichen eine vergleichbare Leistungsfähigkeit wie die menschliche Intelligenz, jedoch sind die Fähigkeiten dieser intelligenten Systeme auf spezifische Anwendungsbereiche beschränkt. Ein Beispiel hierfür ist die Fähigkeit, perfekte Texte zu verfassen, aber nicht in der Lage zu sein, mit Menschen zu kommunizieren oder andere Aufgaben zu bewältigen. Diese Schwache KI bedient sich mathematischer Methoden und Informatik, die speziell auf die jeweiligen Anforderungen hin entwickelt wurden.

Schwache KI-Systeme sind nicht in der Lage, ein tieferes Verständnis für die zugewiesenen Problemlösungen zu erlangen und bleiben daher auf einem oberflächlichen Intelligenz-Level stehen. Im Alltag begegnet man schwachen KI-Systemen oft in Form von Zeichen- oder Texterkennungsprogrammen, Navigationssystemen, Spracherkennung und der individuellen Anzeige von Werbung (Stadler,2022, S. 3).

2.6.2 Starke künstliche Intelligenz (strong/general AI):

Superintelligenz oder Strong AI bzw. General AI ist eine Art der künstlichen Intelligenz, die menschliche intellektuelle Fähigkeiten erreichen und sogar übertreffen soll. Diese Systeme sollen in der Lage sein, eigenständig, intelligent und flexibel zu handeln und nicht nur auf die Lösung eines konkreten Problems beschränkt zu sein.

Bisher ist es jedoch nicht gelungen, eine starke künstliche Intelligenz zu entwickeln und die Diskussion darüber, ob dies überhaupt möglich ist, hält weiter an. Wenn es in Zukunft tatsächlich gelingen sollte, eine starke KI zu entwickeln, würde sie folgende Eigenschaften aufweisen:

- Logisches Denkvermögen, um komplexe Zusammenhänge zu erkennen und zu verstehen.
- Entscheidungsfähigkeit trotz möglicher Unsicherheit, um auch in unvorhersehbaren Situationen handlungsfähig zu bleiben.
- Planungs- und Lernfähigkeit, um Erfahrungen aus der Vergangenheit zu nutzen und zukünftige Herausforderungen zu meistern.

- Kommunikationsfähigkeit in natürlicher Sprache, um auf komplexe Fragen und Anforderungen von Menschen reagieren und entsprechend handeln zu können.
- Fähigkeit, alle Fähigkeiten zu kombinieren, um ein übergeordnetes Ziel zu erreichen.

(Stadler,2022, S. 3)

2.6.3 4 KI-Typen

Neben der Einteilung in eine schwache und eine starke KI lässt sich Künstliche Intelligenz weiter in vier verschiedene Typen unterteilen.

Typ 1: reaktive Maschinen (Reactive Machines):

Der Typ 1 der künstlichen Intelligenz wird als der ursprüngliche Typ der schwachen KI betrachtet. Reaktive Maschinen sind in der Lage, eine einzige Aufgabe zu erfüllen, für die sie speziell programmiert wurden. Ein bekanntes Beispiel für eine solche Maschine ist der IBM Schachcomputer DeepBlue, der im Jahr 1997 gegen den amtierenden Schachweltmeister antrat und gewann.

DeepBlue konnte alle möglichen Züge abwägen und denjenigen auswählen, der am schnellsten zu einem Schachmatt führte, was ihm zum Sieg verhalf. Diese Maschine war jedoch nur in der Lage, diese eine spezifische Aufgabe, nämlich das Spielen von Schach, zu erfüllen und war in sämtlichen anderen Lebensbereichen nutzlos (Stadler,2022, S. 4).

Typ 2: begrenzte Speicherkapazität (Limited Memory):

Im Unterschied zu Reactive Machines sind Limited Memory KI in der Lage, auf Basis von gesammelten Erfahrungen aus vergangenen Ereignissen Entscheidungen zu treffen und diese in aktuellen Situationen zu berücksichtigen. Dabei sind sie in der Lage, Daten zu sammeln, zu speichern und diese für zukünftige Entscheidungen abzurufen. Ein bekanntes Beispiel für eine KI des Typs 2 sind selbstfahrende Autos, die durch die Analyse von vergangenen Situationen und die Berücksichtigung von Verkehrsregeln und Verhaltensweisen anderer Verkehrsteilnehmer in der Lage sind, sich im Straßenverkehr zurechtzufinden. Sie sind in der Lage, aus Erfahrungen zu lernen und ihre Entscheidungsprozesse zu verbessern. KI des Typs 2 sind heutzutage die am weitesten verbreitete Form von KI und finden sich in zahlreichen Anwendungen wie persönlichen Smartphone-Assistenten, Suchmaschinen und sozialen Medien wieder (Stadler,2022, S. 4).

Typ 3: Theorie des Geistes (Theory of Mind):

Theory of Mind KI, auch als Typ 3 KI bezeichnet, ist eine Art von künstlicher Intelligenz, die bisher nur in der Theorie existiert und zur starken künstlichen Intelligenz gezählt wird. Künstliche Intelligente Maschinen (KIM), die diese Art von KI erreichen, sollen in der Lage sein, menschliche Emotionen wahrzunehmen, zu verstehen und ihr Verhalten entsprechend anzupassen. Sie sollen ein Gedächtnis haben und ihr Bild von der Welt basierend auf Gelerntem erweitern können.

Im Gegensatz zu den vorherigen Formen der KI stellt die Entwicklung von Theory of Mind KI eine große Herausforderung für die Wissenschaft dar. Die Gründe hierfür sind, dass Emotionen und zwischenmenschliche Interaktionen höchst komplex sind und technisch schwer nachzubilden sind. Die Entwicklung von KIM, die in der Lage sind, menschliche Emotionen und Verhaltensweisen zu erkennen und zu berücksichtigen, erfordert daher ein tiefes Verständnis von menschlicher Psychologie und sozialer Interaktion sowie eine entsprechend fortschrittliche Technologie (Stadler, 2022, S. 4-5).

Typ 4: Selbstwahrnehmung (Self Awareness):

Der Typ 4 der Künstlichen Intelligenz, auch Self Awareness KI genannt, ist die höchste Stufe der künstlichen Intelligenz und kommt dem menschlichen Bewusstsein am nächsten. Eine solche KI wird in der Lage sein, die Welt vollständig wahrzunehmen und menschliche Emotionen, Absichten und Reaktionen nachvollziehen und darauf basierend handeln zu können. Sie wird sich ihrer eigenen Existenz bewusst sein und ein Selbstbild haben, das sich aus ihren Erfahrungen und ihrem Gedächtnis speist. Damit geht sie den entscheidenden Schritt von "Ich denke" zu "Ich weiß, dass ich denke". Sie wird an die menschliche Intelligenz heranreichen und diese möglicherweise sogar übersteigen (Stadler, 2022, S. 5).

2.6.4 Turing Test

Die Ursprünge der Künstlichen Intelligenz reichen etwa 90 Jahre zurück, als der britische Mathematiker und Informatiker Alan Turing darüber nachdachte, was eine intelligente Maschine ausmacht und wie man ihre Intelligenz messen könnte. Im Jahr 1950 formulierte Turing eine Idee, wie man die Intelligenz einer Maschine durch einen Test feststellen könnte, den sogenannten Turing-Test. Dieser Test wird durchgeführt, indem ein menschlicher Fragesteller über eine Tastatur und einen Bildschirm eine Unterhaltung mit zwei unbekannten Gesprächspartnern führt, von denen einer ein Mensch und der andere eine Maschine ist. Ohne Sicht- oder Hörkontakt muss der Fragesteller anhand des Gesprächs entscheiden, welcher Gesprächspartner die Maschine ist. Wenn der Fragesteller nach intensiver Befragung nicht in der Lage ist, den Gesprächspartner als Maschine zu

identifizieren, hat die Maschine den Turing-Test bestanden und es wird ihr menschliche Intelligenz attestiert (Stadler,2022, S. 5).

2.7 Anwendungsmöglichkeiten von K.I

Durch den Einsatz von Algorithmen können Daten genau analysiert und komplexe Aufgaben leicht gelöst werden. Künstliche Intelligenz ist sogar der menschlichen Wahrnehmung bei der Bild- und Texterkennung überlegen. Eine semantische Textanalyse ermöglicht die Interpretation sowohl analytischer als auch intuitiver Absichten des Verfassers. Deep Learning, auch bekannt als maschinelles Lernen, ermöglicht es, maschinelle Lernprozesse in Gang zu setzen, die exponentiell Informationen aufnehmen können. Die Künstliche Intelligenz entwickelt sich ständig weiter und öffnet damit neue Dimensionen. Wenn sie ethisch ausgerichtet und zum Wohl der Menschheit eingesetzt wird, könnte sie dazu beitragen, menschliche Tätigkeiten und Arbeiten zu optimieren und zu rationalisieren. Allerdings könnte der Einsatz der Künstlichen Intelligenz auch dazu führen, dass der Abstand zwischen Arm und Reich zunimmt. Daher sollte die Implementierung menschengewollter Zielsetzungen ein Hauptziel bei der Entwicklung der Künstlichen Intelligenz sein, um sicherzustellen, dass ihr Einsatz im Einklang mit dem Wohl der Menschheit steht. Ein zentraler Unterschied zwischen Künstlicher Intelligenz und menschlicher Intelligenz ist die Bewusstheit des Menschen, die der Künstlichen Intelligenz fehlt. Allerdings könnten mithilfe der Künstlichen Intelligenz auch neue Sinneserfahrungen entstehen, die über die menschliche Sinneswahrnehmung hinausgehen. Dies wirft komplexe ethische und philosophische Fragen auf. Die Künstliche Intelligenz könnte einerseits das Erkenntnisvermögen der Menschen erweitern und ganz neue Sinneswahrnehmungen ermöglichen. Andererseits besteht die Gefahr, dass die Künstliche Intelligenz den Menschen dominiert und in seinen Lebensbereichen einschränkt (Kitzmann,2022, S.116-117).

Die Tatsache, dass Künstliche Intelligenz (KI) mittlerweile kein bloßes Hype-Thema mehr ist, zeigt sich unter anderem auch durch die zahlreichen verfügbaren Anwendungen, die auf KI-Technologien basieren oder diese nutzen. In den letzten 25 Jahren wurde die Leistungsfähigkeit von KI immer wieder durch Siege über menschliche Spieler der breiten Öffentlichkeit verdeutlicht. So schlug beispielsweise im Jahr 1996 der IBM Computer "Deep Blue" den amtierenden Schachweltmeister Garri Kasparov. Im Jahr 2015 schlug die von Google entwickelte KI "Alpha Go" den Europameister Fan Hui im Brettspiel "Go". Im Jahr 2016 wurde auch der Weltmeister Lee Seedol von "Alpha Go" geschlagen. Im Jahr 2017 präsentierte Deep Mind die neueste Version namens "Alpha Go Zero". Während "Alpha Go" noch mit Expertenwissen trainiert werden musste, wurden "Alpha Go Zero" lediglich die Spielregeln zur Verfügung gestellt. Nach einem nur

dreitägigen Reinforcement Learning, bei dem das System gegen sich selbst spielte, wurde es zum besten Spieler aller Zeiten. Während "Alpha Go" noch 30 Millionen Trainingsspiele benötigte, waren für "Alpha Go Zero" nur noch 3,9 Millionen Spiele notwendig (Haarmeier,2021, S.27).

Im Jahr 2018 ließen Forscher der Universität Freiburg das Atarispiel "Q*bert" von KI-Algorithmen spielen, wobei auch hier Reinforcement Learning zum Einsatz kam. Das KI-System fand einen Fehler im Design des Spiels, nutzte ihn aus und erzielte bisher unerreichte Höchststände bei den Punkten. Ebenfalls im Jahr 2018 schlug das System "OpenAI Five" der Organisation OpenAI ein menschliches Team in dem Multiplayer-Online-Battle "Dota 2". Hierbei arbeiteten fünf Algorithmen in Teamwork zusammen (Haarmeier,2021, S.27).

In verschiedenen Bereichen haben Fortschritte in der Künstlichen Intelligenz (KI) zu bedeutenden Erfolgen geführt. Ein Beispiel hierfür ist die Anwendung der KI bei der Diagnose von Hauterkrankungen. Im Jahr 2019 haben 157 Hautärzte aus Universitätskliniken gemeinsam mit einem trainierten Algorithmus 100 Bilder untersucht, um zu bestimmen, ob es sich um ein Muttermal oder schwarzen Hautkrebs handelt. Die Bewertung des Algorithmus war in diesem Fall präziser als die der Ärzte (Haarmeier,2021, S.28).

Während in den 2010er Jahren oft von Erfolgen in der KI berichtet wurde, bei denen die Technologie besser abschnitt als der Mensch, sind KI-Technologien mittlerweile auch im Massenmarkt verbreitet. Die großen IT-Plattform-Unternehmen sind hierbei die treibende Kraft. So war beispielsweise Siri als KI-Sprachassistent bereits im Jahr 2011 im iPhone verfügbar. 2014 wurde Microsoft Cortana eingeführt, das ab dem Rollout von Windows 10 im Jahr 2015 für alle Nutzer zugänglich war. 2015 folgte der Amazon Smart-Speaker Alexa, 2016 der Google Sprachassistent und 2017 der Samsung Sprachassistent Bixby (Haarmeier,2021, S.28).

2.7.1 K.I Anwendungen im Alltag

Eine Umfrage ergab, dass 85% der Endnutzer in Deutschland mindestens ein Gerät besitzen, auf dem ein Sprachassistent vorinstalliert ist. Die Geschwindigkeit, mit der die Verbraucher diese Technologie annehmen und regelmäßig nutzen, ist vergleichbar mit der Einführung der ersten Smartphones. Dies wird durch die kurzen Lebenszyklen der eingesetzten Hardware und die Tatsache unterstützt, dass die Lösungen in die Betriebssysteme integriert sind und mit einem neuen Software-Release allen Anwendern standardmäßig zur Verfügung gestellt werden. Die schnelle Bereitstellung von Innovationen und neuen Funktionen im Softwarebereich wird durch die Nutzung von Cloud-Architekturen ermöglicht, die ein einfaches oder kontinuierliches Update der Lösungen erlauben. Ein Beispiel dafür sind KI-Technologien, die zunehmend in unseren täglichen Arbeitsabläufen

eingesetzt werden. Microsoft zeigt dies anhand verschiedener Anwendungen wie Cortana in Outlook, welche E-Mails vorliest und wichtige Details zusammenfasst sowie eine Reihe von Terminplanungsaufgaben übernehmen kann. In Excel können durch OCR fotografierte Tabellen erkannt und die Daten in Tabellenform in die Excel-Zellen eingelesen werden. PowerPoint zeigt Live-Untertitel in Echtzeit an und unterstützt zwölf gesprochene Sprachen sowie Untertitel in über 60 Sprachen auf dem Bildschirm. Der PowerPoint Presenter Coach gibt sofortiges Feedback zu häufigen Fehlern wie schlechtem Tempo und der Verwendung von Füllwörtern. MyAnalytics nutzt KI, um Arbeitsmuster zu analysieren und Vorschläge zu unterbreiten, wie die Arbeit anders organisiert werden kann, um Unterbrechungen zu minimieren und die Konzentration zu erhöhen. Microsoft Pix verbessert Fotos mit KI, um sie schärfer und klarer zu machen. Teams nutzt KI, um den Hintergrund unscharf werden zu lassen und andere Hintergründe oder den Präsentator virtuell vor den Folien einzublenden. Mit dem AI Builder können benutzerdefinierte KI-Modelle erstellt oder vorgefertigte Modelle ausgewählt werden, um Geschäftsprozesse zu optimieren (Haarmeier, 2021, S.28-29).

2.7.2 Künstliche Intelligenz in der Forschung

Die Bedeutung von Künstlicher Intelligenz (KI) beschränkt sich nicht nur auf den Alltag, sondern hat auch in Forschung und Wissenschaft eine bedeutende Rolle eingenommen. KI-Anwendungen werden in verschiedenen wissenschaftlichen Bereichen wie Physik, Klimaforschung und Medizin erfolgreich eingesetzt und gelten als mächtiges Werkzeug, das die Wissensarbeit verändert (Gethmann, Buxmann, Distelrath, Humm, Lingner, Nitsch, Schmidt, Spiecker, 2022, S.39).

Die Automatisierung von Routine-Tätigkeiten durch KI kann zu einer höheren Effizienz in der Forschung führen, indem (Forschungs-)Arbeiten beschleunigt werden. Es ist jedoch zu beachten, dass sich dadurch auch Tätigkeitsbereiche und Stellenprofile verändern können. Beispiele hierfür sind KI-gestützte Laborauswertungen in den Lebenswissenschaften. Die Erhöhung des Automatisierungsgrades kann dazu führen, dass Forschende einen weniger tiefen Einblick in die zu untersuchende Materie bekommen, was dazu führen kann, dass sie Ergebnissen von KI-Anwendungen vertrauen, ohne sie nachvollziehen und kritisch hinterfragen zu können (Gethmann, Buxmann, Distelrath, Humm, Lingner, Nitsch, Schmidt, Spiecker, 2022, S.39).

KI ermöglicht eine größere Genauigkeit in der Beantwortung von Forschungsfragen, insbesondere durch die Verfügbarkeit großer Datenmengen. Beispiele hierfür sind KI-gestützte Auswertungen von Versuchen in der Physik. Es ist jedoch wichtig zu beachten, dass falsch eingesetzte KI-Methoden zu falschen Interpretationen von Ergebnissen führen können, und es keine Garantie gegen

Falschaussagen gibt Gethmann, Buxmann, Distelrath, Humm, Lingner, Nitsch, Schmidt, Spiecker, 2022, S.40).

KI und verwandte Technologien eröffnen in vielen Wissenschaftsdisziplinen neue Möglichkeiten, die bislang unmöglich waren, wie zum Beispiel die Genomanalyse und darauf aufbauende Forschungen. Besonders die Hypothesenbildung aufgrund von KI-Techniken eröffnet in vielen Wissenschaftsdisziplinen ganz neue Möglichkeiten (Gethmann, Buxmann, Distelrath, Humm, Lingner, Nitsch, Schmidt, Spiecker, 2022, S.40).

2.7.3 Weitere Anwendungsmöglichkeiten von K.I

Neben den bereits erwähnten Anwendungsmöglichkeiten, sollen folgend weitere erläutert werden.

Maschinelle Übersetzung

Die maschinelle Übersetzung, auch als MÜ oder MT (engl. Machine translation) bezeichnet, wurde erstmals in den 1960er-Jahren entwickelt. Das US-Militär war Vorreiter auf diesem Gebiet und entwickelte ein rudimentäres Russisch-Englisch-Übersetzungsprogramm, dessen Qualität sich jedoch als mangelhaft erwies. Aufgrund dieser mangelhaften Leistung stufte das US-Verteidigungsministerium bald darauf die automatische Übersetzung als unrealisierbar ein, was zu einem Stillstand in der Forschung bis in die 1980er-Jahre führte (Negovec, 2022).

Japan erlangte schließlich den Durchbruch auf dem Gebiet der maschinellen Übersetzung, indem es die weltweit ersten kommerziellen Übersetzer für PCs entwickelte und somit eine Vorreiterfunktion einnahm. Seitdem hat sich die Technologie der Machine Translation stark weiterentwickelt und ist seit 2006 dank Google auch privaten Nutzern zugänglich. Anfänglich hatte der Google-Übersetzer jedoch aufgrund seiner geringen Qualität einen schlechten Ruf. Doch seit 2016 kommen KI-Übersetzer zum Einsatz, die für enorme Qualitätssprünge gesorgt haben. Trotzdem gibt es nach wie vor Schwierigkeiten bei der Übersetzung von feinen Sprachnuancen, die mithilfe künstlicher Intelligenz (fast) gänzlich überwunden werden könnten. Die Forschung auf dem Gebiet der maschinellen Übersetzung wird jedoch stets vorangetrieben, da man sich davon eine effizientere globale Kommunikation in Wirtschaft, Diplomatie und Militär verspricht (Negovec, 2022).

Gesichtserkennung

Die Gesichtserkennung hat sich in den letzten Jahren als eine leistungsfähige Technologie zur Identifikation von Personen etabliert. Durch den Einsatz von 3D-Verfahren und künstlicher Intelligenz ist es möglich, spezifische Merkmale eines Gesichts zu erfassen und mit dem gespeicherten Gesicht des Besitzers

abzugleichen. Die Gesichtserkennung wird mittlerweile nicht nur in der Unterhaltungselektronik eingesetzt, sondern auch in sicherheitsrelevanten Bereichen wie der Sicherheitstechnik, Kriminalistik und Forensik, um Personen eindeutig identifizieren und verifizieren zu können (Stadler, 2022, S.8).

Online-Handel

Auch im Online-Handel hat der Einsatz von Künstlicher Intelligenz eine hohe Bedeutung erlangt. Vor allem Amazon setzt in großem Umfang auf KI-Technologien, um den Handel zu optimieren. Von personalisierter Werbung bis hin zur dynamischen Preisgestaltung sollen zukünftig auch Bestellvorgänge vollständig automatisiert abgewickelt werden. Hierbei kommen Algorithmen zum Einsatz, die das Kaufverhalten der Kunden analysieren und Bestellungen automatisch verarbeiten. Zusätzlich bietet Amazon mit Alexa eine intelligente Einkaufshilfe an, die den Kunden bei der Produktauswahl und Bestellung unterstützt. Zahlreiche Unternehmen profitieren bereits von der Optimierung ihrer Prozesse durch den Einsatz von KI-Technologien im Online-Handel. Die Automatisierung und Effizienzsteigerung durch KI eröffneten neue Möglichkeiten für Unternehmen, um im Wettbewerb erfolgreich zu sein (Stadler, 2022, S.8).

Medizin und Pflege

Die Anwendung von künstlicher Intelligenz im Gesundheitswesen wird angesichts des Fachkräftemangels als vielversprechende Möglichkeit zur Effizienzsteigerung und Kosteneinsparung betrachtet. In diesem Zusammenhang haben insbesondere Machine-Learning-Algorithmen, wie z.B. Deep-Learning, in den letzten Jahren große Fortschritte in der automatischen Diagnose von Krankheiten erzielt. Sie können große Datenmengen analysieren und Muster erkennen, die von menschlichen Ärzten möglicherweise übersehen werden. Auf diese Weise können potenziell lebensbedrohliche Krankheiten früher erkannt und behandelt werden. Darüber hinaus werden KI-Anwendungen auch in anderen Bereichen des Gesundheitswesens eingesetzt. Beispielsweise können analytische Vorgänge in den frühen Phasen der Arzneimittelentwicklung verkürzt werden, indem KI-Systeme genutzt werden, um mögliche Nebenwirkungen oder Interaktionen zwischen Wirkstoffen vorherzusagen. Zudem können KI-Systeme auch zur patientenabhängigen Personalisierung von Therapien beitragen, indem sie große Datenmengen analysieren und Muster erkennen, die für eine individuelle Diagnose und Behandlung genutzt werden können. Insgesamt tragen KI-Anwendungen im Gesundheitswesen dazu bei, jahrelange Arbeit und hohe Investitionskosten einzudämmen und damit den Zugang zu qualitativ hochwertiger Gesundheitsversorgung zu verbessern (Stadler, 2022, S.9).

Autonomes Fahren

Autonome Fahrzeuge stellen eine vielversprechende Technologie dar, die das Potenzial hat, die Sicherheit auf den Straßen deutlich zu verbessern. Innerhalb der nächsten 15 Jahre könnte es möglich werden, dass intelligente, autonom fahrende Autos auf Straßen unterwegs sind. Die Entwicklung autonomer Fahrzeuge erfolgt in verschiedenen Stufen, die als Level bezeichnet werden. Dabei gibt der Mensch schrittweise seine Verantwortung am Steuer an das zunehmend intelligente Fahrzeug ab. Die Entwicklung beginnt bei den bekannten Fahrassistenzsystemen wie Einparkhilfen oder dem Tempomat und geht bis zur vollständigen Autonomie des Fahrzeugs, bei der der Fahrer vollständig zum Passagier wird (Stadler, 2022, S.9).

2.8 Einfluss von KI auf die Arbeitswelt

Neben den genannten Anwendungsmöglichkeiten werden die zunehmende Automatisierung und Digitalisierung zweifellos erhebliche Veränderungen in der Arbeitswelt mit sich bringen. Viele traditionelle Berufe werden verschwinden, während gleichzeitig neue Berufsfelder entstehen werden. Die Automatisierung wird vor allem Routineaufgaben ersetzen, die durch Künstliche Intelligenz ausgeführt werden können. Die Art der Arbeit selbst wird sich ebenfalls verändern, da die Arbeitnehmer zunehmend umfassendere Fähigkeiten benötigen, um in verschiedenen Bereichen tätig zu sein. Ein Beispiel hierfür wäre ein Techniker, der auch als Informatiker, Kommunikationsexperte und Handwerker agiert. Es ist daher wahrscheinlich, dass die Berufe, die sich aufgrund der überschneidenden Tätigkeitsbereiche entwickeln, zunehmen werden (Kitzmann, 2022, S.66).

