Täuschend echte Bilder von Personen, **die es nicht gibt**

Neuronale Netze in selbstfahrenden Autos brauchen riesige Mengen an Aufnahmen aus dem Straßenverkehr, um zu lernen, wie man Menschen erkennt. Deren Verwendung setzt aber das Einverständnis der abgebildeten Personen voraus, was einen enormen Aufwand bedeutet. Eine Forschungsgruppe um Prof. Laura Leal-Taixé hat für dieses Problem verblüffende Lösungen gefunden.

Link
dvl.in.tum.de/team/lealtaixe

Bildquelle: Magdalena Jooss, L. Leal-Taixé

Full Article (PDF, EN): www.tum.de/faszination-forschung-28

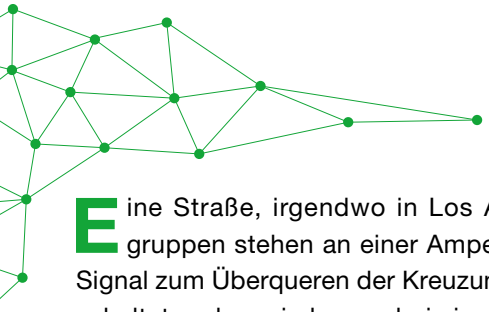
Deceptively Realistic Images of People Who Don't Actually Exist

E

No field of artificial intelligence is untouched by the impact of neural nets: they have become the workhorses of AI in recent years, learning to perform tasks by training with vast quantities of data. This often involves images of people, such as when an autonomous vehicle needs to learn to identify people and avoid them. Yet this type of data is also sensitive and the people in these images need to consent to its use. Prof. Laura Leal-Taixé and her team are researching ways to generate images of people without raising any data protection concerns. She is pursuing numerous approaches, one of which involves replacing real people in images with computer-generated people who do not exist in reality. Another approach is using fully computer-generated images. The researcher draws on the popular video game Grand Theft Auto V to generate images with an astounding level of realism. □

Prof. Laura Leal-Taixé

leitet die Abteilung Dynamic Vision and Learning an der TUM. Die aus Barcelona stammende Computerwissenschaftlerin schrieb ihre Masterarbeit in Boston und promovierte an der Leibniz Universität Hannover, bevor sie in Michigan und Zürich forschte. 2016 wechselte sie an die TUM und gewann ein Jahr darauf den Sofja Kovalevskaja-Preis sowie 2020 den Google Faculty Award. Ausgleich findet sie im Sport: In München hat Leal-Taixé das Bouldern für sich entdeckt.



Eine Straße, irgendwo in Los Angeles. Fußgängergruppen stehen an einer Ampel und warten auf das Signal zum Überqueren der Kreuzung. Als die Ampel umschaltet, gehen sie los, wobei sie einander ausweichen, um die andere Seite zu erreichen. Es sind auf den ersten Blick ganz normale Menschen, wie sie in einer Großstadt leben, unterschiedlich in Kleidung, Alter, Geschlecht und Hautfarbe. Dass sie beobachtet werden, scheinen sie nicht zu bemerken.

Die Person, die ihre Aktivitäten auf dem Bildschirm verfolgt, ist die Computerwissenschaftlerin Prof. Laura Leal-Taixé. Gemeinsam mit ihrem Team erstellt sie einen umfangreichen Datensatz, mit dem intelligente Computerprogramme lernen können, Menschen in Videobildern zu erkennen und zu verfolgen – eine Schlüsselfähigkeit für selbstfahrende Autos. Dazu sind große Mengen detaillierter Aufnahmen von realistischen städtischen Umgebungen nötig, und tatsächlich sind die Bilder, die über Leal-Taixés Bildschirm flimmern, von hoher Qualität. Sogar die Gesichter sind gut erkennbar. Doch wie steht es um die Persönlichkeitsrechte der abgebildeten Menschen? Niemand von ihnen hat sein Einverständnis dafür gegeben. Bei großen Menschengruppen, wie sie neuronale Netze benötigen, wäre das Einholen von Einverständniserklärungen ein enormer logistischer Aufwand. Und was geschieht, wenn eine Person auf dem Bild die Einwilligung verweigert?

