17. Juni 2020

Thesenpapier zum innenpolitischen Frühstücksgespräch „Polizeiliche Gesichtserkennung“

Prof. Dr.-Ing. Florian Gallwitz, Fakultät Informatik, TH Nürnberg

1. Hintergrund
Ich vertrete in der Informatik schwerpunktmäßig das Fachgebiet Mustererkennung mit den Schwerpunkten Bild- und Spracherkennung, maschinelles Lernen und Deep Learning. Das Thema Gesichtserkennung ist eine Anwendung dieser Technologien mit weitreichenden gesellschaftlichen Folgen. Die Entwicklungen auf diesem Gebiet und auch die Pilotstudie am Berliner Südkreuz habe ich deshalb intensiv verfolgt.

2. Zum Stand der Technik bei der Gesichtserkennung
Bei der automatischen Gesichtserkennung sind in den vergangenen Jahren dramatische Fortschritte zu beobachten. Das zeigt sich in Forschungspublikationen und auch in den Vergleichstests kommerziell verfügbarer Systeme, die die US-Standardisierungsbehörde NIST regelmäßig durchführt.

Einen wesentlichen Anteil an dieser Entwicklung hat der Einsatz von künstlichen Neuronalen Netzen, sogenannten Deep-Learning-Verfahren. Hinzu kommt die Verfügbarkeit von teils enorm großen Datenmengen für das Training der Systeme. Zum Teil wurden hierzu mehrere hundert Millionen Bilder von Gesichtern genutzt.

Fertig trainierte Gesichtserkennungssysteme rechnen Bilder von beliebigen Gesichtern, also auch solchen, die nicht in den Trainingsdaten enthalten waren, in mathematische Repräsentationen dieser Gesichter um, sogenannte „Face Embeddings“. Der mathematische Abstand dieser Face Embeddings entspricht der Ähnlichkeit von Gesichtern. Unterschreitet dieser Abstand für ein gegebenes Bildpaar einen geeigneten Schwellenwert, so kann man davon ausgehen, dass beide Bilder mit hoher Wahrscheinlichkeit die gleiche Person zeigen. Auf dieser Basis lassen sich gleichermaßen Verifikationssysteme (z.B. zur Zugangskontrolle) und Erkennungssysteme (z.B. zur Fahndung) realisieren.

Die besten verfügbaren Gesichtserkennungssysteme sind bei der Beurteilung der Frage, ob zwei Bilder die gleiche Person zeigen, weitaus genauer als der Mensch. Selbst sogenannte Super Recognizer, die für die Gesichtserkennung eine besondere Begabung besitzen, können mit den leistungsfähigsten Systemen nicht mithalten.

Dennoch sind Gesichtserkennungssysteme bei weitem nicht fehlerfrei und werden das auch nie sein. Wie die menschliche Fähigkeit zur Erkennung von Gesichtern wird auch die Erkennungsgenauigkeit automatischer Systeme durch zahlreiche Faktoren negativ beeinflusst. Hierzu gehören schlechte Lichtverhältnisse, die teilweise Verdeckung des Gesichts (Schal, Sonnenbrille, Schirmmütze, Bart), Alterung, ungünstige Perspektiven, unterschiedliche Kameraoptiken (Tele vs. Weitwinkel) sowie unterschiedliche Kameraauflösungen.

3. Zur Studie am Berliner Südkreuz

In der Studie wurden drei Gesichtserkennungssysteme unter Mitwirkung von freiwilligen Probanden unter praxisnahen Bedingungen getestet. Zwei der Systeme erzielten eine auf den ersten Blick brauchbare Erkennungsgenauigkeit, jedoch bei einer für einen praktischen Einsatz aus meiner Sicht viel zu hohen Rate an Fehlalarmen. Durch UND-Verknüpfung der beiden besten Systeme ließe sich die Falsch-Positiv-Rate mit vertretbaren Einbußen bei der Erkennungsrate deutlich reduzieren.

Die in der Studie ermittelten Erkennungsraten sind jedoch aus meiner Sicht mit Vorsicht zu betrachten. Sie sind aus mehreren Gründen nicht auf einen praktischen Einsatz zur Fahndung nach Kriminellen übertragbar:

- In der deutlich erfolgreicheren „Testphase 2“ standen für jeden gesuchten Probanden Bilder aus den Videoströmen der gleichen Kameras zur Verfügung, die dann auch zur Gesichtserkennung genutzt wurden. Damit wurden gleich mehrere Freiheitsgrade, die die Erkennungsgenauigkeit von Gesichtserkennungssystemen beeinträchtigen, ausgeschaltet bzw. erheblich gemindert: Kameraoptik, Kameraperspektive, Kameraauflösung und Lichtverhältnisse. Es ist vollkommen unrealistisch anzunehmen, dass für gesuchte Kriminelle Referenzbilder aus jeder Kamera vorliegen, mit der nach ihnen gefahndet werden soll.
- Für jeden Probanden standen gleich mehrere dieser Bilder zur Verfügung (zwei bis fünf). Das erleichtert die Erkennung. Vermutlich wurden diese zudem so ausgewählt, dass diese variable Beleuchtungsverhältnisse und Perspektiven beinhalteten, was die Erkennung weiter vereinfacht.
- Alle Referenzbilder entstanden jeweils kurz vor der Durchführung der beiden Testphasen, wodurch Alterungseffekte ausgeschlossen sind und Merkmale wie Bart und Brille in den Referenzbildern eine hohe Übereinstimmung mit den Testaufnahmen besitzen. Bei der Fahndung etwa auf Basis eines fünfzehn Jahre alten Pass- oder Führerscheinfotos wäre mit einer Vervielfachung der Fehlerrate zu rechnen.
- Die Teilnehmer des Pilottests haben vermutlich bewusst oder unbewusst häufiger in Richtung der installierten Kameras geschaut. Ein gesuchter Krimineller, der von der Existenz eines solches Fahndungssystems Kenntnis hat, würde dagegen den Blick in die Kameras zu vermeiden suchen. Stattdessen könnte er sich einer Erkennung mit einfachen Mitteln entziehen, etwa durch das Tragen einer Schirmmütze und/oder einer Sonnenbrille und durch konsequentes Nach-unten-schauen, etwa durch Blick auf sein Smartphone.

Im Ergebnis wäre die Erkennungsrate im realen Einsatz erheblich niedriger als die in der Studie ermittelten Werte, der Anteil falsch-positiver Treffer unter den gemeldeten Treffern erheblich höher.

4. Zur Gesichtserkennung als Fahndungsinstrument

Sowohl die hohe Genauigkeit als auch die verbleibende Ungenauigkeit bieten Anlass für Kritik an Gesichtserkennungssystemen:

- Die Möglichkeit, Personen auf größere Entfernung zu identifizieren, ohne dass diese davon etwas mitbekommen, unterscheidet Gesichtserkennung grundlegend von anderen biometrischen Erkennungsverfahren. Verknüpft man, wie die amerikanische Firma Clearview, Gesichtserkennungssysteme mit im Internet verfügbaren Bildern zu einer Art Gesichts-Suchmaschine, so eröffnet sich die Möglichkeit, Menschen in der Öffentlichkeit per Smartphone-App auf Knopfdruck zu identifizieren. Der Staat könnte alternativ durch Zugriff auf die gespeicherten Passfotos der Bürger ein umfassendes Überwachungssystem realisieren. Die Möglichkeit, sich anonym in der Öffentlichkeit zu bewegen, würde damit der Vergangenheit angehören.
- Besonders bei der Echtzeit-Überwachung stellen Fehlalarme von Gesichtserkennungssystemen ein erhebliches Problem dar. Würde ein solches System etwa dazu genutzt, in einem Südkreuz-ähnlichen Szenario nach Schwerkriminellen oder Terroristen zu fahnden, so wäre deutschlandweit täglich mit Dutzenden von Fehlalarmen zu rechnen. Der Ausschluss solcher falsch-positiver Treffer würde i.d.R. nicht durch den Blick eines Polizeibeamten auf das Kamerabild erfolgen können, denn das Gesichtserkennungssystem ist dem Menschen in dieser Hinsicht weit überlegen. Es ist also zu erwarten, dass die fälschlich erkannte Person der gesuchten Person tatsächlich sehr ähnlich sieht. Stattdessen müsste jeweils eine Personenkontrolle erfolgen, unter der Annahme, dass es sich

tatsächlich um den gesuchten, vielleicht bewaffneten Schwerkriminellen handeln könnte. Ein großer Teil, vermutlich sogar der weit überwiegende Anteil dieser Kontrollen, würde jedoch zu Unrecht erfolgen.

5. Fazit

Gesichtserkennungssysteme besitzen mittlerweile eine beeindruckende Erkennungsgenauigkeit, die die des Menschen deutlich übersteigt. Das macht sie zu einem attraktiven Fahndungsinstrument, aber auch zu einer bedrohlichen Überwachungstechnologie.

Die trotz aller technischen Fortschritte verbleibenden Fehlalarme sind insbesondere dann ein Problem, wenn die Fahndung wie am Bahnhof Südkreuz in Echtzeit erfolgen soll.

Eine strenge gesetzliche Regulierung des Einsatzes von Gesichtserkennungssystemen sowohl durch staatliche Stellen als auch durch Privatpersonen und Firmen erscheint mir sinnvoll und geboten.