FORSCHUNG KOMPAKT

KI für die Anomalieerkennung

Manipulation automatisierter Gesichtserkennung verhindern

Von der Entsperrung von Smartphones bis zu schnelleren Einlasskontrollen am Flughafen: Die Verbreitung der automatisierten Gesichtserkennung zur Identifikation von Personen nimmt zu. Doch diese Authentifizierungsmethode ist anfällig gegenüber Morphing-Angriffen: Kriminellen bietet sie die Möglichkeit, zwei Gesichtsbilder zu einem zu verschmelzen. Wird ein Reisepass mit einem derart manipulierten Foto ausgestattet, können zwei Personen den gleichen Ausweis nutzen. Fraunhofer-Forscherteams entwickeln gemeinsam mit Partnern ein System, das diese Art von Angriffen vereitelt. Dabei bedienen sie sich Methoden des maschinellen Lernens.

Wer regelmäßig in die USA reist, ist es gewohnt, bei der Passkontrolle in eine Kamera schauen zu müssen. Das elektronische Foto wird blitzschnell mit dem im biometrischen Pass gespeicherten Bild verglichen. Bei dieser biometrischen Gesichtserkennung ermittelt ein Programm die digitalen Daten des Live-Bildes, vergleicht sie mit den Daten des Chip-Bildes und kann so feststellen, ob die Gesichter auf den Fotos in den individuellen Merkmalen übereinstimmen. Auch Smartphones und Tablets lassen sich durch Gesichtserkennung entsperren. Die Methode soll den Zugriff auf sensible Daten und unberechtigten Personen den Zutritt verweigern. Doch sie ist anfällig gegenüber gezielten Angriffen, wie bereits mehrfach in Tests nachgewiesen wurde. »Kriminelle sind in der Lage, die Gesichtserkennungssysteme – wie sie auch bei der Grenzkontrolle eingesetzt werden – so auszutricksen, dass zwei Personen denselben Pass verwenden können«, weiß Lukasz Wandzik, Wissenschaftler am Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik IPK in Berlin. Gemeinsam mit seinen Kolleginnen und Kollegen vom Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, HHI und weiteren Partnern (siehe Kasten) entwickelt er Verfahren, die Bildanomalien erkennen, die bei der digitalen Bildverarbeitung in Morphing-Prozessen auftreten. »Der Morphing-Angriff wird ausgeführt, indem zwei Gesichtsbilder zu einem synthetischen Gesichtsbild verschmolzen werden, das die Eigenschaften beider Personen enthält«, erklärt Wandzik. Mit diesem Foto im Reisepass werden beide Personen durch ein biometrisches Gesichtserkennungssystem authentifiziert.

Kontakt
Janis Eitner | Fraunhofer-Gesellschaft, München | Kommunikation | Telefon +49 89 1205-1333 | presse@zv.fraunhofer.de
Claudia Engel | Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik IPK | Telefon +49 30 39006-238 | Pascalstraße 8 - 9 | 10587 Berlin | www.ipk.fraunhofer.de | claudia.engel@ipk.fraunhofer.de
Anne Rommel | Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, HHI | Telefon +49 30 31002-353 | Einsteinufer 37 | 10587 Berlin | www.hhi.fraunhofer.de | anne.rommel@hhi.fraunhofer.de

Die Attacken könnten beispielsweise vor oder beim Beantragen von Ausweisen stattfinden. Im Projekt ANANAS, kurz für »Anomalie-Erkennung zur Verhinderung von Angriffen auf gesichtsbildbasierte Authentifikationssysteme«, widmen sich die Partner diesem Problem, indem sie simulierte Bilddaten analysieren und erforschen. Dabei werden moderne Methoden der Bildverarbeitung und des maschinellen Lernens angewandt, insbesondere tiefe neuronale Netze, die explizit für die Verarbeitung von Bilddaten konzipiert wurden. Diese komplexen Netze bestehen aus zahlreichen Ebenen, die in vielschichtigen Strukturen miteinander verknüpft sind. Sie beruhen auf Verbindungen zwischen mathematischen Berechnungseinheiten und bilden die Neuronenstruktur des menschlichen Gehirns nach.

Identitätsdiebstahl mit neuronalen Netzen vermeiden

Um die zu entwickelnden Verfahren und Systeme testen zu können, erzeugen die Projektpartner im ersten Schritt die Daten, mit denen die bildverarbeitenden Programme trainiert werden, um Manipulationen zu erkennen. Hierfür werden verschiedene Gesichter zu einem gemorpht. »Um zu entscheiden, ob ein Gesichtsbild authentisch ist oder durch einen Morphing-Algorithmus erstellt wurde, haben wir tiefe neuronale Netze auf gemorphte und reale Gesichtsbilder trainiert. Diese können manipulierte Bilder anhand der dadurch entstehenden Veränderungen erkennen, speziell auch in semantischen Bereichen wie in Gesichtsmerkmalen oder Glanzlichtern in den Augen«, erläutert Prof. Peter Eisert, Abteilungsleiter Vision & Imaging Technologies am Fraunhofer HHI, die Vorgehensweise.