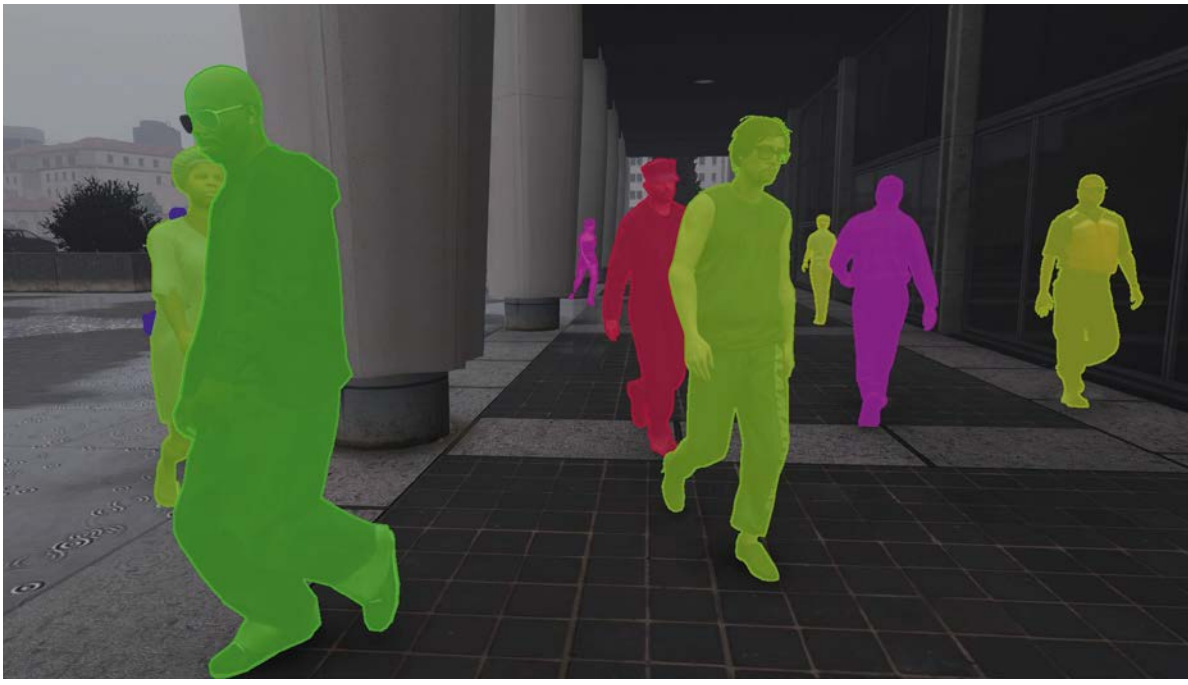
Mit derlei Spitzfindigkeiten muss sich Leal-Taixé nicht aufhalten, denn bei genauerer Betrachtung zeigt sich, dass der Schein trügt. Die Aufnahmen auf ihrem Monitor sind nicht real, sondern stammen aus einem der populärsten Videospiele der Welt. Ihr Team hat sich Methoden der Videospielszene bedient, um das Spiel für die Wissenschaft nutzbar zu machen. Dabei hat es eine überraschende Lösung für eines der kniffligsten Probleme der Computerwissenschaften gefunden.

Datenschutz und Technologie im Konflikt

Laura Leal-Taixés Fachgebiet ist die Bilderkennung, speziell das Erkennen von Menschen in Kamerabildern. Dabei handelt es sich um eine sensible Technologie, wie sie betont: „Solche Methoden können für üble Zwecke verwendet werden. Doch Branchen wie die Automobilindustrie brauchen diese Fähigkeiten dringend." Es ist ein Widerspruch, mit dem sich die Forscherin nicht zufriedengeben will. „Es muss möglich sein, eine Balance zwischen beiden Aspekten herzustellen. Dieses Problem wollen wir lösen", so Leal-Taixé. ▷

Bildquelle: Magdalena Jooss

△ **Bilder von Menschen in Videospielen** eignen sich für Datensätze zum Training neuronaler Netze. Das Forschungsteam hat einen gewaltigen Satz an Trainingsdaten auf der Basis solcher Bilder erzeugt.



▷ **Videospielbilder** bringen die benötigten Informationen für das Tracken von Fußgängern und das Ableiten ihrer Pose bereits mit.

Wie sich das Generieren der enormen Datenmengen mit den Persönlichkeitsrechten der Menschen vereinbaren lässt, denen die Daten gehören, ist ein Schlüsselproblem auf dem Gebiet der Künstlichen Intelligenz, und es fehlt nicht an warnenden Stimmen, die ein technologisches Zurückfallen Europas prophezeien, wenn die Regeln nicht für Forschungsarbeiten gelockert werden. Forschungsinteresse sei über Persönlichkeitsrechte zu stellen, so der Tenor.