Die Auswirkungen der Automatisierung und Digitalisierung beschränken sich jedoch nicht nur auf technische und digitale Fähigkeiten, sondern betreffen auch soziale und emotionale Kompetenzen, die in Zukunft noch wichtiger werden. Da die menschliche Kommunikation nach wie vor die Grundlage jeder beruflichen Tätigkeit bildet, sind Fähigkeiten wie Empathie und Kreativität von entscheidender Bedeutung. Die Arbeitswelt wird zunehmend mobiler und virtueller, was bedeutet, dass der direkte menschliche Kontakt abnehmen wird und die Kommunikation zunehmend digital erfolgen wird. Infolgedessen werden die Fähigkeiten des menschlichen Einfühlungsvermögens und der Kreativität in einer neuen Arbeitswelt immer wichtiger (Kitzmann, 2022, S.66-67).

Trotzdem bleibt der direkte zwischenmenschliche Kontakt unverzichtbar für die menschliche Bedürfnisbefriedigung. Die zunehmende Versachlichung der Arbeit kann zu einem Mangel an menschlicher Interaktion und sozialer Bestätigung führen. Es ist denkbar, dass in der Zukunft Berufe entstehen könnten, die

ausschließlich auf die Befriedigung sozialer Bedürfnisse ausgerichtet sind. Es wird wichtig sein, die grundlegenden menschlichen Bedürfnisse in der Arbeitswelt nicht aus den Augen zu verlieren, um das Wohlbefinden der Arbeitnehmer zu gewährleisten. Eine Möglichkeit, diesem Mangel entgegenzuwirken, besteht in der Schaffung von Berufsbildern, die den persönlichen zwischenmenschlichen Kontakt fördern und eine direkte menschliche Interaktion ermöglichen (Kitzmann, 2022, S.67-68).

2.9 Rechtliche Aspekte beim Einsatz von KI

Die Nutzung von KI-Systemen wirft zahlreiche rechtliche Fragestellungen auf, insbesondere im Hinblick auf den urheberrechtlichen Schutz von generierten Produkten sowie die Möglichkeit der Patentierung von KI-Systemen als Erfindungen.

2.9.1 Urheberrechtsschutz

KI-Technologien haben mittlerweile die Fähigkeit Bilder zu malen, Musik zu komponieren und vorzutragen. Sie können Computerprogramme schreiben und Designs gestalten. Solche Outputs könnten durch Urheber- oder Leistungsschutzrechte geschützt sein. Erst durch einen solchen Rechtsschutz werden sie verkehrsfähig (können lizenziert und verkauft werden) und sind gegen die ungefragte Übernahme und Weiterverwendung geschützt. Sind keine Immaterialgüterrechte (Oberbegriff für Urheber-, Patent- und Markenrechte; englisch Intellectual Property Rights, kurz: IPR), gegeben, sind die Inhalte gemeinfrei und können völlig legal einfach übernommen und weiterverwendet werden. Ein Rechtsschutz ist dann nur sehr eingeschränkt über Verträge möglich. Die Frage, ob an KI-generierten Schöpfungen Urheber- oder Leistungsschutzrechte entstehen können, wird derzeit viel diskutiert. Hierbei stellt sich generell die Frage der Rechtszuordnung. Wenn solche Rechte entstehen, wem würden sie zustehen? Der entwickelnden Person des automatisierten Systems oder Algorithmus, dessen Arbeitgeber*in oder der KI selbst? (Kreutzer, Christiansen, 2021, S.54-55)

Es gibt eine grundlegende Schwierigkeit bei der Zuordnung von Rechten an KIs. Maschinen oder Algorithmen haben keine Rechte und können keine Rechte besitzen, da sie nicht rechtsfähig sind. KIs können keine Verträge im eigenen Namen abschließen, können nicht verklagt werden und auch keine Straftaten begehen. Diese rechtliche Einschränkung bleibt solange bestehen, bis es Künstliche Intelligenzen gibt, die wirklich intelligente und eigenständige Persönlichkeiten darstellen. Auch eine Zuordnung von Immaterialgüterrechten (v. a. Urheber- und Leistungsschutzrechten) an die Entwickler:innen der KI, deren

Arbeitgeber:innen oder die Eigentümer:innen des Systems wirft erhebliche Schwierigkeiten mit der geltenden Rechtsordnung auf. Solche Rechte werden in aller Regel denjenigen zugeordnet, die das Werk geschaffen (beim Urheberrecht) oder die geschützte Leistung erbracht haben (bei Leistungsschutzrechten). Programmierer:innen eines Selbstlernprogramms sind jedoch nicht Schöpfer:innen der Ergebnisse, die dieses erzeugen (ebenso wenig wie beispielsweise die Entwickler:innen von Microsoft Schöpfer:innen von in MS Word geschriebenen Texten sind). Ihnen Urheber- oder Leistungsschutzrechte am Output zuzuordnen ist daher schwer zu begründen (Kreutzer, Christiansen, 2021, S.54-55).

Bereits heute malen KI-Technologien Bilder, schreiben Bücher und komponieren Lieder und dies mehr oder weniger vollständig autonom. Allerdings ist es fraglich, ob es sich hierbei um „Schöpfungen“ handelt, die urheberrechtlich geschützt sein können. Für die urheberrechtliche Schutzfähigkeit von KI-erzeugten Daten und Inhalten ist es entscheidend, ob und inwiefern sie auf die gestalterische Tätigkeit einer menschlichen Person zurückzuführen sind. Das Urheberrecht ist ein personenbezogenes Schutzrecht. Es gilt nur für persönliche geistige Schöpfungen eines Menschen. Weder fallen rein technische Gestaltungen noch solche von Tieren unter das Urheberrecht. Wenn eine KI eine Grafik, einen Text oder ein Musikstück ohne menschliche Intervention erstellt, gibt es keine Urheberrechte auf das Werk. Die KI-Werke sind nicht urheberrechtsfähig. Wenn jedoch ein Mensch den Output der KI bearbeitet, kann ein Urheberrecht auf die bearbeitete Fassung bestehen, sofern die menschliche Bearbeitung kreativ ist und ein Mindestmaß an Originalität aufweist. Je autonomer die KI bei der Schöpfung arbeitet, desto unwahrscheinlicher ist es, dass der Output urheberrechtlich geschützt ist (Kreutzer, Christiansen, 2021, S.56).

2.9.2 Leistungsschutz

Die Frage, ob der Output von KI-Technologien durch Leistungsschutzrechte geschützt werden kann, hängt von der Art des Arbeitsergebnisses ab. Leistungsschutzrechte können (zumindest teilweise) auch auf nicht menschliche Erzeugnisse anwendbar sein. Die Voraussetzungen für die verschiedenen Schutzrechte sind jedoch sehr unterschiedlich.

Fotos:

Technisch gesehen ist es möglich, dass maschinelles Lernen verwendet wird, um automatisch Fotografien, Bildkollagen oder andere Arten von Abbildungen zu generieren. Künstliche Intelligenzen können beispielsweise durch Trainingsdaten aus Bildern in die Lage versetzt werden, eigene Bilder zu erzeugen, vorausgesetzt, dass die erforderlichen Inhalte für die Neukomposition vorhanden sind. Solche nicht schöpferischen Abbildungen, die keine menschlichen, eigenschöpferischen

Züge aufweisen, könnten möglicherweise nur durch das Lichtbildrecht geschützt werden. Obwohl das Lichtbildrecht keine menschliche Kreativität in der Abbildung voraussetzt, ist es bei reinen KI-Erzeugnissen nicht anwendbar. Es herrscht allgemeine Zustimmung in der Rechtsprechung und Literatur, dass das Lichtbildrecht nur für menschliche Leistungen gewährt wird. Daher sind Fotos und andere Abbildungen, die von Tieren oder Maschinen erzeugt werden, nicht durch das Lichtbildrecht geschützt. Dies gilt auch für Videos oder Filme. Selbst wenn es sich um einfache, kurze Videos ohne schöpferische Leistungen handelt, können sie ähnlich wie einfache Fotos durch das Laufbildrecht (§ 95 UrhG) geschützt sein. Dies gilt jedoch nicht für rein technische Erzeugnisse (Kreutzer, Christiansen, 2021, S.56).

Musik- und Filmproduktionen:

Musikproduktionen werden herkömmlich durch das Tonträgerherstellerecht geschützt. Anders als das Urheberrecht an der Komposition bezieht sich dieses Leistungsschutzrecht auf die technisch-organisatorische Leistung, Töne und akustische Werke erstmalig auf einem Tonträger festzuhalten. Neben herkömmlichen Musikproduktionen (Aufnahmen, Erstellung eines Masters) fallen hierunter auch alle anderen Arten von Tonaufnahmen. Die Schutzfähigkeit des auf dem Tonträger verkörperten Tons selbst ist unerheblich und bezieht sich nicht nur auf die Darbietung urheberrechtlich geschützter Werke wie Musikstücke. Es ist möglich, dass Tonträgerherstellerrechte auch an Aufnahmen von Vogelstimmen oder Geräuschen bestehen können. Laut der Gesetzesbegründung ist der Schutzgegenstand das immaterielle Gut, das in der im Tonträger verkörperten Herstellerleistung besteht. Die Person, die die organisatorische Hoheit über die Aufnahme besitzt, wird als Tonträgerhersteller bezeichnet. Der Tonträgerhersteller erbringt maßgebliche technische und wirtschaftliche Leistungen und führt bei Bedarf das Rechtemanagement durch, einschließlich des Abschlusses von Verträgen mit Kunstschaffenden und anderen Beteiligten.

Es ist also möglich, dass ein Unternehmen oder eine natürliche Person, die maschinelles Lernen zur Musikproduktion, Mischung und Mastering einsetzt, Tonträgerherstellerrechte an der entstandenen Aufnahme besitzt, unabhängig davon, ob die durch die KI eingespielte Musik selbst geschützt ist oder nicht. Wenn die KI jedoch gleichzeitig Interpret und Produzent ist und eigenständig den gesamten Prozess von der Tonerzeugung bis zur fertigen Aufnahme organisiert, dürfte ein Tonträgerherstellerecht nicht in Betracht kommen. Nach Artikel 3c des Internationalen Abkommens über den Schutz der ausübenden Kunstschaffenden, der Hersteller von Tonträgern und der Sendeunternehmen vom 26.10.1961 (auch bekannt als Rom-Abkommen) kann nur eine natürliche oder juristische Person Tonträgerhersteller:in sein. Wenn es keine solche Person gibt, die die organisatorische Hoheit über den Produktionsprozess hatte, gibt es auch kein

Tonträgerherstellerrrecht. Es sei darauf hingewiesen, dass eine KI keine Person im Rechtssinne ist.

Für Filmproduktionen gilt dasselbe, lediglich mit dem Unterschied, dass es nicht um die „auf den Ton beschränkte Festlegung der Töne einer Darbietung oder anderer Töne“, sondern um die Festlegung von visuellen und / oder audiovisuellen Leistungen geht.

(Kreutzer, Christiansen, 2021, S.56-57)

Künstlerische Darbietungen wie Gesang, Schauspiel oder Tanz:

Grundsätzlich gibt es "Rechte der ausübenden Künstler" an den Leistungen von Interpreten und Interpretinnen. Nach Artikel 3a des Rom-Abkommens werden als "ausübende Künstler" Schauspieler, Sänger, Musiker, Tänzer und andere Personen definiert, die Werke der Literatur oder Kunst aufführen, singen, vortragen, vorlesen, spielen oder auf irgendeine andere Weise darbieten. Gemäß § 73 UrhG ist ein "ausübender Künstler im Sinne dieses Gesetzes" eine Person, die ein Werk oder eine Ausdrucksform der Volkskunst aufführt, singt, spielt oder auf eine andere Weise darbietet oder an einer solchen Darbietung künstlerisch mitwirkt. Diese Definitionen stellen klar, dass nur Darbietungen von natürlichen Personen durch das Leistungsschutzrecht der ausübenden Künstler:innen geschützt werden können. Da Künstliche Intelligenzen keine Personen, sondern technische Konstrukte sind, kommen ein Schutz ihrer Gesangseinlagen oder anderen Darbietungen, wie beispielsweise rein computergenerierte schauspielerische Leistungen (zum Beispiel in Animationsfilmen), nicht in Betracht (Kreutzer, Christiansen, 2021, S.57-58).

2.9.3 Patentrecht

Es ist möglich und wahrscheinlich, dass KI-Technologien als eigenständige Erfindungen patentiert werden können. Ebenso können Erfindungen, bei denen KI-Technologien als Hilfsmittel und zur Unterstützung eingesetzt werden, als "computer aided inventions" patentierbar sein. Im Falle von rein von KI erfundenen Technologien ("computer generated inventions") sieht die Situation jedoch anders aus. Das Patentrecht, ähnlich wie das Urheberrecht, basiert grundsätzlich auf der Person des Erfinders. Nach geltendem Recht sind Erfindungen nur dann schützbar, wenn sie von einem menschlichen Erfinder stammen. Gemäß § 6 PatentG steht das Recht auf das Patent dem Erfinder oder dessen Rechtsnachfolger zu. Folglich entsteht kein Anspruch auf ein Patent, wenn die Erfindung von einer KI gemacht wurde, da diese keine menschliche Person ist. Dementsprechend können derzeit in den meisten Rechtsordnungen der Welt keine

Patente für rein technische Erfindungen erteilt werden, die von einem Computer generiert wurden. Es ist allerdings möglich, dass sich die diesbezügliche Rechtslage irgendwann ändert (Kreutzer, Christiansen, 2021, S.59).

2.9.4 Geheimnisschutz

Werden Know-how oder Geschäftsgeheimnisse von Algorithmen erzeugt, so können diese nach den Regeln des Gesetzes zum Schutz von Geschäftsgeheimnissen (GeschGehG) gegen eine unbefugte Verwendung geschützt werden. Die Art und Weise, wie diese Geschäftsgeheimnisse entstanden sind, sei es durch menschliche oder maschinelle Generierung, ist dabei unerheblich (Kreutzer, Christiansen, 2021, S.59).

2.10 Risiken und Herausforderungen

Künstliche Intelligenz birgt enorme Potenziale, aber auch Herausforderungen, Befürchtungen und Risiken. In der vorliegenden Arbeit werden fünf Arten von wesentlichen Risiken der KI benannt: Bugs, Cybersecurity, das „Sorcerer’s Apprentice“ Phänomen, eine geteilte Autonomie sowie mögliche negative sozioökonomische Auswirkungen (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49).

Das Risiko von Bugs bezieht sich auf Programmierfehler in der KI-Software. Ähnlich wie bei den Gefahren, die durch Cyber-Attacken entstehen, unterscheidet sich dieses Risiko nicht von den derzeitigen IT-Systemen (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49).

Das „Sorcerer’s Apprentice“-Risiko bezieht sich auf KI-Systeme, die falsche Handlungen ausführen, wenn unklare Instruktionen durch den Nutzer vorliegen. Dies ist ein technisches Problem, das durch Programmierung behoben werden kann, beispielsweise indem Systeme so programmiert werden, dass sie um Klärung bitten oder festgelegte Grenzen programmiert werden. Gleichzeitig besteht auch ein menschlicher Problemfaktor, der abnehmen sollte, wenn Menschen sich daran gewöhnen, mit KI-Maschinen zu interagieren und Wissen entwickeln (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49).

Das Problem der geteilten Autonomie bezieht sich auf KI-Systeme, bei denen die Maschine und der Mensch gemeinsam (auch sequentiell) Dinge kontrollieren. Das Problem kann entstehen, wenn die Übergabe an einen Menschen zu plötzlich ist und dieser nicht bereit ist, die Kontrolle zu übernehmen. (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49).

KI-Systeme beziehungsweise automatisierte Entscheidungsprozesse laufen häufig als Hintergrund-prozesse, die für Menschen, welche durch ihre Entscheidungen beeinflusst werden, oft nicht sichtbar sind. Ihre Entscheidungswege (wie etwa beim Deep Learning) sind häufig nicht nachzuvollziehen, was wiederum auch zur Frage bezüglich der Gerechtigkeit beziehungsweise Fairness der Systeme führt (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49).

Eine bedeutende Herausforderung und ein potenzielles Risiko in diesem Zusammenhang besteht in Verzerrungen, die durch Algorithmen entstehen können, auch bekannt als das "algorithmic bias problem". Die Leistungsfähigkeit von KI hängt stark von der Qualität der zugrunde liegenden Daten ab, aus denen sie lernt, und diese Daten können Verzerrungen enthalten. Wenn maschinelles Lernen auf inhärent verzerrten Daten durchgeführt wird, können die Ergebnisse ebenfalls verzerrt sein. KI-Systeme erkennen Muster in Daten ohne wesentlichen menschlichen Eingriff, was bedeutet, dass Entstehung und Zusammensetzung der Daten eine entscheidende Rolle spielen. Da Daten jedoch nicht immer objektiv und ausgewogen entstehen (z. B. Medienberichte, Einträge in sozialen Netzwerken), können diese Ungleichgewichte in den Daten von der KI übernommen und sogar verstärkt werden (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49).

Verzerrte Daten können aus zwei Gründen entstehen: Erstens, wenn sie die Realität nicht vollständig oder korrekt widerspiegeln, z.B. aufgrund von ungenauen Messmethoden, unvollständiger Datenerfassung oder anderen Fehlern bei der Datenerhebung. Zweitens, wenn der zugrundeliegende Sachverhalt selbst strukturelle Ungleichheiten aufweist, z.B. wenn Beschäftigungsdaten einer Branche ausgewertet werden, in der Männer systematisch gegenüber Frauen bevorzugt werden. KI-Systeme, die mit solchen verzerrten Daten trainiert werden, laufen Gefahr, Modelle zu entwickeln, die diese Verzerrungen replizieren und sogar verstärken (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.49-50).

Ein weiteres Problem im Umgang mit KI-Systemen und automatisierten Entscheidungsprozessen besteht darin, dass maschinelle Lernalgorithmen Schwierigkeiten haben, Ungleichgewichte in Probengrößen zu bewältigen. Dies ergibt sich direkt aus der Tatsache, dass maschinelle Lernalgorithmen auf statistischen Methoden basieren und daher den Gesetzen der statistischen Stichprobengrößen unterliegen. Insbesondere können Lernalgorithmen Schwierigkeiten haben, spezifische Effekte zu erkennen, wenn die Bevölkerung stark segmentiert ist. Da die Lernalgorithmen statistische Schätzmethoden verwenden, wird ihr Schätzfehler umso kleiner, je größer die Datenbasis ist (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.50).

Die Integration von KI-Systemen in soziale und ökonomische Bereiche erfordert die „Umformulierung“ von sozialen Problemen in technische Probleme, so dass

diese von KI-Systemen gelöst werden können. Dies ist keine neutrale Übersetzung: Ein reales Problem ist so zu gestalten, dass es von einem KI-System bearbeitet werden kann, was den Rahmen für Annahmen und möglichen Lösungsvorschläge für das Problem verändert und auch beschränkt (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.50).

Die Entwicklung und Implementierung von KI-Systemen erfordert erhebliche Investitionen in IT-Infrastruktur und Datenressourcen. Die Entwicklung von KI-Systemen erfordert sowohl Rechenleistung als auch große Datenmengen, die mit hohen Kosten verbunden sind. Dadurch werden die Möglichkeiten für KI-Innovationen eingeschränkt und es entsteht ein Wettbewerbsnachteil für Akteure, die nicht über ausreichende Ressourcen verfügen. Dies kann dazu führen, dass nur ein kleiner Teil von Unternehmen in der Lage ist, den Markt für KI-Systeme zu dominieren, was die Wettbewerbsfähigkeit anderer Unternehmen beeinträchtigen kann (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.50).

Obwohl die Automatisierung durch KI-Systeme wahrscheinlich nicht zu einer vollständigen Ersetzung von Arbeitsplätzen durch Maschinen führen wird, wird sie dennoch erhebliche Auswirkungen auf die zukünftige Arbeitswelt haben. Derzeit unterstützen KI-Systeme Menschen bei ihrer Arbeit und verändern Managementstrukturen. Es ist jedoch vor allem wahrscheinlich, dass Berufe mit niedrigen Qualifikationsanforderungen einem höheren Risiko ausgesetzt sind, durch intelligente Maschinen ersetzt zu werden. Diese Entwicklungen werfen wichtige Fragen auf bezüglich bestehender sozialer Sicherheitsnetze und der gerechten Verteilung von Humanressourcen in einer Welt, in der Maschinen zunehmend Arbeiten übernehmen (Scherk, Pöchhacker-Tröscher, Wagner, 2017, S.50).

3 Künstliche Intelligenz im Kreativbereich, dem Digital Design und der digitalen Content-Produktion

Im folgenden Kapitel werden Bereiche, beziehungsweise Anwendungen vorgestellt, die Künstliche Intelligenz in Bezug auf kreative Tätigkeiten anwenden. Überdies werden K.I Anwendungen im Bereich des Digital Design und der digitalen Content-Produktion vorgestellt.

Künstliche Intelligenz wird zunehmend (und oft irrtümlicherweise) mit menschlicher Kreativität und künstlerischer Praxis assoziiert. Da K.I Fähigkeiten wie "Sehen", "Hören", "Sprechen", "Bewegung" und "Schreiben" gezeigt hat, wurde sie in Bereichen und Anwendungen wie Audio-, Bild- und Videoanalyse, Gaming, Journalismus, Drehbuch schreiben, Filme machen, Social-Media-Analyse und Marketing eingesetzt. Eine der frühesten KI-Technologien, die seit mehr als zwei Jahrzehnten verfügbar ist, ist beispielsweise Autotune, das automatisch stimmliche Intonationsfehler korrigiert (Anantrasirichai, Bull, 2021, S.601).

Im Jahr 2016 erregte ein Projekt namens "The Next Rembrandt", geleitet von Microsoft in Zusammenarbeit mit privaten und öffentlichen Institutionen, weltweit große Aufmerksamkeit. Ein Gemälde, das scheinbar aussah, als käme es aus dem Pinsel des niederländischen Meisters, wurde in Amsterdam vorgestellt. Es basierte auf den Ergebnissen eines Deep-Learning-Algorithmus, der über 300 Scans der bestehenden Werke von Rembrandt analysierte und ihre charakteristischsten Merkmale herausfand. Es wurde mithilfe von Deep-Learning-Algorithmen und Gesichtserkennungstechniken erstellt (Zylinska, 2020, S.50).

Kreativität wird im Cambridge Dictionary definiert als "die Fähigkeit, originelle und ungewöhnliche Ideen zu produzieren oder etwas Neues oder Imaginäres zu schaffen". Kreative Aufgaben erfordern im Allgemeinen eine gewisse Originalität, umfangreiche Erfahrung und ein Verständnis des Publikums, während Produktionsaufgaben im Allgemeinen repetitiver oder vorhersehbarer sind, was sie für Maschinen besser geeignet macht. Bisher haben KI-Technologien gemischte Ergebnisse erzielt, wenn sie zur Generierung origineller kreativer Werke eingesetzt wurden. Zum Beispiel erstellt die Plattform GumGum ein neues Kunstwerk, indem es eine kurze Idee des Benutzers als Eingabe erhält. Das Modell wird trainiert, indem es die bevorzugten Werkzeuge und Prozesse aufzeichnet, die der Künstler zur Erstellung eines Gemäldes verwendet. Ein

Turing-Test zeigte, dass es schwierig ist, diese von einer KI generierten Produkte von denen zu unterscheiden, die von Menschen gemalt wurden. KI-Methoden produzieren oft ungewöhnliche Ergebnisse, wenn sie eingesetzt werden, um neue Handlungsstränge für Bücher oder Filmskripte zu erstellen. Die Anwendung Botnik verwendet einen KI-Algorithmus, um Texte vorhandener Bücher automatisch neu zu vermischen und ein neues Kapitel zu erstellen. In einem Experiment fütterte das Team die sieben Harry-Potter-Romane durch den Textalgorithmus, und der "Bot" erstellte sowohl seltsame als auch amüsante Sätze. Wenn K.I jedoch verwendet wird, um weniger strukturierte Inhalte zu erstellen (beispielsweise musikalische Werke), so kann sie angenehme Unterschiede aufzeigen (Anantrasirichai, Bull, 2021, S.601).

