Projektarbeit

Deepfakes und Social Engineering

vorgelegt von

Julian Faigle (Matrikelnummer: 86292)
Studiengang ITS

Max Ernstschneider (Matrikelnummer: 86464)
Studiengang AIT

Semester 6

Hochschule Aalen

Hochschule für Technik und Wirtschaft

Betreut durch Prof. Roland Hellman

15.08.2024

Erklärung

Wir versichern, dass wir die Ausarbeitung mit dem Thema „Deepfakes und Social Engineering“ selbstständig verfasst haben und keine anderen Quellen und Hilfsmittel als die angegebenen benutzt haben. Die Stellen, die anderen Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind in jedem einzelnen Fall unter Angabe der Quelle als Entlehnung (Zitat) kenntlich gemacht worden. Das Gleiche gilt für beigefügte Skizzen und Darstellungen.

Aalen, den 30. Juni 2024

Ort, Datum

Julian Faigle

Autor

Max Ernstschneider

Autor

Inhaltsverzeichnis

Akronyme

LAN Local Area Network

Glossar

Motion Tracking	Motion Tracking ist eine Technik, die verwendet wird, um die Bewegung von Objekten oder Personen in einem Video oder einer animierten Szene zu verfolgen und zu verfolgen. Dies kann in 2D oder 3D erfolgen.
WLAN	Wireless Local Area Network

1. Einführung in Deepfakes

1.1 Definition

Der Begriff Deepfake setzt sich aus den englischen Begriffen Deep Learning und Fake zusammen. Hierbei steht Deep Learning für eine Methode des maschinellen Lernens und Fake für eine Fälschung.
"Bei Deepfakes handelt es sich um einen Teilbereich synthetischer audiovisueller Medien: die Manipulation oder auch synthetische Erzeugung von Abbildungen, Videos und/oder Audiospuren menschlicher Gesichter, Körper oder Stimmen, zumeist mithilfe von KI."[1] [1]
Deepfakes werden mit Hilfe von künstlicher Intelligenz und Deep Learning Technologien erstellt, um Personen realistische Handlungen ausführen oder Worte sagen zu lassen, in Form von Video, Bild oder Audio. Es handelt sich hierbei um gefälschte Darstellungen, die möglichst realitätsnah dargestellt werden.[2]

1.2 Hintergrund

Deepfake ist eine Manipulationstechnik, die es Benutzern ermöglicht, das Gesicht einer Person mit einer anderen Person auszutauschen. Eine optimale Manipulation wird durch Verwendung mehreren Hunderten oder Tausenden Fotos der Zielperson erreicht. Das führt dazu, dass oft prominete Personen als Zielperson gewählt werden, da von ihnen viele Bilder im Internet existieren.
Bild- und Videomanipulationstechnologien bauen auf Techniken aus dem Bereich der künstlichen Intelligenz auf, welcher das Ziel verfolgt, menschliche Denkprozesse und Verhaltensweisen zu verstehen. Da maschinelles Lernen einem System ermöglicht aus Daten zu lernen, ist diese Technik wichtig für das erstellen von Deepfakes.
Deepfakes sind aus zwei Gründen beliebt: erstens wegen der Fähigkeit aus Daten wie Fotos und Videos, realistische Ergebnisse erzeugen zu können und zweitens die Verfügbarkeit der Technik, da diese für jeden leicht zu erreichen und durchzuführen ist. Es gibt Apps, welche die Schritte des Deepfakes-Algorithmus erklärt und so Personen mit wenig Kenntnissen über maschinelles Lernen oder Programmierung die Möglichkeit bietet ein Deepfake Bild oder Video zu erstellen.
Das führt zu einem Problem der heutigen Gesellschaft, da Deepfakes hauptsächlich aus Rache, Erpressung einer Person oder Verbreitung von Fake News einer höheren Person (bspw. eines Politikers) ausgenutzt werden.[3]

1.2.1 Geschichte

Das Manipulieren von Bildern wurde nicht erst in den letzten Jahren bekannt. Denn auch schon früher wurden Bilder zum Beispiel von Hitler, Stalin, oder Breschnew manipuliert, um so die Geschichte zu ihren Gunsten verändern zu können. Damals erforderte

