

Deep Fake Detection: Was ist ein "reales" Bild?



Die Erkennung von Deepfakes ist von entscheidender Bedeutung, um die negativen Auswirkungen manipulierter Videos und Bilder zu minimieren. Deepfakes können genutzt werden, um Desinformation zu verbreiten, Personen zu erpressen, Rufschädigung zu betreiben oder Vertrauen in mediale Inhalte zu untergraben. Ihre Verbreitung stellt eine wachsende Bedrohung für die Gesellschaft dar, insbesondere in Bereichen wie Politik, Wirtschaft oder sozialen Medien. Um effektive Erkennungsmethoden zu entwickeln, sind gute Benchmarks unerlässlich. Sie bieten eine standardisierte Grundlage, um die Leistungsfähigkeit von Algorithmen objektiv zu messen, Schwächen aufzudecken und verschiedene Ansätze miteinander zu vergleichen. Ohne

robuste Benchmarks wäre es nahezu unmöglich, Fortschritte zu bewerten, neue Herausforderungen durch immer realistischere Deepfakes zu adressieren und die Technologie langfristig sicherer zu gestalten. In diesem Projekt soll ein geeigneter Benchmark zur Evaluation von Detektions-Algorithmen entwickelt werden, der die bisherigen Biases [1] vermeidet und so die Entwicklung von praxistauglichen Algorithmen ermöglichen soll.

Betreuer

Prof. Dr.-Ing. Janis Keuper

- janis.keuper@hs-offenburg.de
- <https://www.keuper-labs.org>

Beteiligte Institute und Firmen

Das Projekt wird am Institute for Machine Learning and Analytics durchgeführt.

Ziele des Projekts

- Systematische Untersuchung aktueller Deep Fake Detection Methoden und deren Evaluation
- Erstellungen eines neuen, praxistauglichen Benchmarks
- Wissenschaftliche Publikation der eigenen Ergebnisse

Diese Werkzeuge/Qualifikationen werden erlernt

- Verständnis und Anwendung von generativen Modellen
- Praktische Entwicklung mit Python und Deep Learning Frameworks
- State of the Art Deep Fake Detection Methoden
- Evaluations Methoden

Literatur + Weiterführende Informationen

- IMLA: <https://imla.hs-offenburg.de/>
- [1] Grommelt, P., Weiss, L., Pfreundt, F. J., & Keuper, J. (2024). Fake or JPEG? Revealing Common Biases in Generated Image Detection Datasets. DMLR Workshop, ECCV 24