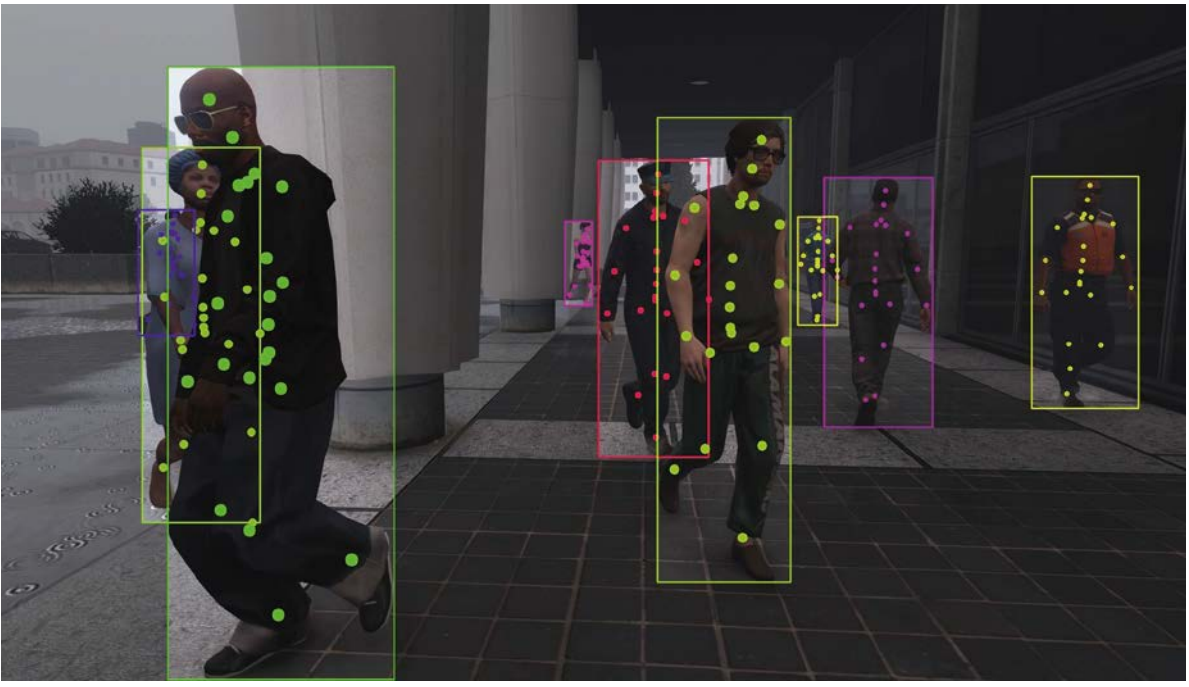
Genau an diesem Punkt setzt die Gruppe von Leal-Taixé an, indem sie verschiedene Wege aufzeigt, wie Trainingsdaten für Algorithmen erzeugt werden können, ohne die Privatsphäre von Menschen zu verletzen.

Täuschend echte Videospiele

Eine Lösung ist das Erstellen von Trainingsbildern, die komplett aus dem Computer stammen. Die Forschung muss diese Bilder nicht erst von Grund auf neu programmieren, sondern kann auf ein Phänomen der Populärkultur zurückgreifen. Die Videospielbranche weist seit vielen Jahren hohe Wachstumsraten auf und hat mit ihren Umsätzen inzwischen die globale Filmindustrie in den Schatten gestellt. Realismus ist dabei ein wichtiges Verkaufsargument und wird dank immer stärker werdender Computerhardware auch zunehmend erreicht. Dass Bilder von Personen aus Videospielen zum Training neuronaler Netze geeignet sind, wurde bereits in der Vergangenheit gezeigt. Nun hat Leal-Taixés Team einen riesigen Satz von Trainingsdaten aus Videospielbildern erstellt. In ihrer Arbeit kam das populäre Spiel „Grand Theft Auto V“ zum Einsatz, wegen seines ausladenden Namens meist nur GTA V genannt. Das Setting des Spiels ist eine Großstadt, die Los Angeles nachempfunden und verblüffend realistisch umgesetzt ist. Spielerinnen und Spieler können sich hier frei bewegen. Neben Realismus bietet GTA V einen weiteren Vorteil: Das Spiel ist ein beliebtes Betätigungsfeld für sogenannte „Modder“, die hobbymäßig Teile des Produkts abändern oder optimieren. Es gibt dafür vorgefertigte Softwaretools, die es erlauben, in das Spiel einzugreifen und etwa Personen an bestimmte Orte zu platzieren.