es allerdings deutlich mehr Zeit und kompliziertere Techniken während der Fotoentwicklung in der Dunkelkammer, um ein Bild zu verfälschen. Doch durch die schnelle Entwicklung der Technologien wurde der Prozess ein Bild zu manipulieren zunehmend schneller. Anfangs begannen ausschließlich Forscher der 1990er Jahre die Entwicklung der Deepfake-Technologie zu übernehmen, diese wurde jedoch später von Amateuren in den Online-Communities unterstützt. Die Akademiker Christoph Bregler, Michele Covell und Malcolm Slaney entwickelten 1997 ein Programm, welches vorhandenes Videomaterial einer sprechenden Person anpassen konnte, dass diese Person die Wörter von einer anderen Audiospur nachahmte. Das Programm baut auf einer älteren Technologie auf, welches bereits Gesichter interpretieren, Audio aus Texten synthetisieren und Lippen im 3D-Raum modellieren konnte. Jedoch war dieses entwickelte Programm von den drei Akademikern das erste, welches alle Komponenten zusammenfügen und überzeugend animieren konnte. So war es möglich eine neue Gesichtsanimation aus einer Audioausgabe zusammenstellen zu können.
Zu Beginn der 2000er Jahre wurde die Entwicklung der Gesichtserkennung mit dem Computer immer weiter vorangetrieben, sodass es zu großen Verbesserungen der Technologie wie Motion Trackings kam, welche die heutigen Deepfakes so überzeugend machen.
In den Jahren 2016 und 2017 gab es zwei Projekt Veröffentlichungen. Einmal das Face2Face-Projekt der Technischen Universität München und einmal das Synthesizing Obama-Projekt der University of Washington.
Das Face2Face Projekt versucht Echtzeitanimationen zu erstellen, indem es den Mundbereich des Zielvideos durch einen Schauspieler ersetzt, während das Synthesizing Obama-Projekt sich damit beschäftigte Videomaterial des ehemaligen Präsidenten Barack Obama zu modifizieren.[4]
Im Jahr 2017 wurde das gefälschte Video des ehemaligen US-Präsidenten Barack Obama veröffentlicht und soll als Warnung der Technologie und deren potenziellen Auswirkungen gelten. Ende 2017 veröffentlichte ein Nutzer auf einer Webseite names Reddit pornografische Inhalte und behauptete, dass diese zu bekannten Personen wie zum Beispiel Taylor Swift oder Scarlett Johansson gehören. Auch wenn diese Bilder und Videos schnell wieder gelöscht wurden, erregte diese auf Deep Learning basierende Gesichtsersatztechnik die Aufmerksamkeit der Medien und verbreitete sich in vielen Internetforen. Alle Inhalte, die mit der Deepfake Technik zu tun hatten, wurden am 7. Februar 2018 auf fast allen Internetforen entfernt und verboten. Trotz des Verbots hat sich die Technik dennoch weiterhin durchgesetzt und wurde weltweit verbreitet. Bei der Person, die die Deepfake-Technik entwickelt hat, soll es sich um einen Software-Ingenieur handeln, der ein Entwicklungs-Kit herausbrachte, mit dem es einem Benutzer selbst ermöglicht, eigene manipulierte Bilder oder Videos zu erstellen. Durch die Hilfe von Open Source Tools und Funktionen von großen Softwareunternehmen wie NVidia und Google wurde die Deepfake-Technike entwickelt. Was bedeutet, dass für die Entwicklung technisches Wissen und Verständnis erforderlich sind, jedoch der Großteil der Software schon zuvor in der Öffentlichkeit zur Verfügung stand. Als klar wurde, dass selbst eine Person ohne viel Wissen in dem Gebiet, beliebig viele visuelle Medien manupulieren kann, wurde die Bedrohung der Deepfake-Technik ernst und das US-Verteidigungsministerium stellte sich

ein. Auch im Jahr 2018 wurde ein Deepfake Video von damaligen Präsidenten Donald Trump in den Medien hochgeladen, in dem die Belgier aufgefordert wurden, aus dem Pariser Klimaschutzabkommen auszusteigen.
Durch solche Veröffentlichungen der Deepfake Videos zeigte sich, dass die Technologie sich schnell weiterentwickelt und in der Lage ist einen großen Teil der Öffentlichkeit in die Irre führen zu können.[3]

1.3 Technische Grundlagen

1.4 Arten von Deepfakes

Deepfakes können in drei Hauptarten unterteilt werden: Video Deepfakes, Audio Deepfakes und Foto Deepfakes. Diese drei Arten lassen sich zusätzlich auch noch miteinander kombinieren.[4]

1.4.1 Video Deepfakes

Bei Video Deepfakes wird zusätzlich zwischen 3 Arten der Manipulation unterschieden. Auf welche Art der Manipulation zurückgegriffen wird, ist davon abhängig, was der Hauptgrund der Nutzung eines Video Deepfakes ist.[4]

Face Swapping

Eine der Arten ist das Face Swapping, bei dem die Gesichter auf Bildern oder Videos durch Fake Gesichter oder Gesichter anderer Personen, wie zum Beipsiel eines Promis, ersetzt wird. Dadurch ist es möglich die Person, dessen Gesicht verwendet wird, in einen anderen Kontext darstellen zu lassen, um beispielsweise in der Filmindustrie den Schauspieler mit einem Stunt Double austauschen zu können, um bestimmte Actionszenen realistischer wirken zu lassen.[4]

Face Morphing

Die zweite Art von Video Deepfakes ist das Face Morphing, welches ein Spezialeffekt ist, um ein Bild oder eine Form durch einen nahtlosen Übergang in ein anderes verändern zu können. Dieser Effekt wird oft in Filmen oder Animationen verwendet.[4]