Im Produktionsbereich beispielsweise hat Twitter automatisches Zuschneiden angewendet, um Thumbnails zu erstellen, die den wichtigsten Teil eines Bildes zeigen. Die BBC hat ein Proof-of-Concept-System zur automatisierten Berichterstattung über Live-Events erstellt. Darin führt das auf KI basierende System automatisch Shot-Framing (Weitwinkel-, Mittel- und Nahaufnahmen), Sequenzierung und Shot-Auswahl durch. Die anfänglichen Ergebnisse zeigen jedoch, dass der Algorithmus noch verbessert werden muss, um menschliche Bediener zu ersetzen. Nippon Hoso Kyokai (NHK, Japans Rundfunkgesellschaft) hat eine neue KI-gesteuerte Rundfunktechnologie namens "Smart Production" entwickelt. Diese extrahiert Ereignisse und Vorfälle aus verschiedenen Quellen wie Social-Media-Feeds (z.B. Twitter), lokalen Regierungsdaten und Interviews und integriert sie in ein für Menschen zugängliches Format (Anantrasirichai, Bull, 2021, S.601).

Es ist herausfordernd zu wissen, was genau im Bereich der Schnittstelle zwischen KI und kreativen Aktivitäten vor sich geht. Die Arbeit an KI hat in den letzten Jahren international stark zugenommen, mit weitaus höheren Forschungsaktivitäten und vielen neuen Entwicklungen. Es ist daher schwierig, einen Überblick darüber zu bekommen, welche Veränderungen stattfinden. Eine zusätzliche Herausforderung besteht darin, dass die kreativen Industrien ein kompliziertes und miteinander verbundenes Set von Bereichen darstellen, was es schwierig macht, die Auswirkungen technologischer Veränderungen in ihnen genau nachzuverfolgen (Davies, Klinger, Mateos-Garcia, Stathoulopoulos, 2020, S.4).

3.1 Auswirkungen von Künstlicher Intelligenz in den Designprozess

Vor einigen Jahren waren Bilder, die mithilfe von künstlicher Intelligenz erstellt wurden, für den ungeübten Betrachter deutlich von Fotografien zu unterscheiden. Es war erstaunlich, dass solche Bilder überhaupt möglich waren. Gleichzeitig war man jedoch überzeugt, dass diese Bilder aufgrund ihrer begrenzten Anwendungsbereiche und des fehlenden täuschenden Potenzials, nicht als Massenware die Öffentlichkeit täuschen oder schädigen würden. Die Technologie der generativen künstlichen Intelligenz hat inzwischen einen Punkt erreicht, an dem die von ihr erzeugten Bilder als Deep Fakes bekannt sind und zu einem kulturellen Massenphänomen geworden sind. Unter anderem gibt es Websites wie "This Person Does Not Exist", auf denen auf Knopfdruck Porträts von nichtexistierenden Menschen erzeugt werden können, die mit dem bloßen Auge nicht mehr von echten Porträtfotografien zu unterscheiden sind. Auch Adobe Photoshop bietet neuronale Filter an, mit denen man auf Fotos porträtierte Personen lächeln lassen kann, selbst wenn diese auf dem Originalbild nicht gelächelt haben. Strafverfolgungsbehörden nutzen Deep Fakes, um Bilder pädophiler Straftaten zu generieren und Täter anzulocken. Außerdem werden Deep Fakes in der Werbung eingesetzt, um jüngere Versionen von Schauspielern zu generieren und diese so in den Werbungen selbst zu verwenden. Deep Fakes werden auch in Filmdokumentationen eingesetzt, um Menschen in Ländern zu schützen, in denen sie aufgrund ihrer sexuellen Orientierung verfolgt werden. In diesen Fällen werden die wahren Gesichter der betroffenen Personen durch Deep Fakes ersetzt, um sie unkenntlich zu machen (Engenhardt, Löwe, 2022, S.11).

Durch die rasanten Fortschritte der KI-Technologie haben sich die Bilder von ihrem Original abgelöst und sind zu eigenständigen Entitäten geworden. Dies hat zur Folge, dass Deep Fakes nicht mehr nur in begrenzten Anwendungsbereichen eingesetzt werden, sondern ein enormes Potenzial für Täuschung und Manipulation in der Öffentlichkeit haben (Engenhardt, Löwe, 2022, S.11).

Deep Fakes sind allerdings nur ein Beispiel für eine Vielzahl von Technologien und Methoden, die unter dem Begriff der Künstlichen Intelligenz zusammengefasst werden. KI ist nicht nur auf die Generierung von Bildern beschränkt, sondern kann auch Stimmen, Sprache, Töne, Gesten und Emotionen erkennen und verarbeiten. In der jüngsten Zeit können sogar Bilder durch Sprache generiert werden. Darüber hinaus ist KI in der Lage, Muster in riesigen Datenmengen zu erkennen und somit beispielsweise bei der Analyse von Kaufentscheidungen oder der Identifikation erfolgreicher Design-Features zu helfen (Engenhardt, Löwe, 2022, S.11).

Die Einbindung künstlicher Intelligenz in die Werkzeuge der Gestaltung hat weitreichende Konsequenzen für die Welt der Gestaltung und der Gestaltenden.

Die Nutzung von KI stellt folgende zentrale Fragen, da sie eine Veränderung des bisherigen Gestaltungsprozesses mit sich bringt:

- Welche Wirkung hat KI auf das Design und den Designprozess?
- Wie verändert KI die Aufgabenbereiche und Arbeitsweisen von Designer:innen?
- Welche Möglichkeiten ergeben sich durch KI-Designwerkzeuge?
- Wie sind diese Werkzeuge in den Gestaltungsprozess eingebunden?
- Entstehen durch KI neue Gestaltungs-methoden und -prozesse?

(Engenhardt, Löwe, 2022, S.11)

Neben diesen unmittelbar designspezifischen Fragen ergeben sich aber auch ethische und rechtliche Fragen:

- Wem gehören die Bilder?
- Welche ethischen Konsequenzen hat KI?

Nicht nur wird Künstliche Intelligenz in den Gestaltungswerkzeugen selbst eingesetzt, sondern auch in digitalen Produkten integriert, die von Designer:innen konzipiert und umgesetzt werden. Designer:innen verwenden KI dabei als Designmaterial und kreieren somit innovative "User Experiences" für intelligente Produkte und Medien. Dabei stellt sich die Frage, wie KI das Erlebnis digitaler Produkte verändert und welche Möglichkeiten sich dadurch für die Nutzenden ergeben. Überdies hat K.I. auch Einfluss auf das Rollenverständnis der Designer:innen. Dabei ergeben sich Fragen, wie sich das Verständnis von Kreativität und Gestaltung durch KI verändert und welche neuen Fähigkeiten Designer:innen im Umgang mit KI benötigen (Engenhardt, Löwe, 2022, S.12).

Obwohl Künstliche Intelligenz kein neues Phänomen darstellt und bereits weitreichend in digitale Produkte integriert ist, sind die meisten der oben genannten Fragen bislang nicht hinreichend beantwortet. In vielen Fällen fehlt sogar eine fundierte wissenschaftliche Basis für die Auseinandersetzung mit der KI-Technologie im Designbereich. Es ist daher überraschend, dass die Fragen bezüglich des Designs mit KI bisher weder systematisch beantwortet noch in eine theoretische Ordnung überführt oder aus der Perspektive des Designs umfassend wissenschaftlich diskutiert und ausgewertet wurden (Engenhardt, Löwe, 2022, S.12).

3.2 Anwendungsmöglichkeiten von K.I in der digitalen Content-Produktion

Neben den ethischen Fragen und möglichen Problemen, die der Einsatz von Künstlicher Intelligenz im Kreativbereich mit sich bringt, bietet sie aber auch viele Möglichkeiten und Vorteile, die folgend vorgestellt werden.

3.2.1 Generative K.I

Generative KI ist eine Technologie, die maschinelles Lernen und künstliche Intelligenz einsetzt, um neue Inhalte aus vorhandenen Daten zu generieren. Im Gegensatz zu anderen Ansätzen des maschinellen Lernens, die darauf abzielen, vorhandene Muster zu erkennen und Vorhersagen zu treffen, geht es bei der generativen KI darum, kreative Prozesse zu automatisieren und neue, bisher unbekannte Inhalte zu generieren. Dabei wird ein Netzwerk von zwei miteinander konkurrierenden neuronalen Netzen eingesetzt: Ein Generator-Netzwerk erzeugt neue Daten, während ein Diskriminator-Netzwerk die Echtheit der Daten beurteilt und dem Generator Feedback gibt, um seine Leistung zu verbessern. Die generative KI hat bereits beeindruckende Ergebnisse erzielt und wurde in verschiedenen Anwendungen eingesetzt, darunter der Erstellung von Bildern, Texten und Musik. Sie ermöglicht es, Inhalte zu generieren, die von Menschen kaum zu unterscheiden sind und liefert somit originelle Ergebnisse, die auf menschlichem Denken und menschlicher Erfahrung beruhen. Dabei ist sie nicht auf eine bestimmte Art von Daten beschränkt, sondern kann mit verschiedenen Formaten wie Text, Bildern, Audio und Videos umgehen. Generative KI stellt somit eine wichtige Entwicklung im Bereich der künstlichen Intelligenz dar und hat das Potenzial, kreative Prozesse in verschiedenen Branchen zu automatisieren und zu optimieren (Industry of Things, 2022).

Generative Künstliche Intelligenz ermöglicht die Erzeugung realistisch aussehender Fotos von menschlichen Gesichtern, Objekten und Szenen. Mithilfe von generativer K.I. können Bilder, die bei schlechten Licht- oder Wetterbedingungen aufgenommen wurden, in die gewünschten Bedingungen umgewandelt werden. Ebenso können Farbbilder in Schwarz-Weiß-Bilder und Fotos bei Tag in Fotos bei Nacht umgewandelt werden.

Durch den Einsatz von Gesichtssynthese und Stimmenklonen kann die Originalstimme eines Schauspielers mit einer Lippensynchronisation abgeglichen werden. Dadurch können Filme und Werbungen auf eine effektivere Art und Weise erstellt werden, die authentisch und überzeugend sind. Auch alte, niedrig aufgelöste Bilder und Filme können durch die Anwendung von generativer KI verbessert werden. Durch die Hochskalierung auf 4K und höher können genauere, klarere und detailliertere Bilder erzeugt werden, was zu einem verbesserten

Seherlebnis führt. Die generative K.I. erzeugt dabei 60 Bilder pro Sekunde anstelle von 23 oder weniger, beseitigt Rauschen und fügt Farbe hinzu (Industry of Things, 2022).

Generative K.I. kann auch im Gesundheitswesen für verschiedene Zwecke eingesetzt werden, wie beispielsweise für die Früherkennung bösartiger Tumore. Durch die Anwendung des Generative Adversarial Networks Modeling kann die KI verschiedene Winkel einer Röntgenaufnahme berechnen und das mögliche Ausmaß des Tumors visualisieren. Dies kann Medizinern helfen, schneller eine Diagnose zu stellen und eine angemessene Behandlung zu planen. Darüber hinaus können Mediziner mithilfe von maschinellem Lernen das Verhalten und die Bewegungsmuster von Patienten beobachten, bevor sie mit Hilfe generativer KI das Design einer Prothese berechnen und eine perfekte Prothese entwickeln (Industry of Things, 2022).

Generative KI spielt eine immer wichtigere Rolle bei der Texterstellung und bietet zahlreiche Möglichkeiten für die Automatisierung von Schreibprozessen. Die Fähigkeiten der KI reichen mittlerweile über das einfache Generieren von Anzeigenzeilen, Produktbeschreibungen und Betreffzeilen für Outreach-Kampagnen hinaus. Durch die Eingabe einer Reihe von Schlüsselwörtern kann die KI einen kompletten Text erstellen, der auf die spezifischen Bedürfnisse des Benutzers zugeschnitten ist. Darüber hinaus können diese Tools auch analysieren, welche Texte am erfolgreichsten sind und dementsprechend neue, ähnlich konzipierte, aber bessere Texte erstellen (Industry of Things, 2022).

3.2.2 Künstliche Intelligenz in der Filmindustrie

Die Anwendung von Künstlicher Intelligenz hat eine bedeutende Revolution in der Filmproduktion bewirkt. Von der Erstellung von Drehbüchern bis hin zur Postproduktion hat die K.I. zunehmend Verwendung gefunden. Die Technologie wird nicht nur eingesetzt, um Drehbücher zu schreiben, sondern auch um Szenenbilder sowie Spezialeffekte zu erstellen. Darüber hinaus wird sie zur Optimierung von Marketingkampagnen genutzt, um beispielsweise das Engagement des Publikums zu verfolgen. Die Integration von K.I. hat somit die Filmproduktion und Vermarktung von Filmen grundlegend verändert und eröffnet neue Möglichkeiten, ein breiteres Publikum zu erreichen (Hooksounds, 2023).

Erstellung von Drehbüchern:

Die Nutzung von künstlicher Intelligenz ermöglicht die Erstellung von überzeugenden Charakteren und Geschichten in Drehbüchern. Dabei kann K.I. durch Analyse ausgewählter Daten lernen und auf dieser Grundlage neue Skripte erstellen. Durch die Anwendung von K.I. können auch bereits bestehende Drehbücher überarbeitet und verbessert werden. Hierbei ist es möglich, dass die

K.I. Handlungslücken erkennt oder alternative Enden vorschlägt (Hooksounds, 2023).

Vorproduktion

Die Vorproduktion ist ein weiterer Bereich, in dem K.I. nützlich eingesetzt werden kann. In der Vorproduktion kann K.I. bei der Suche nach Film-Locations helfen, frühere erfolgreiche Locations analysieren, um so neue Orte mit ähnlichen Eigenschaften vorzuschlagen. Auch bei der Organisation der einzelnen Szenen sowie bei der Auswahl von geeigneten Schauspielern kann K.I. eingesetzt werden (Hooksounds, 2023).

Produktion

Künstliche Intelligenz kann auch die eigentliche Produktion effizienter und rentabler gestalten, indem sie die besten Aufnahmen einer Szene erkennen und vorschlagen kann, die Beleuchtung anpasst oder geeignete Kamerawinkel für die entsprechende Szene vorschlägt (Hooksounds, 2023).

Post Produktion

Die Anwendung von künstlicher Intelligenz findet zunehmend auch in der Filmpostproduktion Anwendung. Insbesondere bei der Farbkorrektur, Spezialeffekten und Videobearbeitung kann die K.I. hilfreich sein. Durch den Einsatz von K.I. können langwierige und mühsame Aufgaben wie das Anpassen von Farben zwischen verschiedenen Szenen oder das Erstellen von nahtlosen Übergängen zwischen Szenen automatisiert werden (Hooksounds, 2023).

Künstliche Intelligenz kann dem Menschen im Allgemeinen dabei helfen, automatisierte, datengesteuerte Entscheidungen zu treffen. Durch den Einsatz von K.I. können Daten schnell und effizient gesammelt und analysiert werden, was zu fundierteren Entscheidungen führt. In der Filmindustrie kann die K.I. besonders nützlich sein, um eindrucksvolle und kosteneffektive Filme zu erstellen. Durch den Einsatz von K.I. zur Automatisierung bestimmter Aufgaben in der Filmproduktion kann Zeit und Geld eingespart werden, ohne die Qualität des Endprodukts zu beeinträchtigen (Hooksounds, 2023).

Ein fortschrittlicher Ansatz in der Filmproduktion wurde erstmals im Jahr 2016 durch den Trailer des Films "Morgan" (Regie: Luke Scott, USA 2016) realisiert. Dabei wurde der Schnitt des Trailers nicht von einem Editor realisiert, sondern von IBMs Watson Supercomputer. Der Supercomputer nutzte sein maschinelles Lernverfahren, welches auf Basis von anderen Trailern trainiert wurde, um eigenständig Ausschnitte aus dem Film auszuwählen und in den Trailer einzubauen (Movie-College, 2023).

Mittlerweile sind K.I. Funktionen auch für jeden Anwender an seinem eigenen Arbeitsplatz verfügbar. Verschiedene Software-Unternehmen haben an Möglichkeiten gearbeitet, K.I. in Anwendungen der Film,- und Medienbranche zu integrieren. In Adobe Premiere Pro ist es beispielsweise möglich, dass Aufnahmen für unterschiedliche Ausgabeformate „reframed“ werden. Dabei kann der Ausschnitt des Videomaterials verändert werden. Die K.I. analysiert die Aufnahmen und entscheidet, wie der neue Ausschnitt aussehen sollte. In Adobe Photoshop ist mit „Sensei“ ein Mix aus künstlicher Intelligenz und Machine Learning verfügbar. Diese beinhaltet beispielsweise die Anwendung „Inhaltsbasierte Füllung“ (Content Aware Fill). Dabei muss nur ein Teil des Bildes definiert werden, der entfernt werden soll und die K.I. sucht alle ähnlichen Stellen im Bild und entfernt diese. Dies nennt sich "One Click Delete and Fill" und kann im besten Fall Personen und Objekte vollständig aus einem Bild entfernen und die freie Fläche dank K.I. mit Teilen aus dem umgebenden Raum auffüllen. Das Ergebnis ist abhängig von der Komplexität des Hintergrundes. Die gleiche Funktion auch in After Effects enthalten (Movie-College, 2023).

Adobe hat in Zusammenarbeit mit der Stanford University im Jahr 2017 an Lösungen geforscht, um den Filmschnitt zu vereinfachen. Hierbei wurde eine intelligente Software entwickelt, welche in der Lage ist, die Einstellungen eines Films genau nach den Szenen des Drehbuchs zu sortieren. Die Software erkennt automatisch, um welche Einstellungsgröße es sich handelt und welche Filmfiguren in dem Frame zu sehen sind. Der Benutzer kann dann bestimmte Präferenzen festlegen, wie beispielsweise die Option, dass immer derjenige zu sehen ist, der spricht. Es ist auch möglich, dem Programm vorzugeben, ob Jump Cuts erlaubt sind oder nur sanfte und unsichtbare Übergänge gewünscht sind. Der Benutzer kann der Software auch vorgeben, dass jede Szene mit einer weiten Einstellungsgröße beginnen soll. In gewissem Umfang ist es somit möglich, einen bestimmten Stil vorzugeben, den die Software dann umsetzt. Das Programm schneidet dann das Material in kurzer Zeit.

Allerdings sind die Ergebnisse noch nicht perfekt. Obwohl das Programm mit Dialogen einigermaßen zurechtkommt, gibt es viele andere Aspekte, die für die K.I. noch sehr schwierig sind. Dinge wie Tempo, Action, beobachtende Einstellungen oder Anschlussfehler sind für die K.I. schlichtweg schwer zu verstehen und zu bearbeiten (Movie-College, 2023).

Die Fähigkeit von K.I., Bildinformationen zu interpretieren, eröffnet eine Vielzahl von unterstützenden Funktionen, die die Arbeit von Editoren erleichtern. In absehbarer Zukunft werden Schnittprogramme den Editoren wahrscheinlich Vorschläge machen, wo und wie lange bestimmte Teile des Materials platziert werden sollten. Es macht derzeit jedoch wenig Sinn, einen vollständigen Schnitt ausschließlich von K.I. erledigen zu lassen, da die K.I. eines der Hauptprobleme

hat, dass sie Zusammenhänge nicht wirklich versteht und nicht "fühlen" kann. Es ist interessant zu bemerken, dass KI viele Aufgaben, die für unser Gehirn mühsam sind, mit Leichtigkeit erledigen kann, während andere Dinge, die das Gehirn leicht lösen kann, für K.I. sehr schwierig sind. Kreativität ist ein Bereich, in dem der Mensch der K.I. noch weit überlegen ist (Movie-College, 2023).

Allerdings gibt es immer noch große Herausforderungen bei der Bilderkennung, insbesondere in Bezug auf die semantische Interpretation von Bildinhalten. Während KI-Systeme in der Spracherkennung mittlerweile eine hohe Genauigkeit von 95% erreichen, sind die Ergebnisse der Bilderkennung oft weniger präzise. Obwohl KI-Systeme Muster in Bildern identifizieren und in Datenbanken erfassen können, verstehen sie oft nicht, was diese Muster bedeuten. Um Bedeutungen in Bildinhalten zu erkennen, müssen KI-Systeme umfangreiche Suchläufe in anderen Datenbanken durchführen, um die bestmögliche Übereinstimmung zu finden.

Jedoch fehlen der KI oft die menschlichen Eigenschaften wie Emotionen und der gesunde Menschenverstand, um Bildinhalte zu interpretieren und Bedeutungen zu verstehen. Dies macht es schwierig für KI-Systeme, die subtilen Anforderungen und Erwartungen von Kunden oder Regisseuren im Schneiderraum zu verstehen. Cutter, die oft zwischen den Zeilen lesen und subtile Andeutungen verstehen können, sind derzeit noch unersetzlich in der Film- und Medienbranche. Obwohl KI-Systeme nützliche Funktionen bieten, die die Arbeit von Editoren erleichtern können, sind sie noch nicht in der Lage, die Kreativität und das Einfühlungsvermögen von Menschen zu ersetzen.

Künstliche Intelligenz sollte eher als Assistent betrachtet werden, der einem gewisse Routineaufgaben abnehmen kann. Von der Farbkorrektur über das Entfernen von Objekten bis hin zu visuellen Effekten und Bildstabilisierung reichen die bereits vorhandenen Anwendungen. Die Einsparung von Zeit, die bisher für eher mechanistische Aufgaben benötigt wurden, kann enorm sein und neue Freiräume für mehr Kreativität eröffnen. Künstliche Intelligenz kann so gesehen, größere Kreativität ermöglichen (Movie-College, 2023).

4 Methoden zur Trennung von Vordergrund-Objekten und Hintergrund in Videomaterial

Der Prozess der Erstellung einer Maske für ein Objekt wird durch eine Vielzahl von Begriffen bezeichnet, was darauf zurückzuführen ist, dass er ein konstanter und allgegenwärtiger Bestandteil des Compositing-prozesses ist. Es ist möglich eine Maske für ein Objekt zu "erstellen" oder das Objekt aus einer vorhandenen Szene zu "extrahieren". Wenn der verwendete Prozess überwiegend prozedural ist, fallen häufig Begriffe wie das Erstellen einer Matte ("pulling a matte") und Keying. Keying ist eigentlich ein breiterer Begriff, der oft nicht nur den Prozess der Erstellung einer Matte beschreibt, sondern auch den Prozess des Kombinierens eines Elements mit einem anderen einschließt. (Brinkmann, 2008, S. 189).

Technisch gesehen soll eine Matte eine "opacity map" (Deckkraft) des ausgewählten Vordergrundobjekts innerhalb eines Bildes sein. Sie soll Halb-Transparenzen an den Mischkanten und anderen teilweise transparenten Regionen korrekt wiedergeben (Wright, 2010, S.15).

Ein Beispiel für den Einsatz einer einfachen Matte wäre eine Split-Screen Komposition, bei der zwei Bildquellen kombiniert werden, wobei eine die linke Hälfte und die andere die rechte Hälfte der Komposition ausfüllt. In diesem Fall würde die benötigte Matte durch eine einfache gerade Linie dargestellt werden, die die Grenzen der jeweiligen Hälfte definiert. Wesentlich häufiger wird jedoch ein Objekt in eine Szene eingefügt, dessen Kanten eine deutlich komplexere Form aufweisen. Es wird eine Matte benötigt, die die Grenzen oder Umrisse des Objekts genau darstellt und die alle Bereiche im Inneren des Objekts vollständig umschließt (Brinkmann, 2008, S. 189 -190).

Jedoch selbst bei einem unbewegten Objekt kann es schwierig sein, die Kanten genau einzuzeichnen, sodass die gewünschte Qualität der Kanten eingehalten werden kann. Daher werden für etwaige Aufgaben häufig Software-Algorithmen eingesetzt, die dabei helfen ein Objekt von dessen Hintergrund zu trennen. Bei der Komposition von Bildsequenzen sind Situationen, in denen statische Matten verwendet werden, recht selten. Viel häufiger besteht die Notwendigkeit, eine Matte für ein Objekt zu erstellen, das sich innerhalb des Bildes oder durch das Bild bewegt. Hierfür wird eine bewegte Matte benötigt. Es gibt zwei Ansätze, um diese Mattes zu generieren. Der erste Ansatz wäre das manuelle Zeichnen einer Matte für das betreffende Objekt, jedoch über jeden einzelnen Frame der

Sequenz hinweg. Dieser Prozess, bekannt als Rotoskopie, wird immer noch recht häufig angewendet, jedoch selten als vollständig frame-genauer Vorgang (Brinkmann, 2008, S. 191).

Idealerweise können Verfahren angewendet werden, die eher prozeduraler Natur sind. Halbautomatisierte Prozesse, bei denen einige anfängliche Parameter bestimmt werden, die in der Lage sind, eine Matte zu extrahieren, und dann die Software diese gleichen Parameter über eine Sequenz von Bildern anwenden kann (Brinkmann, 2008, S. 192).