Hilfreiche Hintergrundinformationen

Der Zugang ist nicht nur unproblematisch, was Persönlichkeitsrechte angeht, sondern hat weitere Vorteile. So ist es bei realen Bildern nötig, die tatsächlich darauf befindliche Information zu kennen, etwa die wirklichen Koordinaten der abgebildeten Personen, um sie mit den ▷



Bildquelle: Fabbri et al. 2021: MOTSynth: How Can Synthetic Data Help Pedestrian Detection and Tracking? arxiv.org/abs/2108.09518 (abgerufen am 13.06.2022).

Links: Segmentation Masks

Rechts: Umschließende Kästen und sogenannte Keypoints des Skeletts beschreiben die Pose der Person

Ergebnissen der trainierenden Algorithmen zu vergleichen. Diese sogenannte „Ground Truth“ kann in der Praxis aufwendig zu berechnen sein. Bei Computerspielbildern kennt man aber die zugrunde liegenden Informationen und kann sie direkt aus dem Spiel abgreifen.

Doch sind die synthetischen Bilddaten eines Unterhaltungsmediums wirklich so realistisch, dass Algorithmen für den Straßenverkehr damit trainiert werden können? Leal-Taixé dämpft zu große Erwartungen: „Wenn man nur mit synthetischen Daten trainiert, ist die Performance bei echten Bilddaten nicht perfekt. Neuronale Netze sind sehr sensibel, was die Textur von Bildern angeht.“ Dieses Manko lässt sich beheben, wenn nach dem Training mit Videospielbildern zusätzlich noch mit echten Bildern trainiert wird, die zuvor anonymisiert wurden, um die Persönlichkeitsrechte der Abgebildeten zu wahren.

Weichzeichnen genügt nicht

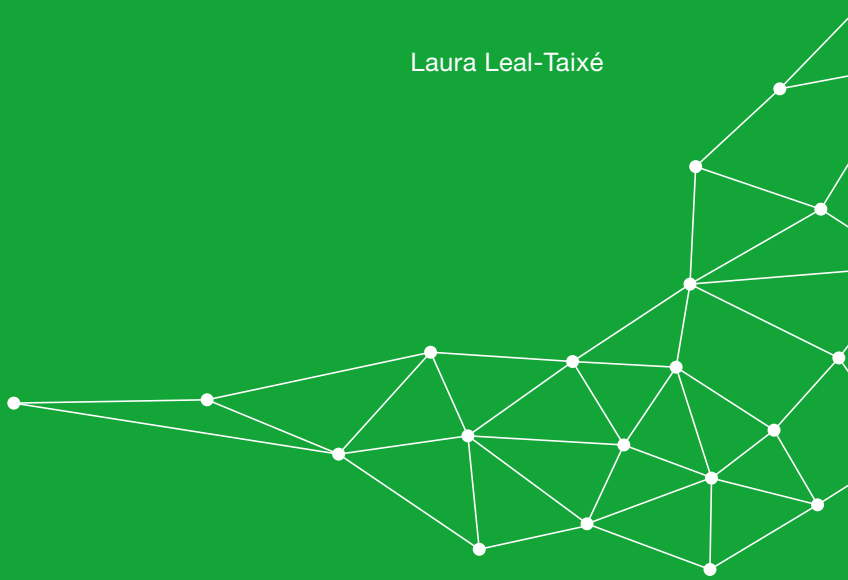
Die Anonymisierung realer Aufnahmen ist ein weiterer Schwerpunkt der Forschungen des Teams von Leal-Taixé. Die gängigste Methode zur Anonymisierung von Personendaten ist das Weichzeichnen oder Verpixeln von Gesichtern. Leal-Taixé weist auf verschiedene Probleme dieses Ansatzes hin. „Wenn man nur das Gesicht anonymisiert, ist die Person womöglich anhand anderer Körpermerkmale identifizierbar. Außerdem lässt sich ein Algorithmus zur Erkennung nicht mit Bildern trainieren, in denen das Gesicht fehlt. Die Bilder müssen so realistisch wie möglich aussehen“, so die Forscherin.