LRP-Algorithmen machen KI-Prognosen erklärbar

Die neuronalen Netze entscheiden sehr zuverlässig, ob es sich um echte oder gefälschte Bilder handelt, die Trefferquote bei den im Projekt erstellten Testdatenbanken liegt bei über 90 Prozent. »Das Problem ist jedoch vielmehr, dass man nicht weiß, wie das Neuronale Netz die Entscheidung getroffen hat«, sagt Eisert. Daher interessiert die Forscherinnen und Forscher am Fraunhofer HHI neben einer Entscheidung über die Echtheit eines Bildes auch der Entscheidungsgrund. Zu diesem Zweck analysieren sie mit eigens entwickelten LRP-Algorithmen (Layer-Wise Relevance Propagation) die Regionen im Gesichtsbild, die für die Entscheidung relevant sind. Dies hilft, verdächtige Bereiche in einem Gesichtsbild zu finden, und Artefakte zu identifizieren und zu klassifizieren, die während eines Morphing-Prozesses erzeugt wurden. Mithilfe der Algorithmen lassen sich gemorphte Bilder als solche erfolgreich identifizieren, wie erste Referenztests bestätigen. Die Gesichtsbereiche, die relevant sind für die Entscheidung, kennzeichnet die LRP-Software entsprechend. Häufig geben die Augen Hinweise, ob es sich um eine Fälschung handelt.

Die Forscher nutzen diese Informationen auch, um die neuronalen Netze robuster zu gestalten, um unterschiedlichste Angriffsmethoden erkennen zu können. »Die Kriminellen können auf immer ausgefeiltere Angriffsmethoden zurückgreifen, zum Beispiel auf KI-Verfahren, die komplett künstliche Gesichtsbilder erzeugen. Indem wir unsere

neuronalen Netze optimieren, versuchen wir, den Fälschern einen Schritt voraus zu sein und zukünftige Attacken zu identifizieren«, sagt der Professor für Informatik.

Eine Demonstrator-Software inklusive Anomalieerkennung und Auswertungsverfahren liegt bereits vor. Sie umfasst verschiedene, miteinander fusionierte Detektormodule der einzelnen Projektpartner. Die vernetzten Module wenden unterschiedliche Erkennungsverfahren an, um eine Manipulation zu ermitteln, woraus am Ende des Prozesses ein Gesamtresultat erzeugt wird. Ziel ist es, die Software in bestehende Gesichtserkennungsysteme an Grenzkontrollen zu integrieren beziehungsweise diese um die Morphingkomponenten zu erweitern und so Fälschungen durch entsprechende Angriffe auszuschließen.

Projekt ANANAS

Anomalie-Erkennung zur Verhinderung von Angriffen auf gesichtsbildbasierte Authentifikationssysteme

Projektpartner:
- Bundesdruckerei GmbH, Berlin
- DERMALOG Identification Systems GmbH, Hamburg
- Fraunhofer-Institut für Produktionsanlagen und Konstruktionstechnik IPK, Berlin
- Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut, HHI, Berlin
- Otto-von-Guericke-Universität, Magdeburg (Verbundkoordinator)

Fördersumme:
4,5 Mio. Euro, davon 74 Prozent Förderanteil durch das Bundesministerium für Bildung und Forschung BMBF

Projektlaufzeit:
Juni 2016 bis Mai 2020

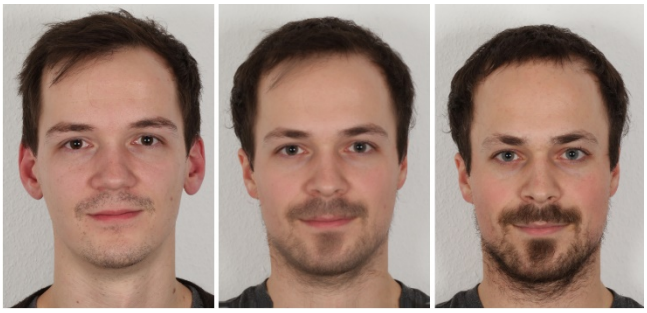

Abb. 1 Illustration eines Face-Morphing-Angriffs. Von links nach rechts: Originalbilder, Mitte: Morphing-Angriff