4.1 Rotoscoping

Rotoscoping ist oft einer der ersten Jobs, den ein angehender digital Artist bei einem etablierten Visual-Effects-Unternehmen bekommt. Obwohl es im Allgemeinen nicht als besonders glamouröse Arbeit angesehen wird, ist es ein äußerst wichtiger Teil der Produktions-Pipeline und oft der Schlüsselfaktor für den Erfolg oder Misserfolg einer bestimmten Szene. Obwohl der Begriff "Rotoscoping" ursprünglich verwendet wurde, um einen Arbeitsablauf zu beschreiben, der auf jedem einzelnen Frame einer Sequenz durchgeführt werden musste, hat er sich seitdem zu einer allgemeineren Beschreibung eines Prozesses entwickelt, der wo immer möglich, prozedurale oder algorithmische Unterstützung einschließt. Konkret bezieht es sich in der Regel auf den Prozess der Erstellung einer Matte für ein bestimmtes Objekt in einer Szene, indem man eine Spline-Form an bestimmten Schlüsselbildern der Bildsequenz definiert und die Software die Form für die Zwischenbilder interpolieren lässt (Brinkmann, 2008, S. 192).

4.2 Luma-Key Mattes

Eine der häufigsten Methoden zur Extraktion einer Matte für ein bestimmtes Objekt basiert auf der Manipulation der Luminanz-Werte in einer Szene. Dies wird in der Regel als Luminanz-Keying bezeichnet und beinhaltet die Anwendung einiger grundlegender Bildverarbeitungsoperatoren, um die Luminanz-Werte auszuwählen, die in der Matte enthalten oder ausgeschlossen werden sollen.

Die Technik ist im Allgemeinen am nützlichsten, wenn das Objekt, das man aus der Szene extrahieren möchte, signifikant heller oder dunkler ist als der Hintergrund, von dem man es trennen möchte. Der ideale Fall wäre ein helles Objekt, das vor einem schwarzen Hintergrund aufgenommen wurde, aber oft gibt es Situationen, in denen der Helligkeitsunterschied zwischen Vordergrund und Hintergrund ausreicht, um eine brauchbare Matte zu extrahieren (Brinkmann, 2008, S. 207).

Der Begriff "Luma-Key" stammt aus dem Bereich der Video-Produktion, wo das Videosignal in Luminanz und Chrominanz aufgeteilt wird. Der Luminanz Anteil (Helligkeit) des Videos wird häufig verwendet, um eine Matte zu erstellen, um ein bestimmtes Element für eine besondere Behandlung zu isolieren. Im Bereich des digitalen Compositing kann dieser Prozess auch als Luminanz Matte bezeichnet werden; jedoch wird diese Funktion in den meisten Compositing-Programmen als Luma-Key-Node bezeichnet. Unabhängig davon, wie es genannt wird, funktioniert es in beiden Bereichen gleich. Ein Teil der Luminanz eines Bildes wird verwendet, um eine Matte zu erstellen. Diese Matte (Luma-Key) kann dann auf vielfältige Weise genutzt werden, um ein bestimmtes Element für eine ausgewählte Manipulation zu isolieren. Luma-Key-Mattes sind einfach zu erstellen und flexibel in ihrer Verwendung, da das Element von Interesse oft dunkler oder heller als der Rest des Bildes ist (Wright, 2010, S.16).

Ein Luma-Keyer nimmt ein RGB-Bild auf und berechnet eine Helligkeitsversion davon, die ein monochromes (einkanaliges Graustufen-) Bild darstellt. Ein Schwellenwert wird festgelegt, und alle Pixelwerte, die größer oder gleich dem Schwellenwert sind, werden auf 100% Weiß gesetzt, während alle Pixelwerte, die kleiner als der Schwellenwert sind, auf Schwarz gesetzt werden. Dies erzeugt eine hart begrenzte Matte, die für viele Anwendungsmöglichkeiten ungeeignet ist. Daher bieten einige Luma-Keyer eine zweite Schwellenwert-Einstellung an, um eine weiche Kante in die Matte einzufügen (Wright, 2010, S.16).

4.3 Chroma-Key Mattes

Chroma-Key-Mattes stammen ebenfalls aus der Videobearbeitung, bei der das Videosignal in Luminanz und Chrominanz getrennt wird. Der Chrominanz (Farb-) Teil des Videos wird verwendet, um eine Matte (Key) zu erstellen, um ein gewähltes Objekt zu isolieren. Die gleichen Prinzipien werden beim digitalen Compositing verwendet. Da der Chroma-Key auf der Farbe eines Objekts basiert, kann er selektiver und daher flexibler angewendet werden als der Luma-Key.

Ein Chroma-Keyer nimmt ein RGB-Bild auf und wandelt es intern in eine HSV- (Hue, Saturation, Value) Darstellung um. Der Grund, warum die ursprüngliche RGB-Version nicht intern verwendet wird, ist, dass der Keyer zwischen verschiedenen Sättigungsstufen unterscheiden muss, was kein RGB-Attribut ist, sondern ein HSV-Attribut. Ein Ausgangswert für RGB wird festgelegt, dann werden zusätzliche Toleranzen festgelegt, die die Aufnahme von Sättigungs- und Wertebereichen (Helligkeit) in die Matte ermöglichen. Ein guter Chroma-Keyer ermöglicht auch eine Art von "Toleranz"-Einstellung, um einen weichen Übergang mit sanften Kanten zu ermöglichen (Wright, 2010, S. 23).

Der Chroma-Key ist aus zwei wichtigen Gründen sehr flexibel. Erstens ermöglicht er es, jede beliebige Farbe auszuwählen. Es muss beispielsweise kein sorgfältig kontrollierter Bluescreen verwendet werden. Zweitens berücksichtigt er die natürlichen Variationen in der Farbe realer Oberflächen, indem er Sättigungs- und Wertebereiche besitzt, um sein Akzeptanzfenster zu erweitern. Beispielsweise bei der Farbe von menschlicher Haut werden die Schattenbereiche nicht nur dunkler, sondern weisen auch eine geringere Sättigung auf. Um eine Matte auf der Haut zu erstellen, darf nicht nur der ausgewählte RGB-Hauptton in die Auswahl miteinbezogen werden, sondern auch die Schatten mit geringerer Sättigung.

Allerdings stellt der Chroma Key keine hochwertige Matte dar. Er neigt dazu, Mattes mit harten Kanten zu erzeugen, die verwischt, erodiert oder anderweitig verarbeitet werden müssen, um brauchbar verwendet werden zu können. Halbtransparente Bereiche können ebenfalls nur sehr schlecht verarbeitet werden (Wright, 2010, S. 24).

4.4 Difference Mattes

Difference Mattes funktionieren durch das Erkennen von Unterschieden zwischen zwei Aufnahmen. In einer Aufnahme befindet sich das gewählte Objekt, das isoliert werden soll. Bei der anderen Aufnahme handelt es sich um eine „Clean Plate“; also einer Aufnahme, mit demselben Hintergrund und Motiv, in dem das gewählte Objekt jedoch nicht zu sehen ist. Zum Beispiel eine Szene mit einem Schauspieler im Bild und dieselbe Szene ohne den Schauspieler. Dies erfordert jedoch, dass die Szene entweder zweimal mit einer statischen Kamera aufgenommen wird oder, bei bewegter Kamera, zwei Durchläufe mit einer Bewegungssteuerung erfolgen. In der Praxis geschieht dies jedoch eher selten. Jedoch ist es möglich eine eigene Clean Plate zu erstellen, in der Bereiche beziehungsweise Teile mehrerer Frames kombiniert werden. In sämtlichen Fällen ist eine Clean Plate erforderlich (Wright, 2010, S. 27).

Ein ausgewähltes Objekt über dessen Hintergrund ist in Abbildung 4 dargestellt. Darin ist ein monochromes Bild mit einem einfachen Kreis auf einem diffusen Hintergrund zu sehen und dient folgend als einfaches Beispiel. Auf der rechten Seite befindet sich die Clean Plate, in der nur der Hintergrund enthalten ist. Um nun eine grobe Difference Matte zu erstellen wird der Unterschied zwischen der Originalaufnahme (links) und der Clean Plate (rechts) herangezogen. Die Pixel, die sich innerhalb des Kreises befinden, weisen einen Unterschied zwischen der Originalaufnahme und der Clean Plate auf.

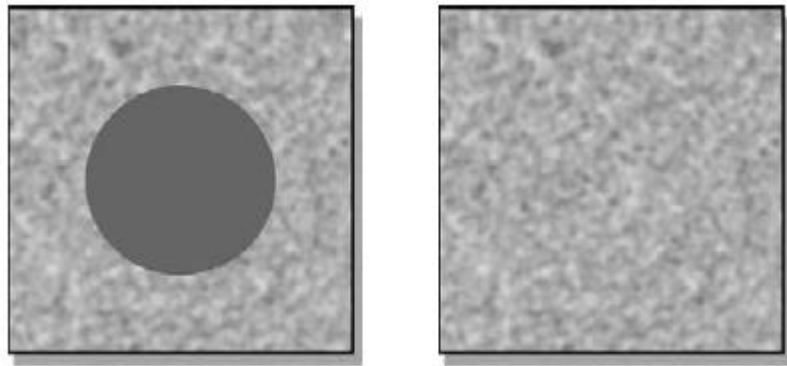


Abbildung 4. Difference Mattes: Originalaufnahme und Clean Plate (Wright, 2010).

In der realen Welt sind die Hintergrundbereiche in der Originalaufnahme und der Clean Plate aufgrund von Filmkörnung und anderen Unterschieden in der Clean Plate, wie zum Beispiel einer leichten Veränderung der Beleuchtung, einem vom Wind gewehten Blatt oder einer verschobenen Kamera, niemals tatsächlich identisch (Wright, 2010, S. 28).

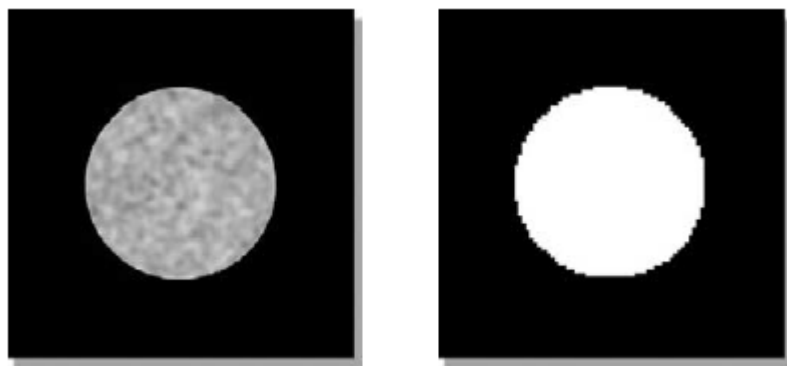


Abbildung 5. raw Difference Matte und binäre Difference Matte (Wright, 2010).

In Abbildung 5 ist links eine grobe („raw“) Difference Matte zu sehen. Dabei handelt es sich um keine saubere Matte, die sofort einsatzbereit wäre. Um diese in eine Difference Matte (rechts) umzuwandeln, wird ein Schwellenwert definiert und Pixel unter jenem Wert werden Schwarz dargestellt; Pixel über dem Schwellenwert werden weiß dargestellt. Dies erzeugt jedoch eine Harte Kante an den Rändern der Objektmaske. Um dies zu vermeiden, wird ein zweiter Schwellenwert definiert, um eine weichere Kante zu ermöglichen (Wright, 2010, S. 28-29).

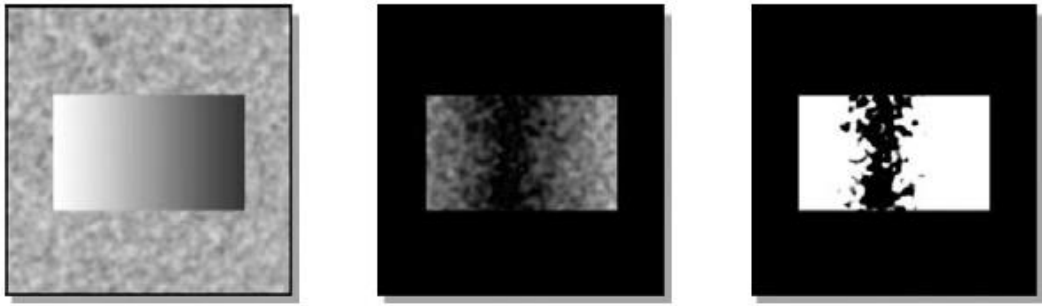


Abbildung 6. Difference: komplexes Objekt (Wright, 2010).

Inhalte mit komplexeren Formen und Objekten stellen jedoch ein Problem für die Verwendung von Difference Mattes dar. In Abbildung 6 ist ein Objekt (im konkreten Fall ein Farbverlauf von Weiß zu Schwarz) zu sehen, in dem sowohl Pixel enthalten sind, die heller als der Hintergrund sind, als auch Pixel, die dunkler als der Hintergrund sind. In der Mitte des Objekts haben die Pixel eine teilweise idente Helligkeit, wie der Hintergrund. Wenn die raw Difference Matte (Mitte) abgeleitet wird, enthält sie einen schwarzen Streifen in der Mitte, da die Pixel im Objekt, die nahe an den Pixelwerten im Hintergrund liegen, zu Differenzen bei oder nahe Null führen. Die linke und rechte Seite des Objekts sind wesentlich heller oder dunkler als der Hintergrund, was zu größeren Unterschieden in den Werten führt. Wird die raw Difference Matte mittels einem Schwellenwert in eine binäre Darstellung umgewandelt, so wird ein grobes „Loch“ in der Mitte der Matte erzeugt. Genau hier zeigt sich das Problem von Difference Mattes. Das zu isolierende Objekt wird häufig Pixelwerte aufweisen, die ähnliche Helligkeiten wie der Hintergrund aufweisen, was zu Löchern in der Matte führt. Um dem gegenzusteuern, muss die Matte stark hochskaliert werden, was jedoch wiederum zu einer harten Kante um das Objekt führt. Es gibt jedoch bestimmte Anwendungsfälle, in denen die Difference Matte gut funktioniert, daher sollte sie nicht völlig ignoriert werden (Wright, 2010, S. 29).

5 Workflow der K.I gestützten Rotoskopie-Anwendungen

In diesem Kapitel werden die jeweiligen K.I gestützten Rotoskopie-Anwendungen und deren Workflow dargestellt und beschrieben, die im empirischen Teil dieser Arbeit zum Einsatz kommen.

5.1 Adobe After Effects

In Adobe After Effects wurde das **Roto-Pinsel 2** Werkzeug verwendet um die Personen in den verwendeten Videos K.I basiert zu rotoskopieren. Beim herkömmlichen, manuellen Rotoscoping werden häufig animierte Masken verwendet, um ein Element in einem Videoclip zu isolieren, was mühsam und zeitaufwändig sein kann. Mit dem Roto-Pinsel 2 Werkzeug werden die Bewegungen des gewählten Objekts (in diesem Fall Personen) mittels künstlicher Intelligenz verfolgt. Dadurch ist es möglich, auf eine Verfeinerung jedes einzelnen Frames zu verzichten, was sich sehr zeitsparend auswirkt (Adobe Inc., 2023).

Der Arbeitsablauf mit dem Roto-Pinsel 2 Werkzeug gestaltet sich nun folgendermaßen. Im ersten Schritt wird das Roto-Pinsel 2 Werkzeug in der Symbolleiste ausgewählt. Anschließend wird im Kompositionsfenster mit Doppelklick eine vorhandene Ebene ausgewählt, wodurch die Ebene im Bedienfeld „Ebene“ geöffnet wird (Adobe Inc., 2023).

Im folgenden Schritt wird nun ein passender Frame ausgewählt, um mit dem Rotoscoping zu beginnen. Dabei ist darauf zu achten, dass mit einem Frame begonnen wird, in dem das gewählte Objekt nicht von einem anderen Objekt verdeckt wird. Wenn sich beispielsweise eine Person aus einem Frame hinaus und wieder hinein bewegt, sollte ein Frame ausgewählt werden, in dem sich die Person vollständig in der Szene befindet (Adobe Inc., 2023).

Im nächsten Schritt wird das gewählte Objekt mit dem Roto- Pinsel Werkzeug markiert. Abbildung 7 zeigt auf der linken Seite die Markierung des gewählten Objektes durch einen grünen Strich mittels des Roto-Pinsel Werkzeugs. Das zu erfassende Objekt wird anschließend von einem Rahmen in Magenta umrissen. (rechts) (Adobe Inc., 2023).

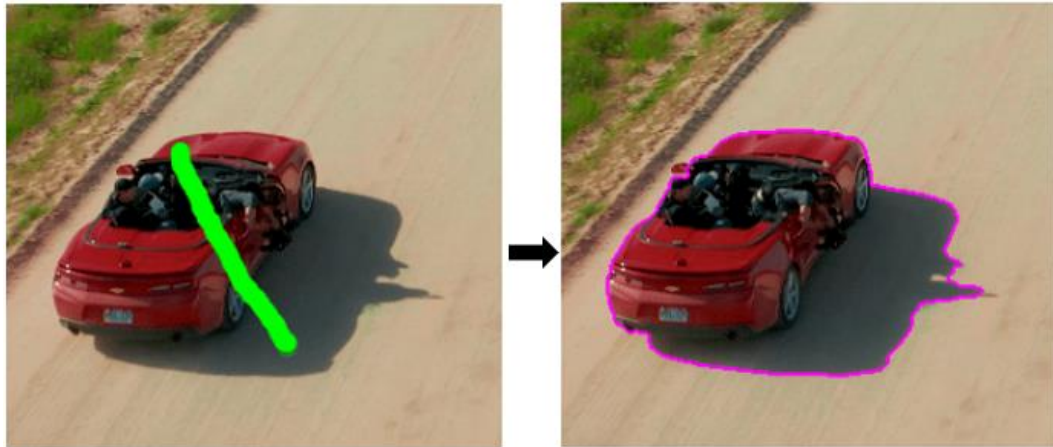


Abbildung 7. After Effects Roto Pinsel 2- Auswahl des Objekts (Adobe Inc., 2023).

Bei der Markierung des Objekts ist darauf zu achten, dass mit einem Strich quer über das Objekt begonnen wird. Darauf soll die Auswahl kontrolliert werden und bei Bedarf Anpassungen und Verfeinerungen vorgenommen werden. Es sollten so wenig Striche wie möglich verwendet werden, um die Auswahl anzupassen. Überdies sollte vermieden werden, eine Kontur um das Objekt zu zeichnen. Stattdessen sollte durch die Mitte gezeichnet und dabei alle Bereiche des Objekts durchlaufen werden, die eine andere Farbe oder Helligkeit aufweisen. Ein Zeichnen über die Ränder sollte ebenfalls vermieden werden, da dies die Auswahl durcheinanderbringen kann (Adobe Inc., 2023).

Sobald das Objekt ausgewählt wurde und ein Rahmen und das Objekt erzeugt wurde, kann die Auswahl im nächsten Schritt verfeinert werden. Dafür werden zusätzliche Striche oder Hintergrundstriche gezeichnet. Mittels Hintergrundstrichen können Teil aus der Auswahl wieder entfernt werden. Dies erfolgt, indem die Alt bzw. Option Taste gedrückt wird, während ein Strich gezeichnet wird. Hintergrundstriche sind rot markiert. Die Magenta-Linie, die das Objekt umgibt, wird nach und nach verfeinert (Adobe Inc., 2023).

Nachdem der Basis-Frame vollständig eingestellt und bearbeitet ist, kann die Propagierung gestartet werden. Durch Drücken der Leertaste wird die Propagierung der Matte vom Basis-Frame zum nächsten Frame ausgelöst. Es ist ebenfalls möglich von Frame zu Frame vorwärts oder rückwärts zu propagieren, um langsamer vorzugehen und die Propagierung genauer zu verfolgen. Falls bei der Propagierung unerwünschte Regionen eingeschlossen werden oder Teile der ursprünglichen Auswahl verloren gehen, kann ein Strich oder Hintergrundstrich im ersten Frame gezeichnet werden, in dem der Fehler auftritt. Dabei sollte ebenfalls darauf geachtet werden, nicht über den Rand der Auswahl hinaus zu zeichnen (Adobe Inc., 2023).

Nachdem die Maske durch alle Frames des Videos propagiert wurde, kann die Propagierung fixiert werden, indem die Schaltfläche „Fixieren“ am unteren Rand des Ebenenfensters geklickt wird. Wenn nach dem Fixieren etwas zu einer Auswahl hinzugefügt oder aus dieser entfernt werden muss, kann die Fixierung der Propagierung mit derselben Schaltfläche aufgehoben werden. Es sollte jedoch grundsätzlich erst fixiert werden, sobald die Maske das Objekt im gesamten Videoclip korrekt umschließt. Das Fixieren setzt die Maske am Ort fest, sodass die Kanten nicht erneut vom Roto-Pinsel propagiert werden müssen (Adobe Inc., 2023).

Das Ergebnis der Propagierung kann im weiteren Verlauf durch die Steuerelemente weiter verfeinert und verbessert werden, beispielsweise mit der Eigenschaft „Kantenrauschen reduzieren“. Überdies kann das Kantenverfeinerungs-Werkzeug für sehr detaillierte Kanten wie Haare verwendet und präziser gesteuert werden. Des Weiteren gibt es Optionen, um Bewegungsunschärfe zu kompensieren und Kantenfarben zu bereinigen. Auch hier ist es möglich Kantenverfeinerungsstriche wieder zu löschen (Adobe Inc., 2023).

Je nach Vorlieben des Anwenders oder je nach Bedarf gibt es mehrere Anzeigeeoptionen beziehungsweise Ansichtsmodi, die gewählt werden können, um die Maske zu bearbeiten und zu verfeinern.

- Alpha: Zeigt den Alphakanal der Ebene an
- Alpharand: Zeigt die Quellebene mit unverändertem Vordergrund und Hintergrund und mit der Segmentierungsgrenze, überlagert als farbiger Umriss
- Alphaüberlagerung: Zeigt die Quellebene mit unverändertem Vordergrund und mit einem einfarbigen Hintergrund

Nachdem die Bearbeitung mit dem Roto- Pinsel 2 Werkzeug abgeschlossen ist, besitzt das ausgewählte Objekt nun einen transparenten Hintergrund und kann vor einen beliebigen neuen Hintergrund platziert werden (Adobe Inc., 2023).

5.2 The Foundry Nuke

In der Compositing Software Nuke von The Foundry wurde der CopyCat node (Knoten) angewendet, um die Personen aus dem verwendeten Videomaterial zu rotoskopieren. Der CopyCat-Node, welcher nur in NukeX und Nuke Studio verfügbar ist, kopiert sequenzspezifische Effekte, wie garbage matting, Beauty-Reparaturen oder Deblurring, von einer kleinen Anzahl von einzelnen Bildern in einer Sequenz und trainiert dann mithilfe von maschinellem Lernen ein Netzwerk, um diesen Effekt auf eine gesamte Sequenz zu replizieren (The Foundry, 2023).

CopyCat nimmt eine beliebige Anzahl von Bildpaaren auf, die als Input- und Ground-Truth-Images bezeichnet werden, welche zusammen das Data Set bilden. Der Datensatz dient dazu ein neuronales Netzwerk mittels maschinellen Lernens zu trainieren, um die Input-Bilder zu manipulieren, und so in weitere Folge die Ground-Truth Images zu replizieren. Wie in Abbildung 8 dargestellt, kann ein Datensatz beispielsweise aus sechs Frames einer Sequenz bestehen, die bearbeitet werden soll und sechs rotoskopierte Masken (The Foundry, 2023).

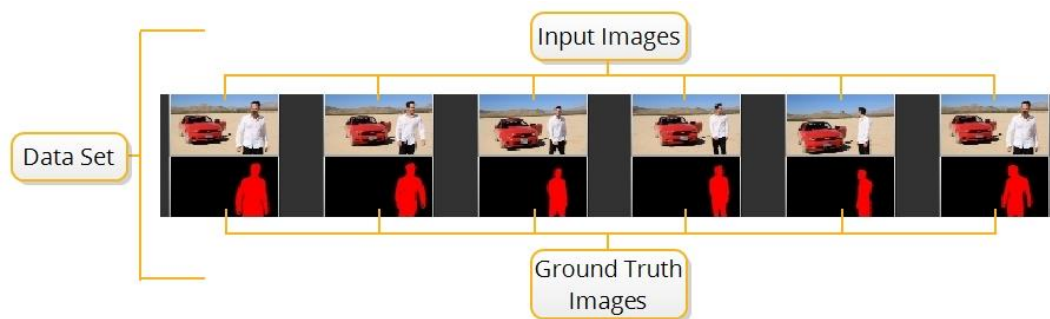


Abbildung 8. Nuke CopyCat dataset (The Foundry, 2023).

Der Workflow, der zur Bearbeitung der Videosequenzen in dieser Arbeit angewendet wurde, wird in den folgenden Absätzen beschrieben.