Seit einiger Zeit gibt es daher Versuche, Bilder von Menschen so zu verzerren, dass sie anhand der Gesichter nicht mehr erkennbar sind. Auch das gelingt nicht immer: Es konnte gezeigt werden, dass in vielen Fällen das ursprüngliche Gesicht wieder rekonstruierbar ist. Auch das Erzeugen sogenannter „Deepfakes“, die als Internetphänomen Bekanntheit erlangt haben und bei denen ein Gesicht mit verblüffender Präzision durch ein anderes ersetzt wird, löst das Problem nicht, denn auch die Besitzerin oder der Besitzer des neuen Gesichts hat Persönlichkeitsrechte. Die Münchner Forschungsgruppe geht daher einen Schritt weiter. Man zeigte, dass es möglich ist, die Personen aus den Bildern zu entfernen und durch vollständig computergenerierte Individuen zu ersetzen, die echt aussehen, aber in der Realität nicht existieren. ▷

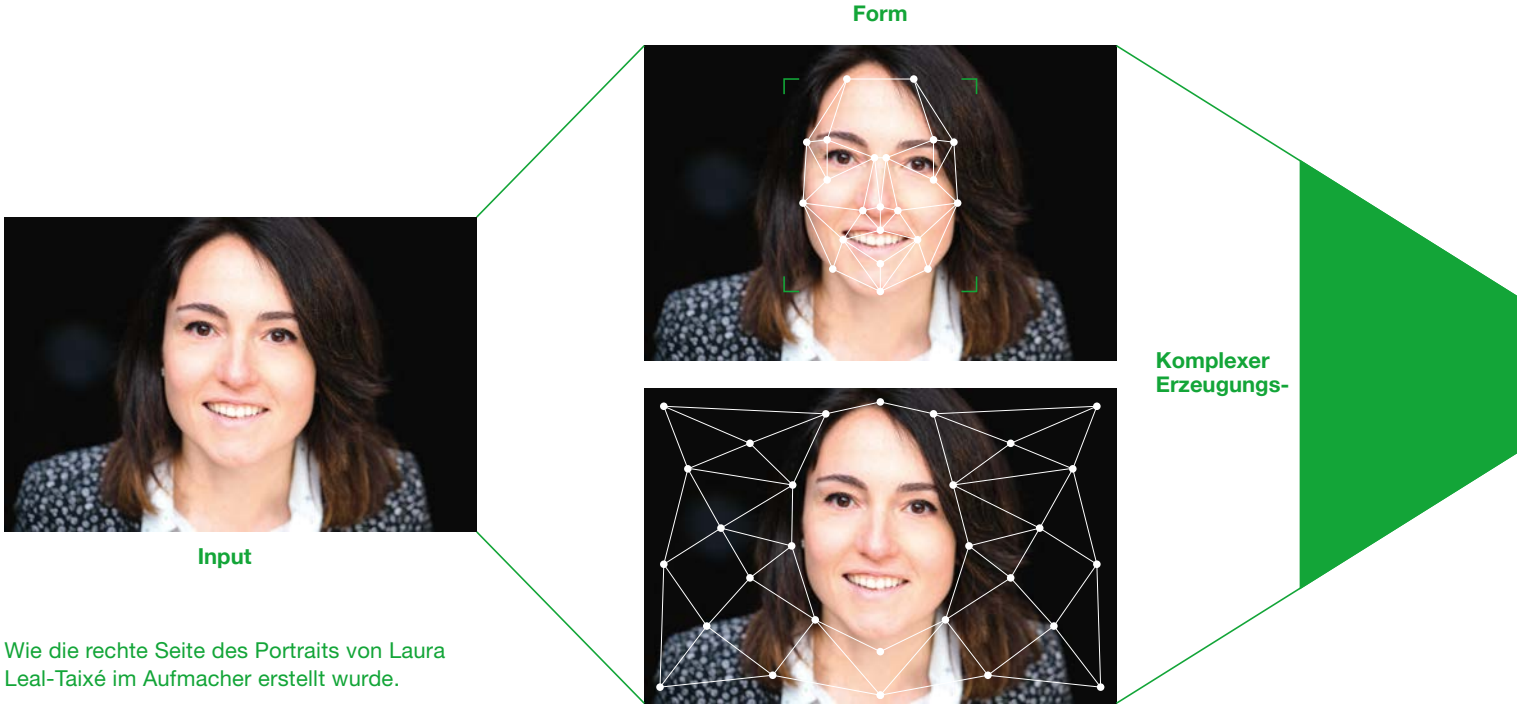
Anonymisierungs-Methoden im Vergleich. Von oben: Original; von Leal-Taixés Team entwickelter CIAGAN-Prozess; Weichzeichnen (17x17 und 9x9); Verpixeln (16x16 und 8x8); Image-to-image Translation (Pix2Pix und CycleGAN).

