

Deepfake-Erkennung für digitale Medien

Erkennung KI-generierter multimedialer Inhalte

Hausarbeit / Bachelorarbeit / Masterarbeit

Generative Modelle erzeugen täuschend echte Fälschungen. Dies stellt nahezu jede Branche, in der verlässliches Material verarbeitet werden soll, vor extreme Schwierigkeiten. In dieser Arbeit sollen derzeitige Ansätze zur Erkennung von KI-generierten Multimedia-Objekten untersucht, verglichen und evaluiert werden. Hierzu wird jeweils ein spezifischer Multimedia-Typ (z.B. Bild, Ton, Video, Text) als Schwerpunkt der Arbeit gewählt und eine Übersicht der derzeitigen Verfahren zur Erstellung von Deepfakes und auch zu deren Erkennung gegenübergestellt. In passenden Experimenten wird die Qualität sowohl der Erzeugung als auch der Erkennung überprüft.

Mögliche Forschungsfragen (Auswahl):

- Wie können Deepfakes eines bestimmten Multimedia-Typs erstellt und verlässlich erkannt werden
- Wie verlässlich sind Erkennungsalgorithmen für Deepfakes eines bestimmten Multimedia-Typs
- Wie lassen sich Erkennungsalgorithmen für Deepfakes in Anwendungen integrieren

Hintergrund: aktuelle Studien zeigen raschen Fortschritt bei der künstlichen Generierung digitaler Inhalte. Die Algorithmen und Verfahren zur Erkennung hinken derzeit hinterher. Allerdings gibt es auch neue Ansätze zur Detektion.

Prof. Dr.-Ing. Stefan Wagenpfeil

Wirtschaftsinformatik | Software-Engineering

E-Mail: s.wagenpfeil@pfh.de

Internet: www.stefan-wagenpfeil.de/pfh

Termine: www.stefan-wagenpfeil.de/termine

PFH Private University of Applied Sciences

Weender Landstraße 3-7

37073 Göttingen