Full body pupperty

Die letzte Art von Video Deepfakes ist die Full body pupperty, bei der einzelne Bewegungen bis hinzu komplette Bewegungsabläufe auf eine andere Person übertragen werden.
Die meisten Deepfakes benötigen viel Zeit für die Erstellung aufgrund der Systeme, welche erst mit dem Ausgangsmaterial trainiert werden müssen, um danach Inhalte verändern zu können. Es gibt aber auch Deepfake-Methoden die in Echtzeit funktionieren, welche die Möglichkeit bietet, Mimik und Lippenbewegungen einer Person zu erkennen und diese anschließend in Echtzeit auf das Videobild einer anderen Person übertragen zu lassen.[4]

1.4.2 Audio Deepfakes

Eine andere Art von Deepfakes sind Audio Deepfakes, bei dem aufgenommene oder live Audio Dateien verändert werden. Wobei hier zwischen Voice Swapping und Text to Speech unterschieden wird.[4]

Voice-swapping

Bei dem Voice-swapping können Audioinhalte so verändert werden, dass ein Text von einer fremden Person gesprochen werden kann. Die Stimme kann mit verschiedenen Effekten verändert werden, sodass zum Beispiel eine Stimme jünger, älter, männlich, weiblich oder auch mit verschiedenen Dialekten versehen werden kann. Dadurch wird dem Höhrer vorgespielt, dass verschiedene Personen sprechen, wobei es sich aber nur um eine Person hält.[4]

Text to Speech

Beim Text to Speech können Audioinhalte einer Aufnahme durch Eingabe eines neuen Textes verändert werden. Dadruch können zum Beispiel falsch ausgesprochene Wörter im nachhinein ersetzt werden, ohne eine neue Aufnahme durchführen zu müssen.[4]

1.4.3 Foto Deepfakes

Die dritte Art der Deepfakes sind Foto Deepfakes, bei denen es sich darum handelt, Fotos zu manipulieren. Dadurch können Fotos nach belieben verändert werden, um beispielsweise eine Person auf dem Bild durch einen Alterungsfilter, den Alterungsprozess der Person dargestellt werden kann.[4]

Face and body-swapping

Mithilfe des Deepfake-Algorithmus, welcher auch bei den anderen Arten verwendet wird, können Änderungen an einem Gesicht und Körper gemacht werden, indem das Gesicht oder der Körper mit einer anderen Person ausgetauscht wird. Eine mögliche Anwendung hierfür wäre das virtuelle anprobieren einer Brille, Haarfarbe oder Kleidung.[4]

1.4.4 Kombination aus Audio und Video Deepfake

Zuletzt gibt es wie oben eine mögliche Kombination der verschiedenen Arten, wie zum Beispiel die Kombination aus Audio und Video Deepfake. Diese Kombination wird auch das Lip-syncing genannt, bei dem Mundbewegungen sowie die gesprochenen Wörter in einem Video verändert und synchronisiert werden. Dadurch ist es möglich eine Person in einem Video scheinbar etwas sagen zu lassen, was sie aber niemals gesagt hat. Dies kann sowohl stark Missbraucht werden, indem zum Beispiel einem Politiker eine falschaussage untergeschoben wird. Es kann aber auch für positive Sachen Verwendung finden, um beispielsweise einen Film oder Werbung in eine andere Sprache zu synchronisieren.[4]

1.5 Anwendungsgebiete

1.5.1 Positive Anwendungsgebiete

1.5.2 Negative Anwendungsgebiete

1.6 Ethik

Test WLANs huhu LAN (Local Area Network)

Literatur

[1] M. Block. (29. Aug. 2023). „Definition und Anwendungsbereiche“ [Online]. Verfügbar: https://link.springer.com/chapter/10.1007/978-3-662-67427-7_2

[2] L. Whittaker. (Juli 2023). „Mapping the deepfake landscape for innovation: A multidisciplinary systematic review and future research agenda“ [Online]. Verfügbar: https://www.sciencedirect.com/science/article/pii/S0166497223000950#abs0015

[3] J. A. Marwan Albahar. (30. Nov. 2019). „DEEPFAKES: THREATS AND COUNTERMEASURES SYSTEMATIC REVIEW“ [Online]. Verfügbar: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.jatit.org/volumes/Vol97No22/7Vol97No22.pdf

[4] J.-T. Kötke. (Feb. 2021). „DEEPFAKE -EINE KURZE EINLEITUNG Deepfake -Eine kurze Einleitung“ [Online]. Verfügbar: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.researchgate.net/profile/Jennifer-Tia-Koetke/publication/373041489_DEEPFAKE_-EINE_KURZE_EINLEITUNG_Deepfake_-Eine_kurze_Einleitung/links/64d4ffddd3e680065aac7ee3/DEEPFAKE-EINE-KURZE-EINLEITUNG-Deepfake-Eine-kurze-Einleitung.pdf