„Wenn man nur das Gesicht anonymisiert, ist die Person womöglich anhand anderer Körpermerkmale identifizierbar. Außerdem lässt sich ein Algorithmus zur Erkennung nicht mit Bildern trainieren, in denen das Gesicht fehlt. Die Bilder müssen so realistisch wie möglich aussehen.“

Laura Leal-Taixé



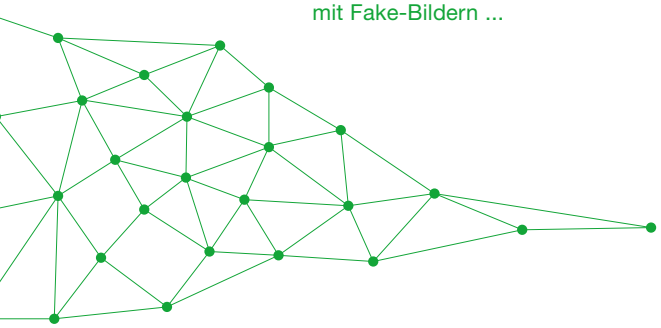
Bildquelle: Maximov et al. 2020: CIAGAN: Conditional Identity Anonymization Generative Adversarial Networks. arxiv.org/abs/2005.09544, Liu et al. 2015: Deep Learning Face Attributes in the Wild. mmlab.ie.cuhk.edu.hk/projects/CelebA.html (abgerufen am 13.06.2022).



Wie die rechte Seite des Portraits von Laura Leal-Taixé im Aufmacher erstellt wurde.

Privatsphäre wahren

mit Fake-Bildern ...