Im ersten Schritt wurde ein „Read Node“ erstellt, mit dem das jeweilige Video geladen wurde. Danach wurden spezifische Einzelbilder in der Videosequenz gesucht. Dabei konnten mit einer diversen Auswahl an Einzelbildern die besten Ergebnisse erzielt werden. Es wurde darauf geachtet, Einzelbilder zu verwenden, auf welchen die zu rotoskopierenden Personen in möglichst unterschiedlichen Bewegungen und Darstellungen zu sehen sind. Des Weiteren wurden ein FrameHold Node jeweils für jedes selektierte Frame erstellt. (The Foundry, 2023).

Unter jeden FrameHold Node wurde ein Roto Node erstellt, um eine Maske für die Personen zu erstellen. Alle einzelnen Roto Nodes wurden mit einem AppendClip Node verbunden. Dem AppendClip Node wurde ein Shuffle Node nachgeschaltet,

welcher den Alpha Kanal in den Rot-Kanal kopiert. Dem Shuffle Node ist ein Remove Node nachgeschaltet, welcher nur den Rot Kanal behält und alle anderen Kanäle entfernt. Dieser Remove Node wurde anschließend mit dem Ground Truth des CopyCat Nodes verbunden. Ein zweiter Remove Node wurde zwischen dem AppendClip und dem CopyCat Node eingefügt, um die Farbkanäle (Rot, Grün, Blau) zu behalten, und wurde mit dem Input des CopyCat Nodes verbunden (The Foundry, 2023).

Die Inputs des CopyCat Nodes sind wie folgt definiert:

- **Input:** Die Image Sequenz, bevor ein oder mehrere Effekte angewandt wurden
- **Ground Truth:** Die Image Sequenz, nachdem ein oder mehrere Effekte angewandt wurden. Dieser Input beschreibt, was das Netzwerk „zu lernen“ versucht
- **Preview:** Diese Eingabe wird verwendet, um anzuzeigen, wie das Modell funktioniert, wenn es auf einen Frame angewendet wird, der nicht Teil des Datensatzes ist, der zum Trainieren des Modells verwendet wurde. Mit fortschreitendem Training sollte sich das Vorschaubild in Richtung des idealen Ergebnisses bewegen, das durch die Ground Truth definiert ist.

(The Foundry, 2023)

Abbildung 9 zeigt die für diese Arbeit verwendeten Nodes und deren Verbindungen zueinander. Dieselbe Struktur und Anordnung der Nodes konnte bei der Bearbeitung aller in dieser Arbeit verwendeten Videoclips, jeweils leicht angepasst, verwendet werden.

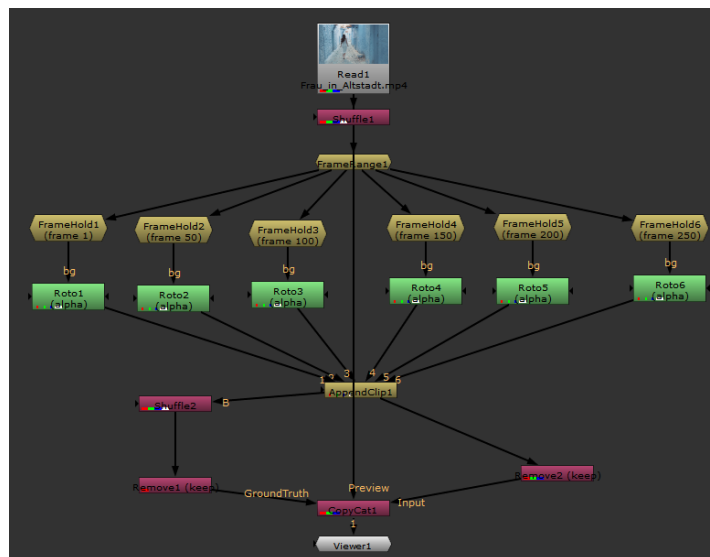


Abbildung 9. Nuke CopyCat Node Graph

Nachdem alle Nodes korrekt verbunden und eingestellt wurden, konnte mit dem Training des Netzwerks begonnen werden. Dafür wurde eine Data Directory definiert, in die die Output files, die während des Training Prozesses generiert wurden, gespeichert wurden.

Die Anzahl der Epochs definiert wie oft CopyCat das gesamte Datenset während des Trainings berechnet. Höhere Werte generieren in der Regel bessere Ergebnisse, erhöhen jedoch die gesamte benötigte Zeit für die Berechnung. In den erweiterten Einstellungen des CopyCat Nodes können weiters die Model Size, Batch Size und Crop Size eingestellt werden.

Model Size: Ermöglicht es, Geschwindigkeit und Speicherverbrauch gegen potenziell bessere Ergebnisse abzuwägen.

Batch Size: Legt die Anzahl der Bildpaare fest, mit denen das Netzwerk in jeder Epoche trainiert wird und wird verwendet, um die Anzahl der Gesamtschritte zu berechnen, die zum Abschließen des Trainings erforderlich sind. Die Batch Size wird standardmäßig automatisch anhand des verfügbaren GPU-Speichers berechnet.

Crop Size: Definiert die Größe der zufälligen Ausschnitte aus den Bildpaaren des Datensatzes. Größere Werte führen im Allgemeinen zu genaueren Ergebnissen, gehen jedoch zu Lasten der Verarbeitungszeit und des Speichers. Kleinere Werte erfordern möglicherweise mehr Wiederholungen, um ein gutes Ergebnis zu erzielen.

Diese Einstellungen erfordern jedoch einen Kompromiss zwischen Trainings Zeit und Qualität der Ergebnisse (The Foundry, 2023).

Nach Einstellung aller Parameter kann mit dem Start des Trainings begonnen werden. Im Graph Tab des CopyCat Nodes kann der Trainings Fortschritt, dargestellt als Diagramm, beobachtet werden.

Während Trainings ist ein Raster, sogenannte Contact Sheets, im Viewer zu sehen. Mit fortschreitendem Training verändern sich die Bilder von zufälligen garbage Mattes in Richtung groundtruth Bilder. Contact Sheets ermöglichen eine Übersicht, ob die Daten richtig ausgerichtet sind und wie sich das Netzwerk bei bestimmten Frames verhält (The Foundry, 2023).

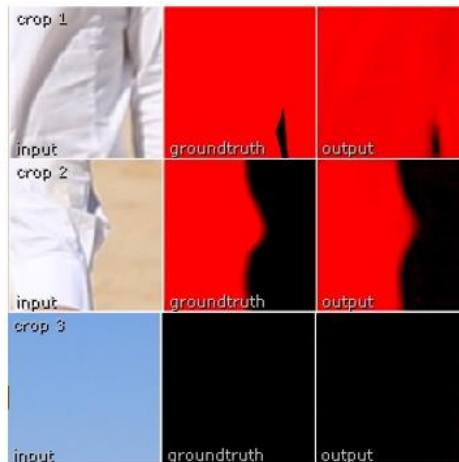


Abbildung 10. Nuke CopyCat Contact Sheets (The Foundry, 2023).

Abbildung 10 zeigt Contact Sheets, die sich in drei Bereiche gliedern. Input zeigt eine zufällige Stichprobe eines Frames aus der Quell-Sequenz, auf die das Training angewendet wird. Ground truth zeigt denselben Bereich, jedoch von der Ground truth Seite der Daten und was das Netzwerk zu replizieren lernt. Output zeigt denselben Bereich mit dem bisherigen Fortschritt, angewandt auf das Quell-Bild (The Foundry, 2023).

Nach Abschluss des Trainings kann das trainierte Netzwerk auf die gesamte Sequenz angewendet werden. Dafür wird ein Inference Node erstellt. Der Inference Node wird mit dem Read Node der Quell Sequenz verbunden, um die Ergebnisse auf dem gesamten Footage zu kontrollieren. Durch Abspielen der Sequenz kann kontrolliert werden, wie die Ergebnisse auf den unterschiedlichen Frames aussehen. Falls Artefakte außerhalb der Maske zurückbleiben, können die Ergebnisse weiterbearbeitet und optimiert werden.

Wenn die Ergebnisse zufriedenstellend sind, kann der maskierte Farbkanal wieder mit der Quell-Sequenz vermischt werden und die finalen Ergebnisse gerendert werden (The Foundry, 2023).

5.3 Runway ML

In der Online Software Anwendung Runway wurde die Funktion Green Screen angewendet, um die Videoclips für diese Arbeit zu bearbeiten. Runway ist eine Online-Software, mit der es möglich ist, mehrere Machine Learning Modelle anzuwenden, ohne selbst programmieren zu müssen. Mit der Funktion Green Screen ist es möglich spezifische Objekte oder Regionen in wenigen Schritten aus Videomaterial zu extrahieren und in weiterer Folge beispielsweise Hintergründe auszutauschen, Objekte zu duplizieren oder andere Effekte anzuwenden. Die Anwendung ist vollständig von einem Browser aus zu bedienen (runwayml, 2023).

Neben Green Screen bietet Runway noch folgende weitere Funktionen an:

- Text to video: Ein A.I System, das es ermöglicht Videos per Texteingabe zu generieren und zu bearbeiten
- Inpaint Content Aware Fill: Ermöglicht Objekte inhaltsbasiert aus Videomaterial zu entfernen
- One-Click Motion Tracking: Automatische Bewegungsverfolgung eines Objekts, das lediglich mit einem Klick ausgewählt werden muss.

(runwayml, 2023)

Runway wird in vier unterschiedlichen Preis-Kategorien mit jeweils mehr oder weniger Anwendungsmöglichkeiten angeboten. Für diese Arbeit wurde das „Team“ Preismodell mit monatlicher Abrechnung gewählt, welches \$28 (26 €) pro Monat kostet und unter anderem die Möglichkeiten bietet, PNG-Sequenzen in einer Auflösung in bis zu 4K zu exportieren (runwayml, 2023).

Bei der Anwendung der Green Screen Funktion wird im ersten Schritt ein neues Projekt erstellt. Ein Videoclip muss anschließend hochgeladen werden, um bearbeitet werden zu können. Ein Upload ist per „Drag and Drop“ in das Browserfenster möglich. Es können auch mehrere Clips im Vorhinein hochgeladen werden, um diese zu einem späteren Zeitpunkt zu bearbeiten.

Um bestmögliche Ergebnisse zu erzielen, empfiehlt Runway Videomaterial zu verwenden, in dem ein gut erkennbares Objekt zu sehen ist und sehr ruhige oder keine Kamera-Bewegungen angewendet wurden. Es wird weiters empfohlen Videomaterial zu verwenden, in dem keine hektischen oder unregelmäßige Hintergründe in niedriger Qualität zu sehen sind, oder sich ein Objekt mehrmals aus dem Bild hinaus und herein bewegt (runwayml, 2023).

Um die vorgestellten Anwendungen bestmöglich zu vergleichen und vor allem deren Grenzen aufzuzeigen, wurde für diese Arbeit jedoch unterschiedliches Videomaterial gewählt, das nicht ausschließlich den genannten Empfehlungen von Runway entspricht.

Nach Hochladen des Videoclips kann mit der Bearbeitung begonnen werden, indem mit einem einzelnen Klick auf das gewählte Objekt eine Selektion vorgenommen wird. Es wird empfohlen möglichst in die Mitte des Objekts zu klicken. Wenn weitere Bereiche des Objekts oder mehrere Objekte selektiert werden sollen, müssen diese ebenfalls jeweils mit einem einzelnen Klick ausgewählt werden. Wenn unerwünschte Bereiche in die Auswahl mitaufgenommen wurden, können diese mit der „Exclude“ Funktion de-selektiert werden. Nach Erzeugung eines ersten repräsentativen Keyframes dient dieser als Referenz für den restlichen Videoclip (runwayml, 2023).

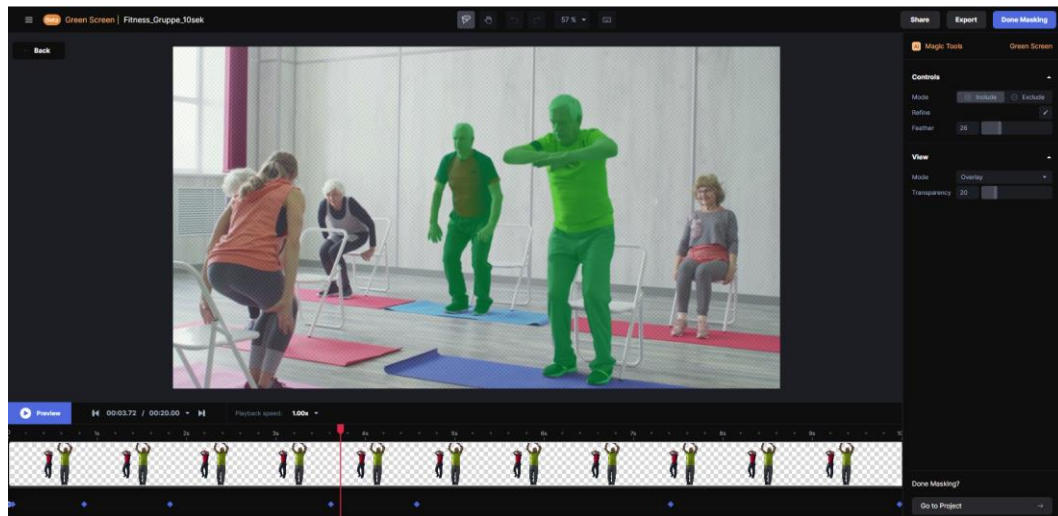


Abbildung 11. Runway Bearbeitung

Sobald eine zufriedenstellende Maske erzeugt wurde, kann der Videoclip abgespielt und beobachtet werden, wie die Bereiche des Objekts zu anderen Zeitpunkten automatisch selektiert und maskiert werden. Falls notwendig, kann die Maske mit weiteren Keyframes verbessert werden (Abbildung 11), wobei möglichst wenige Keyframes erzeugt werden sollten. Überdies ist es möglich mit einem Brush-Tool manuell über detaillierte Bereiche zu malen und so Verbesserungen vorzunehmen. Diese Retuschen werden auf die darauffolgenden Frames des Videoclips übertragen. Zur besseren Ansicht während der Bearbeitung stehen mehrere Ansichts-Modi wie Overlay, Preview oder Alpha Channel zur Verfügung (runwayml, 2023).

Nach der Bearbeitung und sobald ein zufriedenstellendes Ergebnis erreicht wurde, stehen mehrere Optionen bezüglich Export des Videoclips zur Verfügung. Davor können bei Bedarf noch weitere Effekte angewendet werden. Es stehen mehrere Formate und Codecs wie MP4, ProRes oder PNG zur Verfügung. Die Projekte stehen nach dem Export weiterhin zur Bearbeitung online zur Verfügung (runwayml, 2023).

Für diese Arbeit wurde das bearbeitete Videomaterial jeweils als PNG-Sequenz mit transparentem Hintergrund exportiert.

6 Empirischer Teil

6.1 Aufgabenstellung an die Programme

Durch die Literaturrecherche konnte bereits ein Einblick über die Anwendung künstlicher Intelligenz im Alltag, in der Arbeitswelt und in der digitalen Content-Produktion beziehungsweise im Kreativbereich gegeben werden. Um die Hypothesen zu überprüfen und um die restlichen Forschungsfragen beantworten zu können, wurde ein empirischer Teil ausgearbeitet.

Für den empirischen, beziehungsweise praktischen Teil dieser Arbeit wurde Videomaterial verwendet, in dem sich Personen (reale Menschen) in unterschiedlichen Situationen, vor unterschiedlichen Hintergründen und unterschiedlichen Kameraeinstellungen befinden. Das Videomaterial stammt von der Website „<https://www.pexels.com/de-de/>“ und ist lizenztechnisch kostenlos und ohne Namensnennung zu verwenden. Es wurden insgesamt sechs unterschiedliche Videos verwendet, in denen eine oder mehrere Personen aus unterschiedlichen Winkeln und Entfernungen aufgenommen wurden.

Das gleiche Videomaterial wurde jeweils mit den drei unterschiedlichen Programmen (Adobe After Effects, The Foundry Nuke, Runway ML) bearbeitet, wobei die beinhalteten Personen, K.I. unterstützt, rotoskopiert wurden. Die Videosequenzen lagen jeweils als .mp4 Dateien, mit einer Bildrate von 25 fps und einer Auflösung von 1920 x 1080 Pixel vor. Eine kurze bearbeitete Videosequenz in einer Länge von 10 Sekunden wurde jeweils als PNG-Sequenz exportiert.

Im ersten Video ist eine Frau in einem hellblauen Kleid von hinten zu sehen, die durch einen Gang in einer Altstadt läuft, der links und rechts von alten Gebäuden mit blauer Fassade begrenzt ist. Der Körper der Frau ist vollständig aus einer Entfernung von ca. drei Metern zu sehen. Sie besitzt langen, lockiges braunes Haar, das durch die Laufbewegungen auf und ab schwingt. Die Aufnahme verfolgt die Frau in einer durchgehenden Entfernung von ca. drei Metern. Die Frau wurde jeweils in den verwendeten Programmen K.I.-basiert rotoskopiert.

Im zweiten Video ist eine Frau mit blondem, schulterlangem, dichtem Haar und Brille von der rechten Seite bis zu ihren Unterschenkeln zu sehen, die auf einer hellbraunen, hölzernen Bank sitzt. Sie trägt dunkle Kleidung und hält ein Buch in ihren Händen und neben ihr liegt ebenfalls ein Buch auf der Bank. Hinter der Frau befindet sich eine graue Beton-Mauer. Im Videoclip blättert die Frau mehrere Seiten des Buches durch und legt dieses anschließend seitlich neben sich auf die Bank. Vor der Frau sitzt eine weitere Person auf der Bank, die nur unscharf und in

Nahaufnahme zu sehen ist. Diese Person hält ein Handy in der Hand, welches die Frau zeitweise geringfügig verdeckt. Die Aufnahme ist starr. Die blonde Frau wurde jeweils inklusive ihrer beiden Bücher rotoskopiert.

Das dritte Video zeigt einen älteren Mann mit grauem kurzem Haar und Bart, der mit dunkler Kleidung und ohne Schuhe an einem Strand entlangläuft. Der Mann ist Anfangs aus ca. 3 Metern Entfernung vollständig zu sehen; die Kamera bewegt sich jedoch gegen Ende des Clips auf ca. einen Meter in Richtung des Mannes. Die Kamera verfolgt den Mann, sodass dieser immer im Fokus bleibt. Am Boden vor dem Mann sind dessen Spiegelungen im Sand zu sehen. Im Hintergrund ist ein bewölkter grau-blauer Himmel und ein Meer mit starkem Wellengang zu sehen.

Im vierten Video ist eine Frau mit lockigem, schulterlangem braunem Haar in Nahaufnahme bis zum Brustbereich zu sehen. Die Haare bewegen sich im Wind. Die Frau trägt weiße Kleidung und helle Kopfhörer, an denen zwei weiße Kabel jeweils an den Enden herunterhängen. Im Hintergrund ist ein klarer, blauer Himmel und ein Meer samt Strand in Unschärfe zu sehen. Die Aufnahme verfolgt die Frau, die sich durchgehend rhythmisch bewegt und eine Hand an die Kopfhörer hält.

Während die ersten vier Videoclips Personen in hellen Umgebungen zeigen, ist im fünften Video eine dunkelhäutige Person mit kurzen Haaren in einem abgedunkelten Zimmer zu sehen. Das Licht in der Aufnahme besitzt eine gelblich-braune Färbung, die sich auf das T-Shirt der abgebildeten Person überträgt. Die Aufnahme ist starr und zeigt die auf einem gepolsterten Sessel sitzende Person bis zu deren Knien. Zu Beginn des Videoclips sitzt die Person aufrecht und es ist nur deren Silhouette zu sehen. Danach lehnt sich die Person zurück und das Licht fällt langsam auf deren Oberkörper. Zusätzlich bewegt die Person ihre linke Hand auf ihren Kopf zu und hält diesen anschließend bis zum Ende der Aufnahme.

Das sechste und letzte Video zeigt eine Gruppe älterer Menschen in einem geschlossenen, hellen Raum, die von einer Trainerin vorgezeigte Fitness-Übungen ausführt. Insgesamt sind samt Trainerin sechs Personen aus unterschiedlichen Entfernungen jeweils vollständig zu sehen. Zwei männliche Personen in der Mitte der starren Aufnahme wurden jeweils rotoskopiert. Die Personen stehen auf roten und blauen Gummi-Matten, die auf einem grauen Fußboden liegen. Im Hintergrund ist eine graue Wand zu sehen. Die beiden rotoskopierten Personen tragen helle, rote und gelbe Oberteile und graue Hosen sowie Sportschuhe. Im Laufe des Videoclips sind die Personen zuerst stehend, mit hochgestreckten Armen zu sehen. Anschließend setzten sich die Personen auf weiße Klapp-Sessel, die jeweils hinter ihnen stehen und bewegen die Arme nach unten. Der rechte Ellbogen der vorderen Person verdeckt teilweise den linken Unterarm der hinten stehenden Person.

6.2 Objektive Analyse der Anwendungen

Die in dieser Arbeit eingesetzten Anwendungen werden im folgenden Kapitel objektiv analysiert. Dazu wurden während der Bearbeitung die Anzahl der Arbeitsschritte, der Zeitaufwand für den Anwender und die Renderzeit erfasst und protokolliert. Für die Bearbeitung der Videoclips in den drei verwendeten Anwendungen wurde ein Computersystem mit folgenden Eigenschaften verwendet:

- Betriebssystem: Windows 10 Pro, 64 Bit Version
- Installierter Arbeitsspeicher: 64 GB RAM
- Prozessor: Intel Core i7-11700K, 8x 3.60GHz
- Grafikkarte: PNY NVIDIA Quadro P2000 5GB GDDR5

Hierbei ist zu erwähnen, dass sich die Renderzeit auf das verwendete Computersystem und dessen verfügbare Leistung bezieht.

6.2.1 Anzahl der Arbeitsschritte

Die Arbeitsschritte waren einerseits in den Anwendungen Adobe After Effects und Runway ML sehr ähnlich und gut miteinander vergleichbar. Die Anwendung The Foundry Nuke (CopyCat) unterschied sich punkto Arbeitsschritte jedoch von den anderen beiden Anwendungen. Der spezifische Workflow der jeweiligen Anwendungen wurde bereits in Kapitel 5 ausführlich erläutert.

Worin sich die drei Anwendungen jedoch glichen, war die Auswahl eines oder mehrerer spezifischer Frames aus den Videoclips, der/die als Ausgangspunkt für die weitere Bearbeitung dienten. In Adobe After Effects und Runway ML wurde in diesem einen ausgewählten Frame das zu rotoskopierende Objekt mit einem oder wenigen Klicks beziehungsweise Strichen ausgewählt und selektiert. In The Foundry Nuke müssen zuerst mehrere Einzelbilder ausgewählt werden, in denen folglich im zweiten Schritt jeweils eine Maske um das ausgewählte Objekt (Person/en) gezeichnet wird.

In Adobe After Effects und Runway ML wird im zweiten Schritt der Clip abgespielt und dabei die Auswahl verfeinert und verbessert.

Im dritten Schritt wird in Adobe After Effects die Maske propagiert. In Runway ML werden im dritten und letzten Schritt zusätzliche Retuschen und Verbesserungen vorgenommen. Ein vierter Schritt fällt nicht an.

Der dritte Schritt in The Foundry Nuke bezeichnet das Training des erstellten Netzwerkes, was die deutlich längste Zeit in Anspruch nimmt (siehe Abschnitt Renderzeit). Nach Ende des Trainings kann im vierten und letzten Schritt das trainierte Netzwerk auf die gesamte Sequenz angewendet werden.

In Adobe After Effects wird im vierten und letzten Schritt die Maske fixiert, wodurch die Kanten nicht erneut propagiert werden müssen.

Somit ergibt sich punkto Arbeitsschritte folgende Aufteilung:

- Adobe After Effects: 4 Arbeitsschritte
- The Foundry Nuke (CopyCat): 4 Arbeitsschritte
- Runway ML: 3 Arbeitsschritte

6.2.2 Zeitaufwand

Während der Bearbeitung der Videos in den drei verwendeten Programmen wurde die Bearbeitungszeit erfasst und protokolliert. Die folgende Abbildung veranschaulicht die Bearbeitungsdauer für jedes Video in Bezug auf das verwendete Programm.

