

„Erfolgreiche“ Gesichtserkennung mit Hunderttausenden Fehlalarmen

Dortmund/
Essen/Berlin,
30. Oktober 2018

Die Unstatistik des Monats Oktober ist die [Pressemitteilung](#) des Innenministeriums über das „erfolgreiche“ Projekt zur automatischen Gesichtserkennung mit einer Trefferrate von über 80% und Falsch-Alarm-Rate von unter 0,1%.

Das Bundesministerium des Innern, für Bau und Heimat führte im Rahmen des gemeinsamen Pilotprojektes "Sicherheitsbahnhof Berlin Südkreuz" gemeinsam mit der Bundespolizei und der Deutsche Bahn AG einen Test von drei Gesichtserkennungssystemen am Bahnhof Südkreuz in Berlin durch. Das Bundeskriminalamt beriet hierbei. Bundesinnenminister Horst Seehofer war begeistert: „Die Systeme haben sich in beeindruckender Weise bewährt, so dass eine breite Einführung möglich ist“. tagesschau.de wiederholte am 11. Oktober kritiklos zwei Zahlen, die den Erfolg des Projekts belegen sollten: Erstens, eine Trefferrate 80%, das heißt, von je 10 Gesuchten (Testpersonen) wurden 8 richtig erkannt und 2 nicht. Zweitens, eine Falsch-Alarm-Rate von 0,1%, das heißt, von je 1.000 Normalbürgern (die nicht als Testpersonen teilnahmen) wurden 999 als unverdächtig und einer fälschlicherweise als gesuchte Person klassifiziert.

Das klingt doch gut! Also so schnell wie möglich auf allen Bahnhöfen Kameras mit Gesichtserkennung installieren, um uns vor den rund 600 als gefährlich eingestuftem Islamisten und anderen gesuchten Personen zu schützen. Oder gleich wie in China die totale Überwachung aller Bürger auf allen Straßen, öffentlichen Plätzen und Gebäuden einführen?

Als Treffer zählte schon, wenn eines der drei Systeme anschluss

Ein genauer Blick in den Abschlussbericht zeichnet ein anderes Bild. Keines der drei getesteten Systeme hat eine Trefferquote von 80% über beide Testphasen hinweg erreicht; die 80% sind auch nicht der Durchschnitt der drei Systeme. Diese Zahl bekam man nur nachträglich, indem man die Treffer aller drei Systeme addierte (d.h., wenn mindestens eines der Systeme einen Treffer hatte, galt das als Treffer). Die Bilder der „Gesuchten“, die den Systemen zur Verfügung standen, waren anders als Fahndungsfotos so gut wie perfekt: in der ersten Testphase mit 312 Freiwilligen wurden diese mit hochauflösenden Kameras aufgenommen, in der zweiten Phase, an der nur noch 201 Testpersonen weitermachten, wurden dann sogar die von den Überwachungskameras in der ersten Phase aufgenommenen Bilder verwendet, das heißt, am gleichen Ort, wo getestet wurde – damit bekam man bessere Werte als man sie in Realität (in der

man keine erste Testphase hat) je erreichen könnte. Die Zahlen beruhen auch nicht auf allen erhobenen Daten, sondern nur auf einer kleinen Auswahl, wobei nicht angegeben wurde, wie diese ausgesucht wurden. All das macht eine Bewertung unklar.

Der eigentlich interessante Punkt ist aber: Das Problem bei Massenüberwachungssystemen sind Fehlalarme, das heißt, Normalbürger, die vom Gesichtserkennungssystem für Gesuchte gehalten werden. Eine Person betritt einen Bahnhof und wird vom System als „Verdächtiger“ gemeldet. Was ist die Wahrscheinlichkeit, dass sie tatsächlich der Verdächtige ist? Eben nicht 80%. Und auch nicht 0,1%.

Eine Überschlagsrechnung kann diese Frage beantworten. Die Deutsche Bahn berichtet, dass rund 11,9 Millionen Menschen täglich mit der Bahn reisen. Der Abschlussbericht der Studie gibt an, es gäbe derzeit etwa 600 Verdächtige, die als islamistische Gefährder eingestuft werden, welche das System erkennen sollte. Um die Rechnung einfach zu machen, nehmen wir an, dass sich davon täglich 100 an Bahnhöfen aufhalten (etwa der gleiche Anteil wie bei der Normalbevölkerung).

Von den 100 Verdächtigen erwarten wir, dass etwa 80 (80%) erkannt werden, und von den rund 11,9 Millionen anderen Menschen werden täglich nochmals etwa 11.900 (0,1%) fälschlicherweise als gesuchte Person eingestuft. Das heißt, die Wahrscheinlichkeit, dass jemand eine gesuchte Person ist wenn die Kamera Alarm schlägt, ist etwa $80/11.980$, also rund 7 in 1.000, oder 0,7%. In anderen Worten, etwa 99,3% der Einschätzungen des Systems sind falsch.

Monatlich mehr als 350.000 unnötige Personenkontrollen durch Fehlalarme

Jedem Alarm solle ein Einsatz der Polizei zur Abklärung folgen. Da eine flächendeckende Überwachung an Bahnhöfen jedoch 11.900 falsche Alarme pro Tag erwarten ließe, müssten jeden Monat über 350.000 Personen unnötigerweise kontrolliert werden. Das wäre nicht nur enorm aufwändig und teuer, unsere Bahnhöfe würden auch bald wie die Kontrollen an Flughäfen aussehen. Falls statt 600 nach doppelt so vielen Personen gefahndet wird, ändert sich das Ergebnis nur gering: etwa 98,7% der Einschätzungen des Systems sind falsch.

Es gibt noch eine dritte irreführende Zahl im Bericht und der Pressemitteilung: Es wird uns gesagt, dass sich die Falsch-Alarm-Rate durch Zusammenführung aller drei Systeme auf „bis zu 0,00018%“ reduzieren lässt. Was uns nicht gesagt wird, ist, dass dadurch die Trefferrate deutlich reduziert würde. Der ängstliche Bürger merkt's ja nicht.

Diese Unstatistik zeigt, wie ein bisschen statistisches Denken Politikern, Ministerialen, Journalisten und Bürgern helfen könnte, sich nicht durch verdrehte Zahlen beeindrucken zu lassen. Dass das Gesichtserkennungssystem 99,3% der Verdächtigen zu Unrecht verdächtigt, ist mit der Regel von Bayes (ein Satz aus der Wahrscheinlichkeitsrechnung) berechnet, die in Bayern inzwischen jeder Schüler der 11. Klasse lernt. Im Bericht und der Pressemitteilung ist sie nicht zu finden.

Wohin der Schutz des Bürgers durch den Staat führen kann, zeigen die über 300 Millionen Überwachungskameras mit Gesichtserkennungssoftware in China.

Ansprechpartner:

Prof. Dr. Gerd Gigerenzer,
Sabine Weiler (Pressestelle RWI),

Tel.: (030) 82406-0
Tel.: (0201) 8149-213

Mit der „Unstatistik des Monats“ hinterfragen der Berliner Psychologe Gerd Gigerenzer, der Dortmunder Statistiker Walter Krämer und RWI-Vizepräsident Thomas K. Bauer jeden Monat sowohl jüngst publizierte Zahlen als auch deren Interpretationen. Alle „Unstatistiken“ finden Sie im Internet unter www.unstatistik.de.