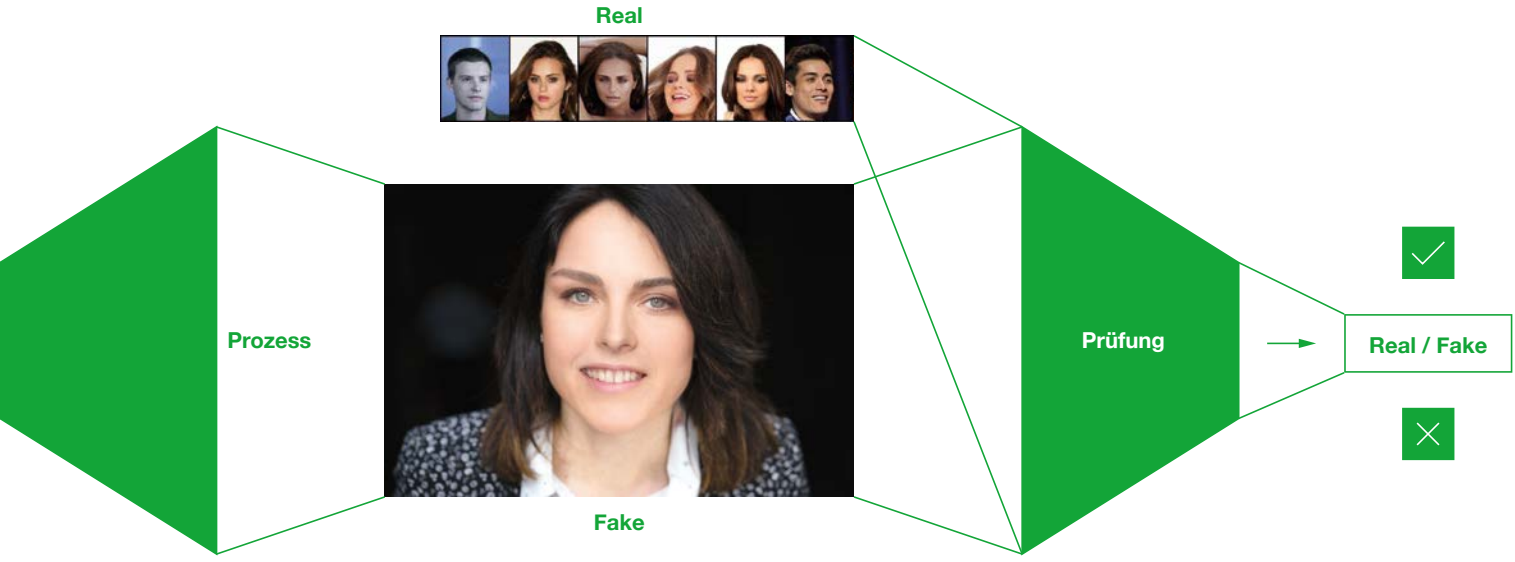
Auch für Videobilder geeignet

Doch genügt das für Anwendungen im Bereich der Smart Mobility? Für bewegte Bilder gelten schließlich besondere Anforderungen. Würde man nur Bild für Bild die Person ersetzen, könnte sich das in abgehackten, unrealistischen Bewegungen äußern. „Wir brauchen Stabilität unseres Bildes über mehrere Einzelbilder des Videos hinweg, um die Verfolgung von Personen trainieren zu können“, so Leal-Taixé. Möglich ist das, indem zuerst sogenannte „Landmarken“ bestimmt werden. „In Gesichtern sind das zum Beispiel die Nasenspitze, die Augen oder der Mund“, erklärt die Forscherin. Auch die Körperhaltung eines Menschen ist eine solche Landmarke. Mithilfe der Bestimmung von Landmarken kann die Pose der Person erkannt und imitiert werden, sodass im Bewegtbild ein realistisches Bewegungsmuster erhalten bleibt.

Während ein neuronales Netz die fiktiven Personen generiert, kontrolliert ein zweites neuronales Netz den Erfolg der Maßnahme, indem es versucht, die Person zu identifizieren. Gelingt die Identifizierung nicht, ist der Nachweis erbracht, dass die Anonymisierung erfolgreich war.

Die Arbeiten befinden sich noch in einem frühen Stadium, doch der Zugang erweist sich als vielversprechend. Leal-Taixés Gruppe gelang es, realistische Bilder von Personen zu erzeugen, die es in der Realität nicht gibt.

Bildquelle: Magdalena Jooss, L. Leal-Taixé, Maximov et al. 2020: CIAGAN: Conditional Identity Anonymization Generative Adversarial Networks. arxiv.org/abs/2005.09544 (abgerufen am 13.06.2022). Grafiken: edundsepp (Quelle: TUM)



Anonymisierung von Gesichtern, sodass sie nicht identifiziert, aber für die Verfolgung detektiert werden können. Das Originalbild und daraus entnommene Landmarken (Augen, Nase, Mund) sowie die Gesichtsform durchlaufen einen komplexen Bilderzeugungs-Prozess, der die Fälschung produziert. Dieses künstlich erzeugte Bild wird abschließend mit realen Bildern abgeglichen, um sicherzugehen, dass keine echten Personen identifiziert werden können.

Junge Arbeitsgruppe

Leal-Taixé setzt bei ihrer Arbeit auf ein junges Team. Kurz nach ihrem Wechsel an die TUM wurde eines ihrer Projekte mit dem Sofja Kovalevskaja-Preis der Alexander von Humboldt-Stiftung ausgezeichnet. Die Dotierung von 1,65 Millionen Euro erlaubte es ihr, eine Arbeitsgruppe aufzubauen. Der Pool an Studierenden wie auch das offene Umfeld seien in München ideal, betont sie: „Die Studierenden sind außergewöhnlich gut vorbereitet und wir würden gerne mehr Diplomarbeiten betreuen als wir können.“ Wenn die Computerwissenschaftlerin nach der Arbeit ihren Rechner ausschaltet und die Universität verlässt, lässt sie die Videospiele übrigens im Büro zurück. Ihr Interesse daran ist rein beruflich, doch das war nicht immer so. Sie spielt auch privat gerne Spiele mit offener Spielewelt wie GTA – ein heimliches Laster, wie sie zugibt. Seit ihre Tochter zur Welt kam, hat sie dafür allerdings kaum noch Zeit. Sie lässt lieber Algorithmen für sich spielen, die dabei lernen, wie sie in Zukunft realen Personen ausweichen.

Reinhard Kleindl

... die Detektion ermöglichen, aber

Identifikation verhindern