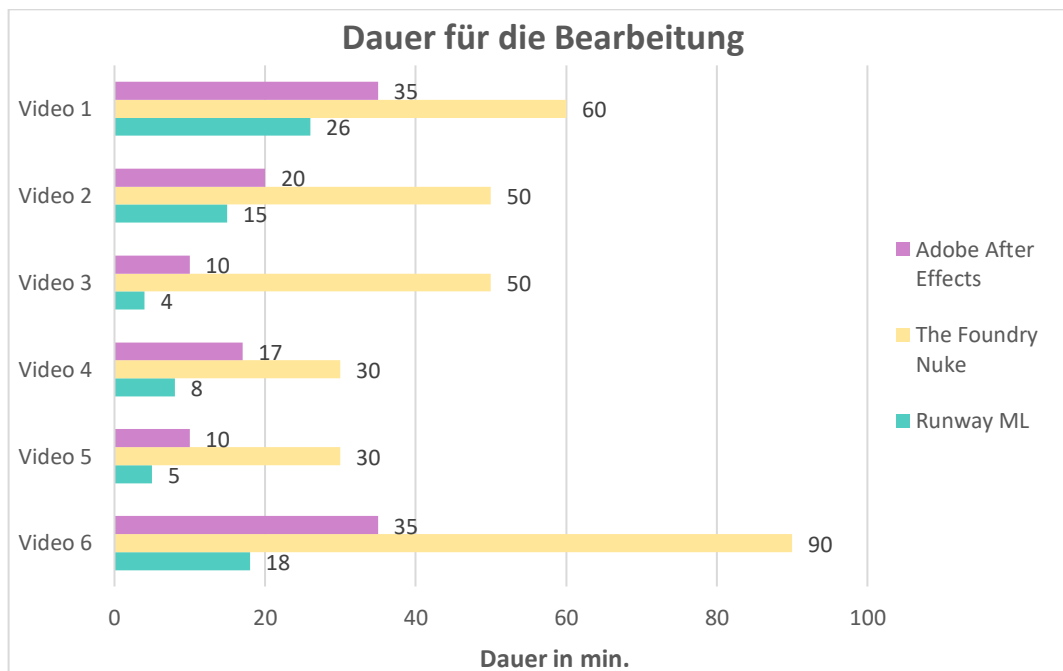


Abbildung 12. Dauer für die Bearbeitung

Die Bearbeitung in The Foundry Nuke nahm jeweils die meiste Arbeitszeit in Anspruch. Speziell das sechste Video, in dem zwei Personen rotoskopiert wurden, nahm in Summe 90 Minuten an Bearbeitungszeit in Anspruch, da in The Foundry Nuke für alle sechs ausgewählten Keyframes jeweils zwei Masken um die beiden Personen gezeichnet werden mussten. In den restlichen Videos ergaben sich in The Foundry Nuke Bearbeitungszeiten zwischen 30 und 50 Minuten.

In Adobe After Effects konnten die Videos wesentlich schneller bearbeitet werden als in The Foundry Nuke, da hier keine vollständigen Masken um die Personen

gezeichnet werden mussten, sondern lediglich mehrere Striche, um die Personen zu selektieren. Auch hier dauerte die Bearbeitung des sechsten Videos am längsten. Das erste Video nahm ebenfalls, wie das sechste Video, 35 Minuten an Bearbeitungszeit in Anspruch. Im ersten Video ist eine schnell laufende Frau mit langem, lockigem Haar zu sehen, dessen Selektierung und Kanten-Verfeinerung mehr Zeit in Anspruch nahm, als in den restlichen Videos.

Die Bearbeitung in Runway ML hatte in sämtlichen Videos die deutlich kürzeste Bearbeitungszeit. Video 3 und Video 5 nahmen lediglich 4 beziehungsweise 5 Minuten in Anspruch. Die Videos 1,2 und 6 konnten in einer Dauer zwischen 15 und 26 Minuten bearbeitet werden.

Somit ergeben sich folgende durchschnittliche Bearbeitungszeiten:

- Adobe After Effects: 21 Minuten
- The Foundry Nuke: 51 Minuten
- Runway ML: 12 Minuten

Es sei hier erwähnt, dass sich die Bearbeitungszeiten auf die Aufzeichnungen und Kenntnisse des Verfassers dieser Arbeit, sowie auf das verwendete Computersystem beziehen.

6.2.3 Renderzeit

Im Anschluss an die Bearbeitung wurde in jeder Anwendung der Renderprozess gestartet und die jeweilige dafür benötigte Dauer gemessen und dokumentiert. Wie bereits erwähnt, bezieht sich die Dauer für den Renderprozess auf die Leistung des verwendeten Computersystems. Die Framerate betrug in allen sechs bearbeiteten Videos 25 fps. In Adobe After Effects wurde der „Fixieren“ Prozess herangezogen und protokolliert. In The Foundry Nuke wurde jeweils mit 10.000 Epochs gerendert und die dafür benötigte Zeit gemessen. In der Anwendung Runway ML wurde der „Exporting“ Prozess für den Vergleich herangezogen.

Die jeweiligen Renderprozesse der Anwendungen unterscheiden sich stark voneinander, was einen direkten Vergleich schwierig macht und sich auch in den jeweiligen Renderzeiten widerspiegelt. In Tabelle 1 sind die jeweiligen Renderzeiten aufgelistet und der jeweilige Durchschnitt pro Anwendung.

	Adobe After Effects	The Foundry Nuke	Runway ML
	Einfrieren/Fixieren	10.000 Epochs	Exporting
	hh:mm:ss	hh:mm:ss	hh:mm:ss
Video 1	00:06:00	04:30:00	00:06:00
Video 2	00:05:00	03:45:00	00:08:00
Video 3	00:03:00	03:00:00	00:07:00
Video 4	00:05:00	03:00:00	00:09:00
Video 5	00:03:00	03:45:00	00:08:00
Video 6	00:05:00	03:00:00	00:09:00
Durchschnitt:	00:04:30	03:30:00	00:07:50

Tabelle 1. Renderzeit.

Es ist ersichtlich, dass in The Foundry Nuke mit Anwendung der CopyCat Funktion wesentlich längere Renderzeiten notwendig waren, als in den anderen beiden Programmen. Mit einer höheren Anzahl an Epochs, die eine deutlich höhere Renderzeit mit sich gebracht hätten, konnten in weiteren Versuchen keine optisch besseren Ergebnisse erzielt werden. Video 1 benötigte mit 4 Stunden und 30 Minuten die längste Renderzeit. Die anderen Videos benötigten zwischen 3 Stunden und 3 Stunden 45 Minuten. Somit ergab sich in The Foundry Nuke eine durchschnittliche Renderzeit von 3 Stunden und 30 Minuten.

Die Renderprozesse in Adobe After Effects und Runway ML lassen sich zeitlich besser miteinander vergleichen. In Adobe After Effects ergab sich eine durchschnittliche Renderzeit von 4 Minuten und 30 Sekunden. In Runway ML war

eine durchschnittliche Renderzeit von 7 Minuten und 50 Sekunden notwendig. So wie sich die Arbeitsschritte in The Foundry Nuke von jenen in Adobe After Effects und Runway ML unterscheiden, so unterscheiden sich auch die Renderzeiten stark voneinander. Ein direkter Zusammenhang mit einer verlängerten Renderzeit und einem daraus resultierenden optisch hochwertigeren Ergebnis konnte nicht festgestellt werden.

6.3 Optische Qualitäten der Ergebnisse

Um die optischen Qualitäten der der Ergebnisse der Umfrage zu vergleichen, wurde eine quantitative Analyse als Methode herangezogen. Dafür wurde ein Fragebogen beziehungsweise eine Umfrage als Mittel gewählt. Der Fragebogen wurde mit der Umfrageverwaltungssoftware „Google Forms“ erstellt. Der Fragebogen wurde anschließend über mehrere Verteiler an Studienkollegen und Bekannte übermittelt und in Online Foren für empirische Umfragen gestellt.

Mit Google Forms können benutzerdefinierte Fragebögen erstellt werden, bei denen die Antwortmöglichkeiten im Voraus festgelegt werden können. Es gibt verschiedene Optionen zur Auswahl, einschließlich Single-Choice, Multiple-Choice und offener Antworten. Darüber hinaus werden die gesammelten Antworten automatisch in die Antwortgrafiken integriert und aktualisiert (Google Inc., 2023).

6.3.1 Analyse der demoskopischen Fragen

Zum Stichtag 04.04.2023 haben insgesamt 81 Teilnehmer und Teilnehmerinnen an der Umfrage teilgenommen.

Die Auswertung bezüglich der Geschlechter der Teilnehmer und Teilnehmerinnen ergibt folgende Zuordnung:

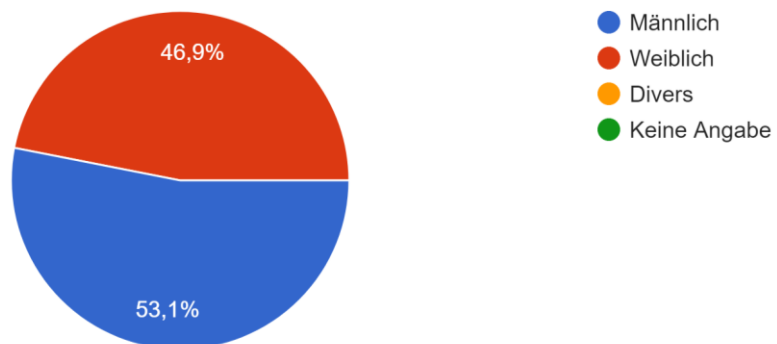


Abbildung 13. Geschlecht der Teilnehmer und Teilnehmerinnen

Abbildung 13 zeigt, dass die Aufteilung bezüglich des Geschlechts der Teilnehmer und Teilnehmerinnen relativ ausgewogen ist. Der Anteil an männlichen Teilnehmern beträgt 53,1 %. Die anderen 46,9 % gaben an, dem weiblichen Geschlecht zugehörig zu sein. Des Weiteren gab es die Möglichkeit, dass Geschlecht mit „divers“ oder gar nicht anzugeben, was jedoch von keiner Person in Anspruch genommen wurde. Da die Anzahl von Teilnehmern und Teilnehmerinnen mit 81 vergleichsweise gering ist, kann hinsichtlich des Geschlechts kein statistisch aussagekräftiger Schluss gezogen werden.

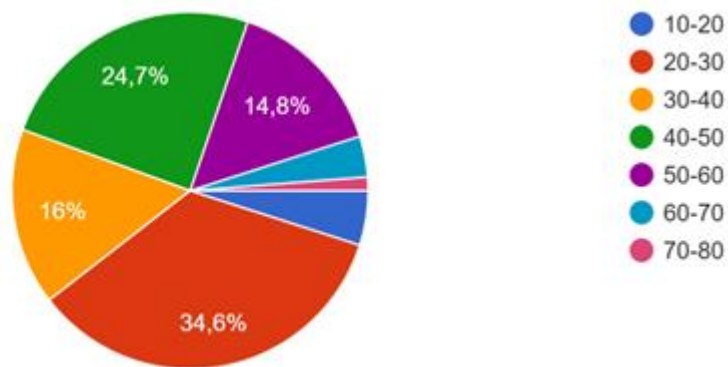


Abbildung 14. Altersgruppen der Teilnehmer und Teilnehmerinnen.

In Abbildung 14 sind die Altersgruppen der Teilnehmer und Teilnehmerinnen dargestellt. Den größten Anteil mit 34,6 % macht jene Gruppe der 20–30-jährigen Teilnehmer und Teilnehmerinnen aus. Die Gruppe der 40-50-jährigen Personen macht den zweit-größten Anteil aus, gefolgt von den 30-40-jährigen und den 50-60-jährigen. Die restlichen Altersgruppen machen, wie in der Abbildung dargestellt, den geringsten Anteil aus. Die Verteilung der Altersgruppen ist somit relativ ausgewogen, lässt jedoch wie bereits erwähnt, aufgrund der eher geringen Anzahl an Teilnehmern und Teilnehmerinnen, keinen statistisch brauchbaren Rückschluss zu.

6.3.2 Analyse der allgemeinen Fragen

Zusätzlich zu den demografischen Fragen wurden allgemeine Fragen gestellt, um Informationen über die Vorkenntnisse der Teilnehmer und Teilnehmerinnen im Bereich der Bild- und Video-Bearbeitung sowie darüber zu erhalten, auf welchen Geräten die Videos angesehen wurden.

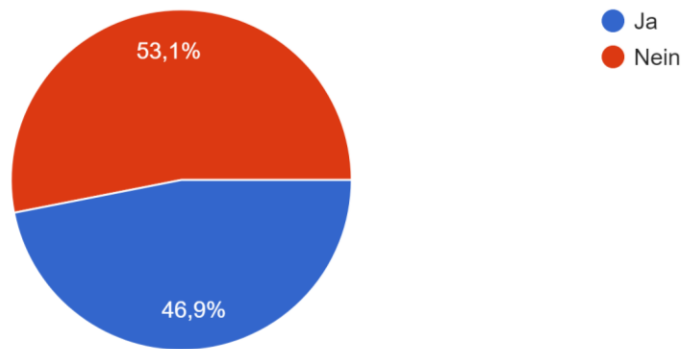


Abbildung 15. Vorkenntnisse der Teilnehmer und Teilnehmerinnen im Bereich Bild- und Video Bearbeitung.

Die Verteilung bezüglich der Vorkenntnisse der Teilnehmer und Teilnehmerinnen im Bereich Bild- und Video Bearbeitung ist ebenfalls relativ ausgewogen. Mit 53,1% gaben etwas mehr als die Hälfte der teilnehmenden Personen an, keine Vorkenntnisse in besagtem Bereich zu besitzen. Die anderen 46,9 % besitzen jedoch Kenntnisse in diesem Bereich.

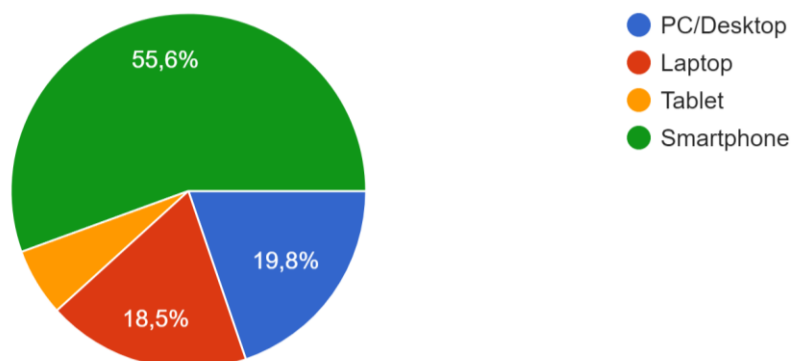


Abbildung 16. Verwendete Geräte zur Betrachtung der Videos.

Am Ende des allgemeinen Abschnittes in der Online Umfrage wurden die Teilnehmer und Teilnehmerinnen darum gebeten anzugeben, auf welchem Gerät sie die Videos betrachten und die Umfrage beantworten. Etwas mehr als die Hälfte der Personen (55,6%) sahen die Videos auf einem Smartphone. 19,8 % verwendeten dafür einen PC bzw. Desktop. Ein Laptop wurde von 18,5 % der Personen verwendet. Die restlichen 6,2 % der Personen verteilen sich auf die Benutzer/innen eines Tablets.

6.3.3 Optische Analyse des ersten Videos

Wie bereits in Kapitel 6.1 „Aufgabenstellung an die Programme“ erwähnt, ist im ersten Video eine Frau mit braunem, lockigem und langem Haar in einem hellblauen Kleid von hinten zu sehen, die durch einen Gang in einer Altstadt läuft, der links und rechts von alten Gebäuden mit blauer Fassade begrenzt ist. Abbildung 17 zeigt auf der linken Seite einen Frame des originalen Videos und auf der rechten Seite denselben Frame nach der Bearbeitung mit den drei verwendeten Programmen. Die Darstellung auf der rechten Seite entspricht der in der Umfrage verwendeten. Die drei Ergebnisse wurden nebeneinander platziert, um einen optimalen Vergleich zu ermöglichen. Unten auf der Abbildung befindet sich eine Nummerierung der Videos. In der Umfrage wurde jedoch nicht erwähnt, welches Ergebnis mit welchem Programm erzielt wurde, um die Neutralität der Umfrage zu wahren.



Abbildung 17. Video 1: Vor und nach Bearbeitung.

Ergebnis 1.1 wurde mit Adobe After Effects, Ergebnis 1.2 mit The Foundry Nuke/Copy Cat und Ergebnis 1.3 mit Runway ML erzielt.

In der Umfrage ergab sich auf die Frage, welches Ergebnis optisch am Besten aussieht folgendes Ergebnis: Die deutliche Mehrheit mit 43,2 % empfand Ergebnis 1.3 als am besten gelungen. Ergebnis 1.2 wurde von 29,6 % der Teilnehmer/innen ausgewählt, gefolgt von Ergebnis 1.1. Jedoch konnten 12,3 % der Personen keinen Unterschied in den drei Ergebnissen erkennen. Die Aufteilung ist in Abbildung 18 ersichtlich. Die Frage „Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?“ beantworteten 60,5 % mit Ja; die anderen 39,5 % konnten jedoch keine Fehler erkennen. Somit wurden von der Mehrheit Bildfehler in den drei Ergebnissen gefunden werden. 42 Teilnehmer/innen konnten in Ergebnis 1.2 Fehler erkennen. In Ergebnis 1.1 konnten noch 33 Teilnehmer/innen Fehler erkennen. In Ergebnis 1.3 nur 30 Personen.

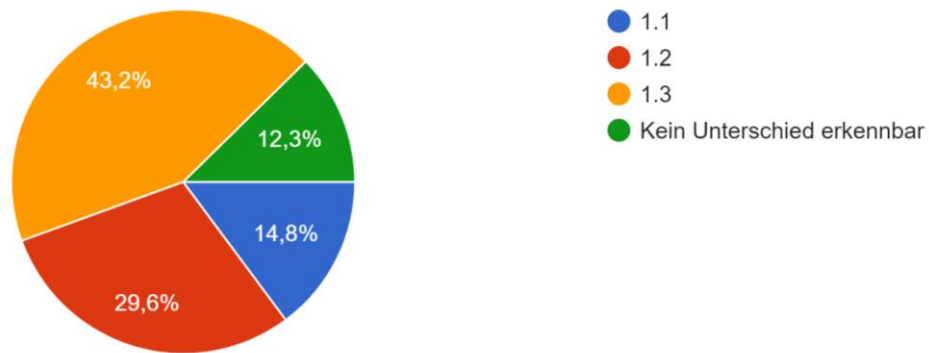


Abbildung 18. Video 1: Ergebnis des optischen Vergleichs.

Außerdem bot die Umfrage die Möglichkeit, in kurzen Antworten zu beschreiben, welche Bildfehler erkannt wurden und wo und wann diese in den Videos zu sehen waren.

Am häufigsten wurde erwähnt, dass sich Bildfehler/Artefakte bei den Haaren beziehungsweise Locken befinden. Es wurde angegeben, dass die Haare teilweise verschwinden oder unscharf werden.

Neben den Haaren wurden Bildfehler im Bereich des Kleides erkannt. Dabei wurde erwähnt, dass am unteren Ende des Kleides zeitweise „Flackern“ zu sehen ist, wenn sich die Dame im Video bewegt. Auch im Bereich der Füße konnten viele Teilnehmer und Teilnehmerinnen einige Artefakte erkennen. Dabei wurde mehrmals erwähnt, dass die Füße beim Aufsetzen/Auftreten am Boden teilweise verschwinden. Bei den Armen konnten Fehler zu jenen Zeitpunkten erkannt werden, wenn sie sich vor oder hinter den Körper der Dame bewegen.

Neben den bereits erwähnten Bild-Fehlern konnten jedoch auch 39,5 % der Teilnehmer/innen keine Fehler in den Ergebnissen erkennen, was mehr als einem Drittel der teilnehmenden Personen entspricht.

Aus den Ergebnissen kann somit abgeleitet werden, dass die drei verwendeten K.I. basierten Rotoskopie-Anwendungen bei schnell bewegten Personen, die noch dazu lockige, bewegte Haare besitzen, eher unzureichende Ergebnisse liefern, die eine weitere Bearbeitung erfordern.

6.3.4 Optische Analyse des zweiten Videos

Das zweite Video zeigt eine Frau mit blondem, schulterlangem, dichtem Haar und einer Brille, die auf einer hellbraunen, hölzernen Bank sitzt. Die Kamera erfasst die Frau von der rechten Seite bis zu ihren Unterschenkeln. Sie trägt dunkle Kleidung und hält ein Buch in ihren Händen, während ein weiteres Buch auf der Bank neben ihr liegt. Hinter der Frau befindet sich eine graue Betonmauer. Im Verlauf des Videoclips blättert die Frau mehrere Seiten des Buches durch und legt es anschließend seitlich neben sich auf die Bank. Vor der Frau sitzt eine weitere unscharfe Person in Nahaufnahme, die ein Handy in der Hand hält und zeitweise die Frau geringfügig verdeckt. Die Aufnahme ist statisch. Die blonde Frau wurde zusammen mit ihren beiden Büchern rotoskopiert.



Abbildung 19. Video 2: Vor und nach Bearbeitung.

Die optischen Ergebnisse des zweiten Videos ergaben in der Umfrage eine relativ ausgeglichene Verteilung.

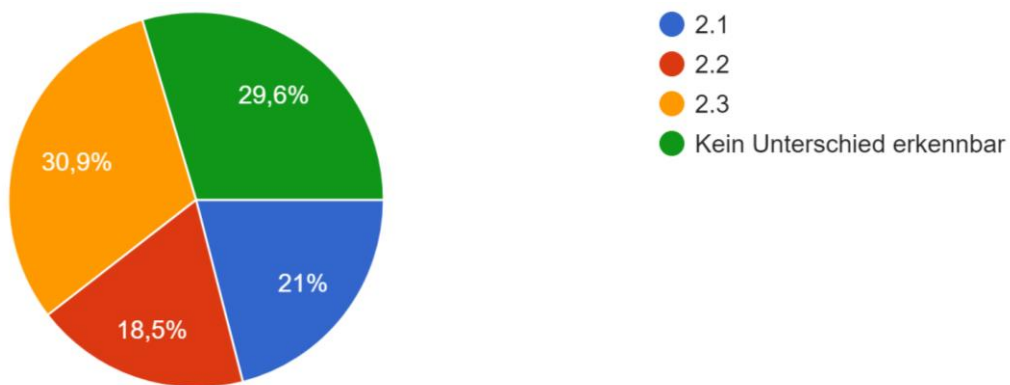


Abbildung 20. Video 2: Ergebnis des optischen Vergleichs.

Wie in Abbildung 20 ersichtlich ist, sind die Antworten, bezüglich des optisch besten Ergebnisses grob geviertelt. Jeweils rund 30 % empfanden das dritte

Ergebnis (2.3) am besten, beziehungsweise konnten keinen Unterschied erkennen. Rund 20 % empfanden jeweils die Ergebnisse 2.1 und 2.2 als optisch am besten gelungen.

Die Frage, ob generell Video-Artefakte/Bildfehler erkannt wurden, wurde von 81,5 % der Teilnehmer/innen mit „Ja“ beantwortet. Obwohl einige Teilnehmer/innen angaben, keine Unterschiede in den drei Ergebnissen zu erkennen, wurden doch mehrheitlich zumindest in einem Ergebnis Fehler erkannt. Mit einem leichten Überhang konnten in Ergebnis 2.2 die meisten Personen (65) Fehler erkennen. In 2.1 sahen 59 Teilnehmer/innen Artefakte. In Ergebnis 2.3 konnten 57 Personen zumindest einen Fehler erkennen.

Bei den individuellen Antworten gaben viele Teilnehmer/innen an, dass sie Bildfehler/Artefakte im Bereich der Haare erkennen konnten. Diese Art von Fehler konnte bereits in Video 1 erkannt werden, obwohl sich die Haare der Personen in Video 1 und Video 2 nicht nur farblich unterscheiden. In Ergebnis 2.2 und 2.3 sind überdies wesentlich härtere Kanten um die Haare zu sehen, als in Ergebnis 2.1, welches mit Adobe After Effects erzeugt wurde.

Ebenfalls wurde häufig erwähnt, dass Fehler im Bereich der Bank und Sitzfläche beziehungsweise Gesäß der Dame erkannt wurden. In jenem Moment, in dem die Dame ihr Buch beiseite auf die Bank legt, sind deutliche Artefakte zu sehen, da hier offensichtlich Teile der Bank und der vom Buch erzeugte Schatten in die Masken miteinbezogen wurden. Hier zeigt sich ein Problem, dass die K.I. Anwendungen mit Licht/Schatten oder Farbwechsel wechsel Probleme bekommen können.

Der wohl größte Bildfehler ergibt sich durch das Handy, dass von jener Person, die vor der Dame sitzt gehalten wird. Das Handy bewegt sich während des Clips kurzzeitig vor die Hüfte der Dame und wird teilweise ausgeschnitten. Dadurch wirkt es so, als wären teilweise Löcher in der Hüfte der Dame zu sehen.

Sämtliche Ergebnisse des zweiten Videos enthalten Bildfehler und Artefakte und könnten ohne weitere Bearbeitung/Verbesserung so nicht brauchbar verwendet werden.

6.3.5 Optische Analyse des dritten Videos

Im dritten Video läuft ein älterer Mann mit grauem, kurzem Haar und einem Bart dunkel gekleidet und barfuß entlang eines Strandes. Der Mann ist am Anfang des Videos aus einer Entfernung von etwa 3 Metern vollständig zu sehen, während die Kamera sich gegen Ende des Clips auf ungefähr einen Meter in Richtung des Mannes bewegt. Die Kamera folgt dem Mann, um ihn immer im Fokus zu halten. Auf dem Boden vor dem Mann sind seine Spiegelungen im Sand zu sehen. Im Hintergrund sind ein bewölkter, grau-blauer Himmel und ein Meer mit starkem Wellengang zu sehen.

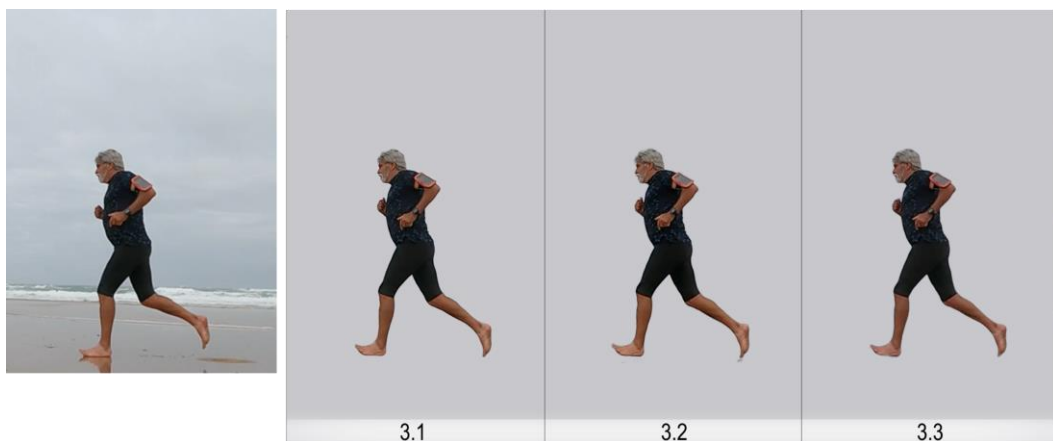


Abbildung 21. Video 3: Vor und nach Bearbeitung.

Die Frage nach dem optisch am besten gelungenen Ergebnis ergab, dass mehr als die Hälfte der Befragten (53,1 %) keinen Unterschied zwischen den drei Ergebnissen erkennen konnten.

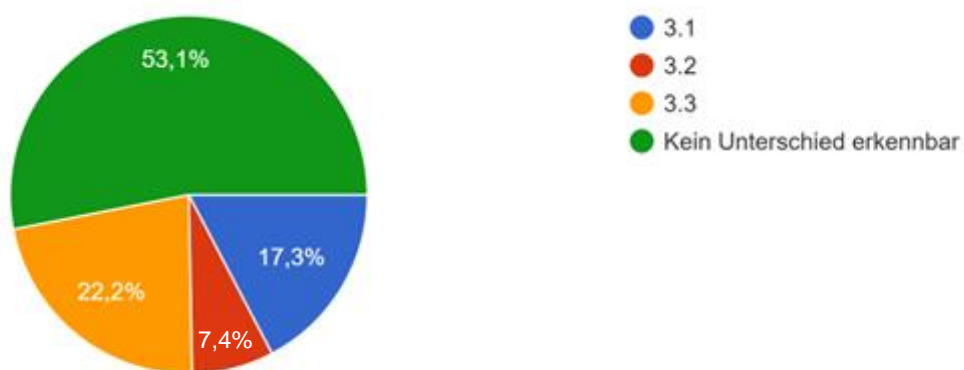


Abbildung 22. Video 3: Ergebnis des optischen Vergleichs.

Dies deckt sich ebenfalls mit der Frage, ob in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt wurden. Diese beantworteten 70,4 % der Teilnehmer/innen mit „Nein“. Ergebnis 3.1 und 3.3 wurden beide jeweils mit rund

20 % als am besten empfunden. Leidglich 7,4 % der Befragten entschieden sich für Ergebnis 3.2.

In Ergebnis 3.2 wurden auch von den meisten Teilnehmer/innen (88 %) Bildfehler erkannt. In 3.1 konnten noch 56 %, in Ergebnis 3.3 44 % der Befragten Artefakte erkennen.

Bei den individuellen Antwortmöglichkeiten wurde am häufigsten erwähnt, dass im Bereich der Füße Fehler erkannt wurden. Im Moment des Auftretens der Füße im Sand wurde häufig erkannt, dass hier Teile des Hintergrunds nicht entfernt wurden und somit in den Ergebnissen kurz zu sehen sind. Da die Teile des Hintergrunds abwechselnd jeweils nur sehr kurzzeitig zu sehen sind und anschließend wieder verschwinden, wird dadurch ein Flackern erzeugt, welches von der Mehrheit der Teilnehmer/innen als störend empfunden wurde.

Aufgrund der Reflexionen der laufenden Person im nassen Sand, kam es im Bereich vor, beziehungsweise unter der Person ebenfalls zu Artefakten, die als „Glitch“ oder „Flackern“ beschrieben wurden. Diese Bildfehler sind jeweils nur weniger als eine Sekunde lang zu sehen, wodurch das erwähnte Flackern entsteht.

Die bereits weiter oben erwähnte Tatsache, dass mehr als die Hälfte der Teilnehmer/innen keinen Unterschied in den Ergebnissen erkennen konnte, ergab auch bei den individuellen Antworten, dass keine Unterschiede oder Bildfehler in allen drei Ergebnissen erkannt werden konnten.

Die Ergebnisse des dritten Videos zeigen, dass die K.I gestützten Rotoskopie Anwendungen bei gleichmäßigen Bewegungen und ruhigem gleichmäßigem Hintergrund schon akzeptable Ergebnisse hervorbringen können. Allerdings ist auch in diesem Fall eine Nachbearbeitung unumgänglich. Die erwähnten Artefakte im Bereich der Füße und die Reflexionen vor der Person fielen in der Befragung doch deutlich auf und müssten nachträglich entfernt werden. Die Artefakte, die durch die Reflexionen im nassen Sand verursacht wurden, könnten jedoch bereits während der Bearbeitung verhindert werden, indem beispielsweise der Videoclip so beschnitten wird, dass besagter Bereich vor/unter der Person gar nicht in den Render-prozess miteinbezogen wird.

6.3.6 Optische Analyse des vierten Videos

Beim vierten Video handelt es sich um eine Nahaufnahme einer Frau mit lockigem, schulterlangem braunem Haar. Die Frau befindet sich in einer offenen Umgebung, mit einem klaren, blauen Himmel und einem Meer samt Strand im Hintergrund, die jedoch unscharf dargestellt sind. Die Frau trägt weiße Kleidung und helle Kopfhörer, an denen zwei weiße Kabel herunterhängen. Die Haare der Frau bewegen sich sanft im Wind und die Kamera verfolgt sie, während sie sich rhythmisch bewegt und eine Hand an ihren Kopfhörern hält.

Da die Inhalte in diesem Clip eher horizontal ausgelegt sind und die Person in einer Nahaufnahme zu sehen ist, wurde die Anordnung für das Vergleichsvideo in der Umfrage geändert. Anstatt die Ergebnisse nebeneinander zu zeigen, wie in den ersten drei Videos, wurde eine Anordnung gewählt, bei der das erste Ergebnis über dem zweiten und dritten Ergebnis angeordnet ist, welche nebeneinander im unteren Teil des Vergleichsvideos zu sehen sind.



Abbildung 23. Video 4: Vor und nach Bearbeitung.

Ähnlich wie im dritten Video konnte beinahe die Hälfte der Befragten (45,7 %) keinen Unterschied in den Ergebnissen erkennen. Die Aufteilung bezüglich des optisch besten Ergebnisses fiel für die drei Ergebnisse annähernd gleich aus. Lediglich Ergebnis 4.3 wurde von 19,8 % der Teilnehmer/innen geringfügig besser empfunden als die Ergebnisse 4.1 und 4.2. In Abbildung 24 ist im Diagramm die genaue Aufteilung zu sehen.

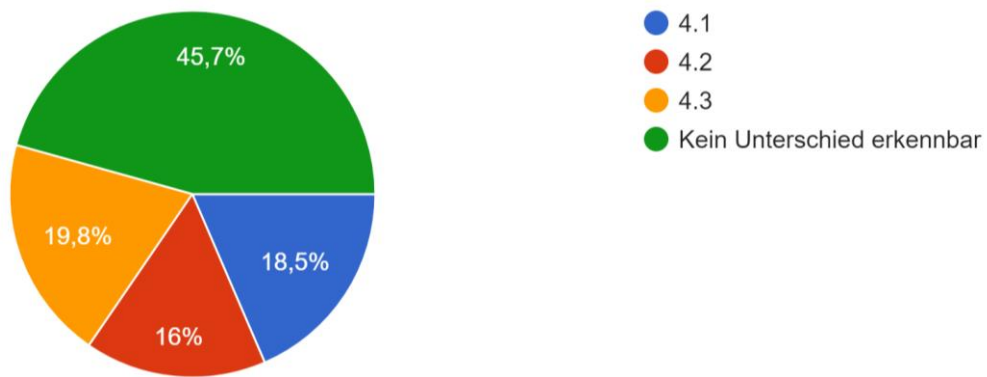


Abbildung 24. Video 4: Ergebnis des optischen Vergleichs.

54 Teilnehmer/innen (66,7 %) konnten in keinem der drei Ergebnisse Video-Artefakte/Bildfehler erkennen. Die anderen 33,3 % fanden jedoch mindestens in einem Ergebnis einen Fehler.

Am häufigsten wurden in Ergebnis 4.2 ein oder mehrere Fehler gefunden, gefolgt von 4.1 und 4.3.

Ähnlich wie in Video 1 und Video 2 wurde bei den Individual- Antworten am häufigsten erwähnt, dass Fehler bei den Haaren der Person erkannt werden konnten. Es wurde von feinen Haarsträhnen berichtet, die verschwinden, sobald sie zu dünn werden. Teilweise wurde beschrieben, dass Teile der Haare nicht mit dem Rest des Kopfes verbunden sind. Einmal wurde eine „Unnatürliche Freistellung der Haare“ erwähnt, womit wohl dieselben bereits erwähnten Fehler gemeint sind.

Circa die Hälfte der Antworten bezog sich auf Artefakte in den Haaren der freigestellten Person. Jedoch wurde am zweit-häufigsten erwähnt, dass keine Fehler oder keine Unterschiede in den Ergebnissen erkannt werden konnten.

Ein möglicher Grund für diese gleichmäßig aufgeteilten Antworten könnte sein, dass, wie in Kapitel „Analyse der allgemeinen Fragen“ beschrieben, circa die Hälfte der Teilnehmer/innen angab, keine Vorkenntnisse im Bereich Bild/Video Bearbeitung zu besitzen, während die andere Hälfte der Befragten jedoch angab, Kenntnisse im besagten Bereich zu besitzen. Möglicherweise achteten jene Personen mit Vorkenntnissen detaillierter auf mögliche Bildfehler, während Personen ohne Kenntnisse diese schlicht nicht wahrnahmen.

6.3.7 Optische Analyse des fünften Videos

Im fünften Video wird eine dunkelhäutige Person mit kurzen Haaren in einem abgedunkelten Zimmer gezeigt. Das Licht in der Aufnahme hat eine gelblich-braune Farbe, die sich auf das T-Shirt der abgebildeten Person überträgt. Die Kamera ist starr und zeigt die Person bis zu den Knien auf einem gepolsterten Sessel sitzend. Zu Beginn des Clips ist nur die Silhouette der Person zu sehen, während sie aufrecht sitzt. Dann lehnt sich die Person zurück und das Licht fällt langsam auf ihren Oberkörper. Zusätzlich bewegt die Person ihre linke Hand auf ihren Kopf zu und hält ihn bis zum Ende der Aufnahme.

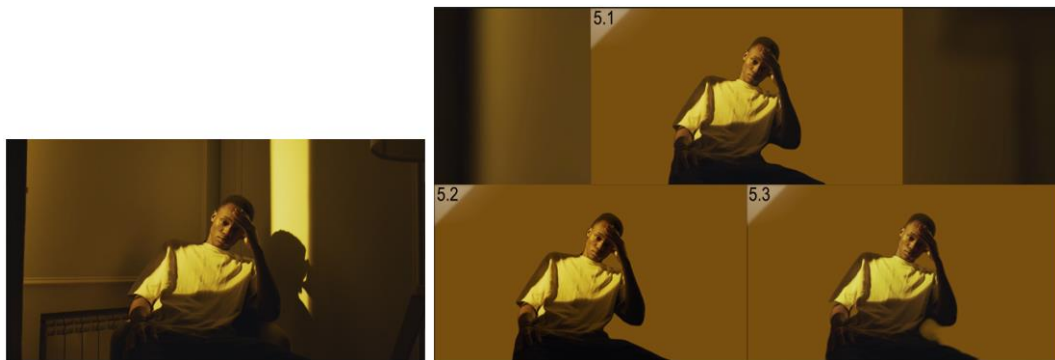


Abbildung 25. Video 5: Vor und nach Bearbeitung.

Für das fünfte Video wurde ebenfalls, wie in Video 4, eine Anordnung für das Vergleichs-video in der Umfrage gewählt, bei der das erste Ergebnis über dem zweiten und dritten Ergebnis angeordnet ist.

Eine deutliche Mehrheit der Teilnehmer/innen (48,1 %) empfand Ergebnis 5.1 als optisch am besten gelungen.

Etwas mehr als ein Viertel der Befragten entschied sich für Ergebnis 5.2. Lediglich 8,6 % wählten Ergebnis 5.3 aus. Allerdings konnten auch 17,3 % der teilnehmenden Personen keinen Unterschied in den drei Ergebnissen erkennen.

In Abbildung 26 ist das dazugehörige Diagramm abgebildet.

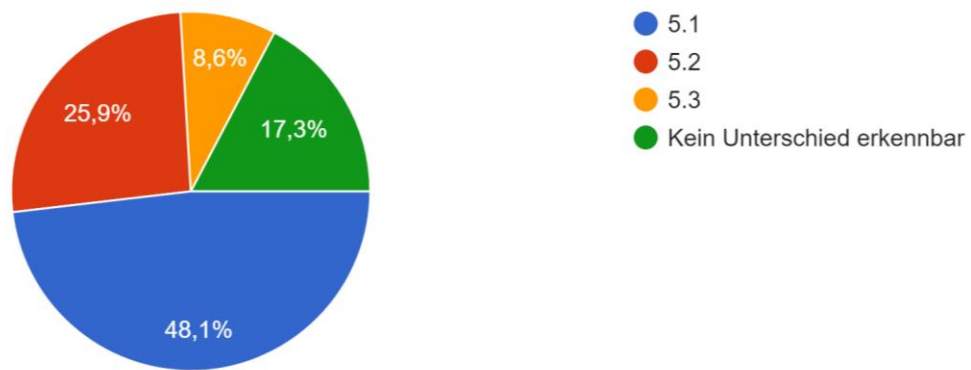


Abbildung 26. Video 5: Ergebnis des optischen Vergleichs.

84 % der Teilnehmer/innen beantworteten die Frage, ob Video-Artefakte/Bildfehler erkannt werden konnten, mit „Ja“. Nur 16 % der Befragten konnten in diesem Video keine Fehler erkennen. Wenn die Befragten einen oder mehrere Fehler erkennen konnten, dann wurden diese in allen drei Ergebnissen entdeckt.

Die häufigste Antwort, die bei den individuellen Antwortmöglichkeiten gegeben wurde, ist, dass bei Ergebnis 5.2 in der Mitte der Person beziehungsweise am Bauch, ein Loch zu sehen ist. Dies erscheint in dem Moment, in dem sich die Person im Sessel vorbeugt und Licht auf den Oberkörper der Person fällt. Durch dieses Loch im Bauch ist der neutrale, gelbe Hintergrund des Vergleichsvideos zu sehen. Dieses Loch tritt lediglich bei Ergebnis 5.2 in Erscheinung. Allerdings empfanden mehrere Teilnehmer/innen die Kanten rund um die Person in 5.2 als besser gelungen als in den anderen beiden Ergebnissen.

Ähnlich wie in den vorangegangenen Videos, wurde auch bei Video 5 ein „Flimmern“ an den Kanten der Person erwähnt, welches auch als „Flackern“ oder „pulsierende Ränder“ mehrmals genannt wurde. Die Kanten wurden auch vielfach als unscharf beziehungsweise „geblurred“ empfunden. Es wurden zwar mehrere unterschiedliche Fehler in den jeweiligen Ergebnissen gefunden, jedoch wurde keines der Ergebnisse als fehlerfrei empfunden, sofern die Teilnehmer/innen generell angaben, Fehler erkannt zu haben.

Die Ergebnisse des fünften Videos zeigen einige Grenzen der K.I. gestützten Rotoskopie-Anwendungen auf. Bei dunkleren und/oder wechselnden Lichtverhältnissen und wenig Kontrast, wurden doch mehrere Artefakte in den Ergebnissen erzeugt. Speziell Ergebnis 5.2 wäre wohl kaum weiter verwendbar, ohne einer umfangreichen Nachbearbeitung.

6.3.8 Optische Analyse des sechsten Videos

Im sechsten und letzten Video ist eine Gruppe älterer Menschen in einem hellen, geschlossenen Raum zu sehen, die von einer Trainerin angeleitete Fitness-Übungen ausführen. Insgesamt sind sechs Personen, einschließlich der Trainerin, aus verschiedenen Entfernungen vollständig zu sehen. Zwei männliche Personen in der Mitte der starren Aufnahme wurden rotoskopiert. Die Personen stehen auf roten und blauen Gummi-Matten, die auf einem grauen Fußboden platziert sind. Im Hintergrund ist eine graue Wand zu sehen. Die rotoskopierten Personen tragen helle, rote und gelbe Oberteile, graue Hosen und Sportschuhe. Zu Beginn des Clips sind die Personen stehend mit hochgestreckten Armen zu sehen. Anschließend setzen sie sich auf weiße Klapp-Sessel, die hinter ihnen platziert sind, und bewegen ihre Arme nach unten. Der rechte Ellbogen der vorderen Person verdeckt teilweise den linken Unterarm der hinten stehenden Person.

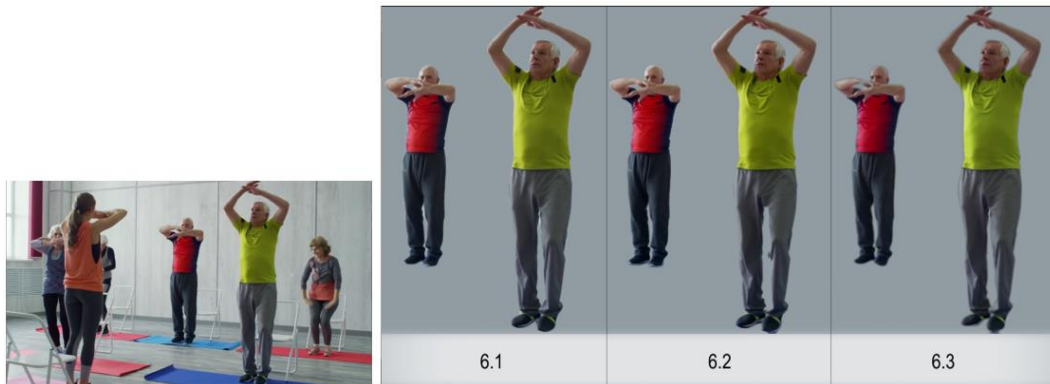


Abbildung 27. Video 6: Vor und nach Bearbeitung.

Mehr als ein Drittel der Befragten (39,5 %) konnten keinen Unterschied in den Ergebnissen erkennen. Wurde ein Unterschied erkannt, so empfanden 32,1 % der Teilnehmer/innen das erste Ergebnis (6.1) als optisch am besten gelungen. Ergebnis 6.2 wurde noch 17,3 % ausgewählt; Ergebnis 6.3 wählten lediglich 11,1 % der Befragten aus. In Abbildung 28 sind die diesbezüglichen Ergebnisse grafisch dargestellt.

Video-Artefakte und/oder Bildfehler wurden von 65,4 % der Teilnehmer/innen erkannt. Diese Bildfehler wurden in sämtlichen Ergebnissen gefunden, in Ergebnis 6.2 jedoch am häufigsten.

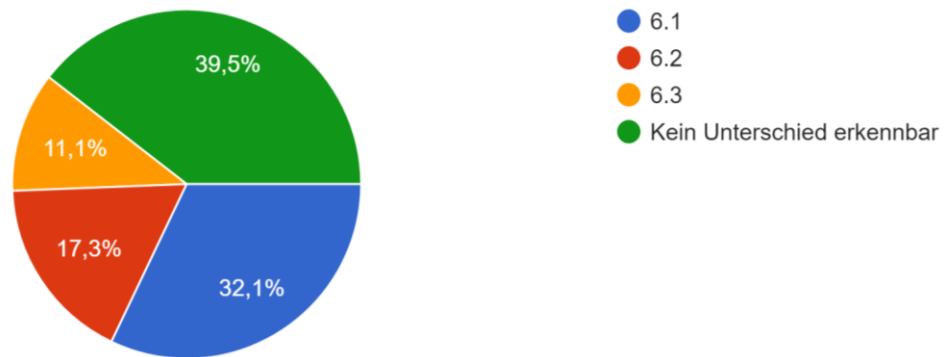


Abbildung 28. Video 6: Ergebnis des optischen Vergleichs.

Die häufigste Antwort, welche bei den individuellen Antworten gegeben wurde, bezog sich auf Fehler an den Armen und am Kopf. In jenem Moment, in dem die beiden dargestellten Personen ihre ausgestreckten Arme herabsenken, überkreuzen sich diese und verdecken kurzzeitig Teile des Kopfes und des Gesichts. In diesem Bildbereich wurde teilweise Elemente des Hintergrundes mit in die Auswahl genommen. Speziell in Ergebnis 6.2 sind deutliche Artefakte an den Händen der vorderen Person zu sehen. In diesem Bereich dürfte zu besagtem Moment die Maske unklar definiert sein, wodurch Teile der Hände abgeschnitten werden und Löcher, beziehungsweise „Flimmern“ entsteht. Auch in den schmalen Bereichen zwischen den Armen und dem Oberkörper und zwischen den Beinen, wurde teilweise der Hintergrund mit in die Auswahl der Maske genommen. Dies tritt vor allem während des Hinsetzens der Personen auf. In Ergebnis 6.3 wird gegen Ende des Clips, somit nach dem Hinsetzen, das Gesicht der hinteren Person teilweise abgeschnitten und eine unscharfe Maskenkante erzeugt.

Einige weitere Antworten bezogen sich auf den Bereich der Fußsohlen und der Füße der Personen. In jenen kurzen Momenten, in denen die Personen ihre Füße anheben und Stampf-Bewegungen ausführen, werden Teile des Bodens mit in die Auswahl genommen und erzeugen so Artefakte und Flimmern an den Kanten der Maske.

Auch die Ergebnisse des sechsten Videos weisen somit einige Fehler und Artefakte auf. Da speziell im Bereich der grauen Haare und der grauen Hosen und Schuhe der Personen, wenig Kontrast zum Hintergrund des Original-Clips besteht, konnten die drei K.I. gestützten Rotoskopie-Anwendungen nur fehlerbehaftete Ergebnisse liefern. Speziell bei überlappenden und überkreuzenden Körperteilen und Objekten wurden einige Bildfehler erzeugt, trotz der eher langsamen Bewegungen der beinhalteten Personen.

7 Fazit

In dieser Masterarbeit wurden Anwendungsmöglichkeiten von künstlicher Intelligenz in der digitalen Video-Post Production und im Digital Design vorgestellt und thematisiert. Künstliche Intelligenz wird zunehmend immer häufiger im Alltag eingesetzt. Sie wird sowohl im beruflichen Kontext als auch im privaten Bereich bereits von vielen Menschen genutzt. Von Anwendungen wie autonom (selbst-) fahrender Autos über KI-Assistenten wie Siri und Cordana und Translation Tools wie Deepl oder Google Translate, reichen die Anwendungsbereiche künstlicher Intelligenz.

Auch in den Bereichen Grafik und Design sowie Animation, Post Production und Visual Effects existieren bereits viele brauchbare Anwendungen und Programme, die einen Mehrwert für ihre Anwender/innen bieten.

Die erste Forschungsfrage befasste sich mit Tasks im Bereich der digitalen Video-Postproduktion, bei denen Artists durch den Einsatz von KI-Systemen aktuell unterstützt werden können. Aus der Literaturrecherche ergab sich folgendes Bild: Künstliche Intelligenz sollte primär als Unterstützung angesehen werden, um bestimmte repetitive Aufgaben zu übernehmen. Die aktuellen Anwendungen reichen von Farbkorrekturen über Objektentfernungen bis hin zu visuellen Effekten und Bildstabilisierungen. Die dadurch eingesparte Zeit, die bisher für mechanistische Tätigkeiten benötigt wurde, kann enorm sein und somit Raum für mehr Kreativität schaffen. In diesem Sinne kann Künstliche Intelligenz größere kreative Freiräume ermöglichen.

Eine weitere Forschungsfrage untersuchte Methoden, um Vordergrund-Objekte in Videomaterial von dessen Hintergrund zu trennen. Hierfür wurden neben dem klassischen Rotoscoping mehrere Arten von Mattes vorgestellt, die für teils unterschiedliche Anwendungsfälle eingesetzt werden können.

Für den praktischen Teil der Arbeit wurden drei Programme/Anwendungen vorgestellt, mit denen K.I gestütztes Rotoscoping möglich ist. Diese Anwendungen waren Adobe After Effects, The Foundry Nuke (unter Einsatz der CopyCat Funktion) und Runway ML. Zuerst wurde der allgemeine Workflow in der jeweiligen Anwendung beschrieben. Anschließend wurden sechs Videos mit den Anwendungen bearbeitet, in denen sich Personen (reale Menschen) in unterschiedlichen Situationen, vor unterschiedlichen Hintergründen und unterschiedlichen Kameraeinstellungen befinden. Eine objektive Analyse wurde bezüglich der Anzahl der Arbeitsschritte, dem Zeitaufwand für die Bearbeitung und der Renderzeit durchgeführt. Diese ergab, dass sich zwei der Anwendungen

(Adobe After Effects und Runway ML) punkto Dauer für die Bearbeitung und Renderzeit ähnlich sind, die dritte Anwendung (The Foundry Nuke) sich jedoch in den genannten Punkten stark unterscheidet. Die Anzahl der Arbeitsschritte war zwar in sämtlichen Anwendungen ähnlich; die Art der Arbeitsschritte jedoch unterschiedlich. Ein direkter Zusammenhang zwischen längerer Renderzeit und optisch besseren Ergebnissen konnte nicht festgestellt werden.

Für die Forschungsfrage, wie sich die Ergebnisse der vorgestellten Anwendungen/Techniken hinsichtlich optischer Qualität unterscheiden, wurde eine quantitative Analyse durchgeführt. Hierfür wurde eine Online-Umfrage mittels „Google Forms“ erstellt und an mehrere Empfänger weitergeleitet. Neben allgemeinen und demoskopischen Fragen konnten die Teilnehmer und Teilnehmerinnen zwischen vordefinierten Antworten wählen und überdies hatten sie die Möglichkeit, eigenes Feedback zu geben. Die Ergebnisse der Umfrage bestätigen die zuvor aufgestellte Hypothese. Mit den Videos, in denen die beinhalteten Personen schnelle oder unregelmäßige Bewegungen ausführen, konnten wesentlich schlechtere Ergebnisse erzielt werden als in jenen Videos mit ruhigen Bewegungen. Speziell an den Objektkanten und im Bereich der Haare konnten einige Fehler identifiziert werden. Einen weiteren Einfluss hatten die Lichtverhältnisse in den bearbeiteten Szenen. Die Hintergründe hatten ebenfalls einen Einfluss auf die Ergebnisse. Bei gleichmäßigen und ruhigen Hintergründen fielen die Ergebnisse besser aus, als bei unregelmäßigen und unruhigen Hintergründen, in denen viele andere bewegte Objekte oder unregelmäßige Muster zu sehen waren. Vor allem im zweiten und im fünften Video konnte eine große Mehrheit der Teilnehmer/innen Bildfehler und Artefakte erkennen.

Abschließend kann somit gesagt werden, dass die K.I gestützten Rotoskopie-Anwendungen bei ruhigen und gleichmäßigen Bewegungen bzw. Hintergründen bereits brauchbare Ergebnisse liefern, mit denen weitergearbeitet werden kann. Eine weitere Bearbeitung ist jedoch im Großteil der Fälle notwendig, um zu wirklich zufriedenstellenden Ergebnissen zu gelangen. Bei schnellen und unregelmäßigen Bewegungen und sich schnell wechselnden oder unruhigen Hintergründen hingegen, ist eine umfangreiche Nachbearbeitung notwendig, um die Ergebnisse verwerten zu können. In manchen Fällen werden auch unbrauchbare Ergebnisse generiert werden. Eine Weiterentwicklung der bestehenden Techniken und Anwendungen wird unumgänglich sein, um zukünftig ohne Nachbearbeitung auszukommen.

Der Einsatz von künstlicher Intelligenz wird zukünftig sowohl Vorteile als auch Nachteile mit sich bringen. Einerseits kann K.I bereits viele gleichmäßige, monotone oder repetitive Aufgaben übernehmen, was den Anwendern Zeitersparnis bringt. Dadurch können sich Anwender/innen auf kreativere Aufgaben konzentrieren. Jedoch existieren nach wie vor zahlreiche Bereiche, in

denen künstliche Intelligenz nicht an das Niveau menschlicher Personen herankommt und vermutlich auch in näherer Zukunft nicht herankommen wird. Bis dahin müssen noch zahlreiche ethische, rechtliche und technische Aspekte geklärt werden. Künstliche Intelligenz wird wohl nicht mehr aus der Gesellschaft und dem Berufsalltag verschwinden. Die Gesellschaft wird lernen müssen mit künstlicher Intelligenz umzugehen, ihr Potenzial sinnvoll einzusetzen und nicht gegen sie zu arbeiten.

Literaturverzeichnis

Buch:

Buchkremer, R., Heupel, T., Koch, O., (2020). *Künstliche Intelligenz in Wirtschaft & Gesellschaft-Auswirkungen, Herausforderungen & Handlungsempfehlungen*. Springer-Gabler

Ertel, W. (2016). *Grundkurs Künstliche Intelligenz: eine praxisorientierte Einführung*. Springer Vieweg

Wittpahl, V. (2019). *Künstliche Intelligenz. Technologie | Anwendung | Gesellschaft*. Springer Vieweg

Wennker, P. (2020). *Künstliche Intelligenz in der Praxis, Anwendung in Unternehmen und Branchen: KI wettbewerbs- und zukunftsorientiert einsetzen*. Springer-Gabler

Haarmeier, M. (2021). *Künstliche Intelligenz für den Mittelstand- Erfolgreiche Einführung und Nutzung von KI-Anwendungen in Unternehmen*. Springer-Gabler

Kitzmann, A. (2022). *Künstliche Intelligenz. Wie verändert sich unsere Zukunft?* Springer

Gethmann, C., Buxmann, P., Distelrath, J., Humm, B., Lingner, S., Nitsch, V., Schmidt, J., Spiecker, I. (2022). *Künstliche Intelligenz in der Forschung*. Springer

Zylinska, J. (2020). *AI Art. Machine Visions and Warped Dreams*. Open Humanities Press

Brinkmann, R. (2008). *The Art and Science of Digital Compositing- Second Edition*. Elsevier

Wright, S. (2010). *Digital Compositing for Film and Video-Third Edition*. Elsevier

Engenhardt, M., Löwe, S. (2022). *Design und künstliche Intelligenz. Theoretische und praktische Grundlagen der Gestaltung mit maschinell lernenden Systemen*. Birkhäuser Basel

Artikel-Journal:

Scherk, J., Pöchlacker-Tröscher, G., Wagner, K. (2017). *Künstliche Intelligenz - Artificial Intelligence*. https://www.bmk.gv.at/dam/jcr:32e07784-1825-4441-880f-39b67cf8fc4b/kuenstliche_intelligenz.pdf

Stadler, M. (2019). *KI im Mittelstand*. mindsquare

Stadler, M. (2022). *Künstliche Intelligenz*. mindsquare

Anantrasirichai, N., Bull, D. (2021). *Artificial intelligence in the creative industries: a review*. <https://doi.org/10.1007/s10462-021-10039-7>

Davies, J., Klinger, J., Mateos-Garcia, J., Stathoulopoulos, K. (2020). *The Art in the Artificial- AI and the creative industries*. <https://cdn2.assets-servd.host/creative-pec/production/assets/publications/PEC-and-Nesta-report-The-art-in-the-artificial.pdf>

Kreutzer, T., Christiansen, P. (2021). *KI in Unternehmen. Ein Praxisleitfaden zu rechtlichen Fragen*. Bertelsmann Stiftung

Bermudez, L., Dabby, N., Lin, Y. A., Hilmarsdottir, S., Sundararajan, N., Kar, S., (2021). A Learning-Based Approach to Parametric Rotoscoping of Multi-Shape Systems. <https://ieeexplore.ieee.org/document/9423380>

Website:

Stadler, M. (2021). Machine Learning-Datenanalysemethode. Abgerufen 22.02.2023 von <https://mindsquare.de/knowhow/machine-learning/>

Adobe Inc. (2023). „Roto-Pinsel“ und „Maske verbessern“. Abgerufen 10.02.2023 von <https://helpx.adobe.com/at/after-effects/using/roto-brush-refine-matte.html>

The Foundry (2023). „Training Networks to Replicate Effects Using Machine Learning“. Abgerufen 13.02.2023 von https://learn.foundry.com/nuke/13.1/content/comp_environment/air_tools/cc-intro.html?Highlight=copy%20cat

The Foundry (2023). „Creating a Data Set for Training“. Abgerufen 13.02.2023 von https://learn.foundry.com/nuke/13.1/content/comp_environment/air_tools/cc-data.html

The Foundry (2023). „Training and Monitoring the Network“. Abgerufen 13.02.2023 von https://learn.foundry.com/nuke/13.1/content/comp_environment/air_tools/cc-train.html

The Foundry (2023). „CopyCat“. Abgerufen 13.02.2023 von https://learn.foundry.com/nuke/content/reference_guide/air_nodes/copycat.html

The Foundry (2023). „Applying and Improving the Results“. Abgerufen 13.02.2023 von https://learn.foundry.com/nuke/13.1/content/comp_environment/air_tools/cc-apply.html

Runwayml (2023). „Green Screen“. Abgerufen 14.02.2023 von <https://help.runwayml.com/hc/en-us/articles/4412269871379-Green-Screen>

Runwayml (2023). „Pricing Plans“. Abgerufen 14.02.2023 von <https://runwayml.com/pricing/>

Negovec, L. (2022). Maschinelle Übersetzung: Methoden und Schwierigkeiten. Abgerufen 13.03.2023 von <https://www.allesprachen.at/blog/maschinelle-uebersetzung/>

Hooksounds (2023). Künstliche Intelligenz in der Filmindustrie. Abgerufen 23.03.2023 von <https://www.hooksounds.com/de/blog/kuenstliche-intelligenz-der-filmindustrie/>

Movie-College (2023). Artificial Intelligence & Film. Abgerufen 27.03.2023 von <https://www.movie-college.de/filmschule/postproduktion/montage/artificial-intelligence-film>

Industry of Things (2022). Was ist generative KI und was kann sie? Abgerufen 27.03.2023 von <https://www.industry-of-things.de/was-ist-generative-ki-und-was-kann-sie-a-8faf44a80c7de6711d3b05875722c122/>

Google Inc. (2023). Google Forms. Abgerufen 04.04.2023 von <https://www.google.com/forms/about/>

Schirlbauer, J. (2021). Anwendung von künstlicher Intelligenz im Digital Design. abgerufen 18.04.2023 von <https://www.crafted.at/b/kuenstliche-intelligenz-im-digital-design/>

Seymour, M. (2018). Rotobot, bringing Machine Learning to Roto. abgerufen 18.04.2023 von <https://www.fxguide.com/quicktakes/roto-robot-bringing-machine-learning-to-roto/>

Abbildungsverzeichnis

Abbildung 1. KI, Machine Learning und Deep Learning in der Entwicklung (Stadler, 2019, S.8).	23
Abbildung 2. Der Aufbau von Neuronen-Schichten im Deep Learning (Stadler, 2019, S.10).	25
Abbildung 3. Traditionelle Programme versus ML (Wittpahl, 2019, S.25).	26
Abbildung 4. Difference Mattes: Originalaufnahme und Clean Plate (Wright, 2010).	58
Abbildung 5. raw Difference Matte und binäre Difference Matte (Wright, 2010).	58
Abbildung 6. Difference: komplexes Objekt (Wright, 2010).	59
Abbildung 7. After Effects Roto Pinsel 2- Auswahl des Objekts (Adobe Inc., 2023).	61
Abbildung 8. Nuke CopyCat dataset (The Foundry, 2023).	63
Abbildung 9. Nuke CopyCat Node Graph	64
Abbildung 10. Nuke CopyCat Contact Sheets (The Foundry, 2023).	66
Abbildung 11. Runway Bearbeitung	68
Abbildung 12. Dauer für die Bearbeitung	72
Abbildung 13. Geschlecht der Teilnehmer und Teilnehmerinnen	75
Abbildung 14. Altersgruppen der Teilnehmer und Teilnehmerinnen.	76
Abbildung 15. Vorkenntnisse der Teilnehmer und Teilnehmerinnen im Bereich Bild- und Video Bearbeitung	77
Abbildung 16. Verwendete Geräte zur Betrachtung der Videos.	77
Abbildung 17. Video 1: Vor und nach Bearbeitung.	78
Abbildung 18. Video 1: Ergebnis des optischen Vergleichs.....	79
Abbildung 19. Video 2: Vor und nach Bearbeitung.	80
Abbildung 20. Video 2: Ergebnis des optischen Vergleichs.....	80
Abbildung 21. Video 3: Vor und nach Bearbeitung.	82
Abbildung 22. Video 3: Ergebnis des optischen Vergleichs.....	82

Abbildung 23. Video 4: Vor und nach Bearbeitung.	84
Abbildung 24. Video 4: Ergebnis des optischen Vergleichs.....	85
Abbildung 25. Video 5: Vor und nach Bearbeitung.	86
Abbildung 26. Video 5: Ergebnis des optischen Vergleichs.....	87
Abbildung 27. Video 6: Vor und nach Bearbeitung.	88
Abbildung 28. Video 6: Ergebnis des optischen Vergleichs.....	89

Tabellenverzeichnis

Tabelle 1. Renderzeit.....	74
----------------------------	----

Anhang

Fragenbogen der Online- Umfrage

UMFRAGE K.I gestützte Rotoskopie-Anwendungen

Sehr geehrte Damen und Herren,

Vielen Dank für die Teilnahme an meiner Umfrage!

Mein Name ist Matthias Bäck, ich bin Student an der FH St. Pölten und beschäftige mich im Rahmen meiner Masterarbeit mit K.I gestützten Rotoskopie-Anwendungen. (K.I = Künstliche Intelligenz). Mit diesen ist es möglich Vordergrund-Objekte, im konkreten Fall Personen, von deren Hintergrund zu trennen. Die Umfrage dauert ca. 10 Minuten.

Bitte nehmen Sie sich ausreichend Zeit, um die Fragen gewissenhaft beantworten zu können. Die mit dieser Umfrage gewonnenen Daten werden ausschließlich zu wissenschaftlichen Zwecken im Rahmen meiner Masterarbeit verwendet. Sie können diesen Fragebogen gerne an Kolleginnen und Kollegen, Verwandte und Freunde weiterleiten. Bei Fragen stehe ich gerne unter mp211504@fhstp.ac.at zur Verfügung.

Matthias Bäck

Student Masterstudium Digital Media Production/ Animation & Visual Effects

FH St. Pölten

Allgemein

Ihnen werden im Folgenden 6 Videos vorgestellt, in denen mit 3 unterschiedlichen Anwendungen/Programmen Personen in Videos mittels K.I. unterstützten Rotoskopie-Anwendungen von deren jeweiligen Hintergrund getrennt wurden. Die Ergebnisse der 3 Anwendungen werden jeweils in einem Video nebeneinander dargestellt. So wird ein direkter Vergleich gewährleistet. Die Videos können entweder direkt in der Umfrage angesehen werden oder, per Klick auf das untenstehende „You Tube“ Symbol, auf You Tube.

Ich bitte Sie darum, die Videos genau anzusehen, die Fragen zu beantworten und Ihre Auswahl auch kurz zu begründen. Vielen Dank!

Bitte geben Sie ihr Geschlecht an.

- ☐ Weiblich
- ☐ Männlich
- ☐ Divers
- ☐ Keine Angabe

Bitte geben Sie ihre Altersgruppe an.

- ☐ 10-20
- ☐ 20-30
- ☐ 30-40
- ☐ 40-50
- ☐ 50-60
- ☐ 60-70
- ☐ 70-80

Besitzen Sie Vorkenntnisse im Bereich Bild/Video Bearbeitung?

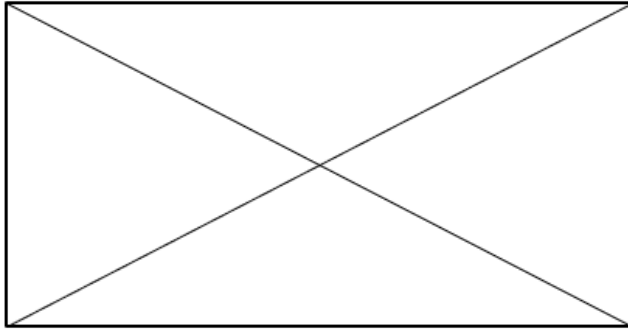
- ☐ Ja
- ☐ Nein

Auf welchem Gerät sehen Sie sich die Videos an?

- ☐ PC/Desktop
- ☐ Laptop
- ☐ Tablet
- ☐ Smartphone

Video 1

Mit Klick auf das "You Tube" Symbol, können Sie das Video in größerer Darstellung auf You Tube ansehen.



Darstellung des Videos

Welches Ergebnis sieht optisch am besten aus?

- ☐ 1.1
- ☐ 1.2
- ☐ 1.3
- ☐ Kein Unterschied erkennbar

Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?

- ☐ Ja
- ☐ Nein

Wenn ja, in welchen Ergebnissen werden Video-Artefakte/Bildfehler erkannt?

- ☐ 1.1
- ☐ 1.2
- ☐ 1.3

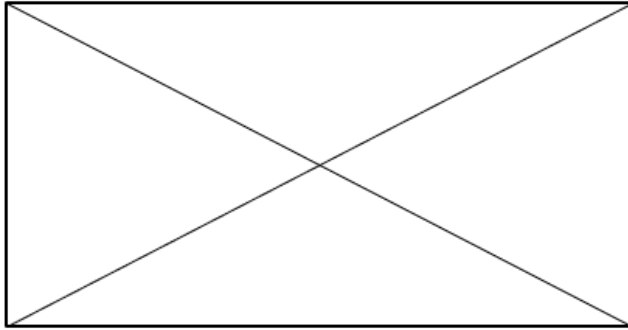
(Mehrfachauswahl möglich)

Bitte begründen Sie ihre Antwort.

Langtextantwort:

Video 2

Mit Klick auf das "You Tube" Symbol, können Sie das Video in größerer Darstellung auf You Tube ansehen.



Darstellung des Videos

Welches Ergebnis sieht optisch am besten aus?

- ☐ 2.1
- ☐ 2.2
- ☐ 2.3
- ☐ Kein Unterschied erkennbar

Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?

- ☐ Ja
- ☐ Nein

Wenn ja, in welchen Ergebnissen werden Video-Artefakte/Bildfehler erkannt?

- ☐ 2.1
- ☐ 2.2
- ☐ 2.3

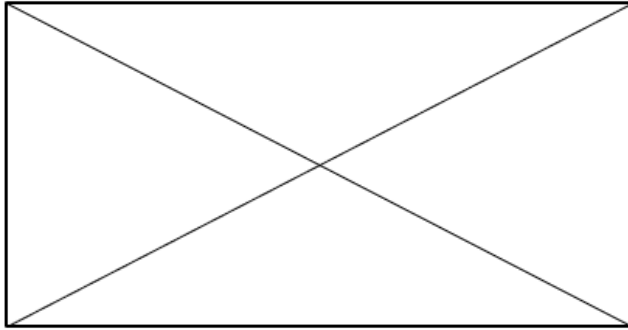
(Mehrfachauswahl möglich)

Bitte begründen Sie ihre Antwort.

Langtextantwort:

Video 3

Mit Klick auf das "You Tube" Symbol, können Sie das Video in größerer Darstellung auf You Tube ansehen.



Darstellung des Videos

Welches Ergebnis sieht optisch am besten aus?

- ☐ 3.1
- ☐ 3.2
- ☐ 3.3
- ☐ Kein Unterschied erkennbar

Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?

- ☐ Ja
- ☐ Nein

Wenn ja, in welchen Ergebnissen werden Video-Artefakte/Bildfehler erkannt?

- ☐ 3.1
- ☐ 3.2
- ☐ 3.3

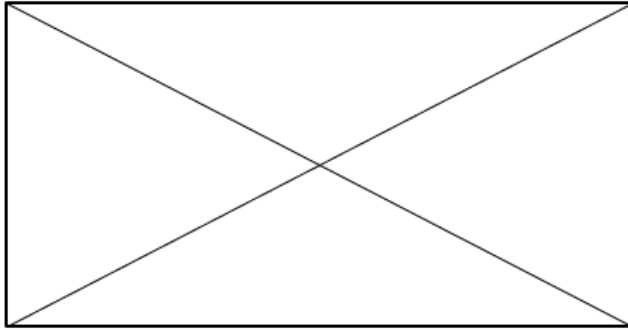
(Mehrfachauswahl möglich)

Bitte begründen Sie ihre Antwort.

Langtextantwort:

Video 4

Mit Klick auf das "You Tube" Symbol, können Sie das Video in größerer Darstellung auf You Tube ansehen.



Darstellung des Videos

Welches Ergebnis sieht optisch am besten aus?

- ☐ 4.1
- ☐ 4.2
- ☐ 4.3
- ☐ Kein Unterschied erkennbar

Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?

- ☐ Ja
- ☐ Nein

Wenn ja, in welchen Ergebnissen werden Video-Artefakte/Bildfehler erkannt?

- ☐ 4.1
- ☐ 4.2
- ☐ 4.3

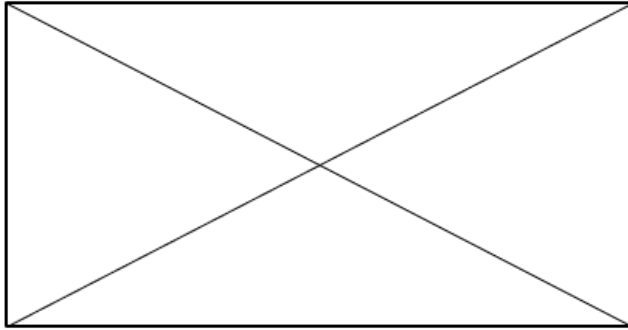
(Mehrfachauswahl möglich)

Bitte begründen Sie ihre Antwort.

Langtextantwort:

Video 5

Mit Klick auf das "You Tube" Symbol, können Sie das Video in größerer Darstellung auf You Tube ansehen.



Darstellung des Videos

Welches Ergebnis sieht optisch am besten aus?

- ☐ 5.1
- ☐ 5.2
- ☐ 5.3
- ☐ Kein Unterschied erkennbar

Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?

- ☐ Ja
- ☐ Nein

Wenn ja, in welchen Ergebnissen werden Video-Artefakte/Bildfehler erkannt?

- ☐ 5.1
- ☐ 5.2
- ☐ 5.3

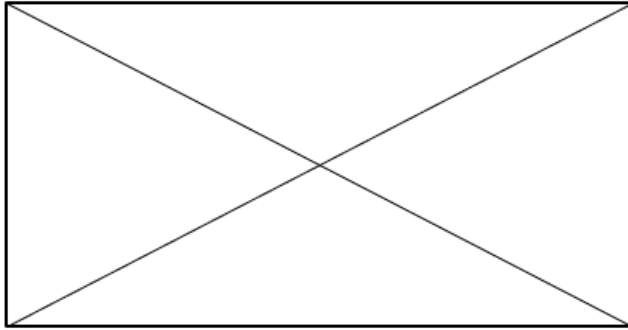
(Mehrfachauswahl möglich)

Bitte begründen Sie ihre Antwort.

Langtextantwort:

Video 6

Mit Klick auf das "You Tube" Symbol, können Sie das Video in größerer Darstellung auf You Tube ansehen.



Darstellung des Videos

Welches Ergebnis sieht optisch am besten aus?

- ☐ 6.1
- ☐ 6.2
- ☐ 6.3
- ☐ Kein Unterschied erkennbar

Werden in einem oder mehreren Ergebnissen Video-Artefakte/Bildfehler erkannt?

- ☐ Ja
- ☐ Nein

Wenn ja, in welchen Ergebnissen werden Video-Artefakte/Bildfehler erkannt?

- ☐ 6.1
- ☐ 6.2
- ☐ 6.3

(Mehrfachauswahl möglich)

Bitte begründen Sie ihre Antwort.

Langtextantwort: