

Pressemitteilung

Universität des Saarlandes

Friederike Meyer zu Tittingdorf

02.06.2023

<http://idw-online.de/de/news815416>

Forschungs- / Wissenstransfer, Forschungsergebnisse
Informationstechnik, Kunst / Design, Medien- und Kommunikationswissenschaften
überregional



UNIVERSITÄT
DES
SAARLANDES

KI-Methode „DragGAN“ verspricht die digitale Bildbearbeitung zu revolutionieren

Kleidungsstücke an einem digitalen Avatar anprobieren und von allen Seiten begutachten? Die Blickrichtung des Haustieres auf dem Lieblingsfoto anpassen? Oder die Perspektive auf einem Landschaftsbild verändern? Diese und ähnliche Fotobearbeitungen haben bisher selbst versierte Profis vor Herausforderungen gestellt. Eine neue Methode verspricht nun, sie auch für Laien zugänglich zu machen – dank KI-Unterstützung ganz einfach mit wenigen Mausklicks. Sie wird von einem Forschungsteam unter Leitung des Saarbrücker Max-Planck-Instituts für Informatik entwickelt, insbesondere von dem dortigen Saarbrücken Research Center for Visual Computing, Interaction and Artificial Intelligence (VIA).

Die neue Methode hat das Zeug dazu, die digitale Bildbearbeitung für immer zu verändern. „Mit ‚DragGAN‘ entwickeln wir derzeit ein Werkzeug, das es dank einer übersichtlichen Nutzeroberfläche auch Laien ermöglicht, komplexe Bildbearbeitungen vorzunehmen. Sie müssen nur die Stellen im Foto markieren, die sie verändern möchten. Dann geben sie in einem Menü an, welcher Art die Veränderung sein soll – und mit nur wenigen Mausklicks kann jeder Laie dank KI-Unterstützung die Pose, den Gesichtsausdruck, die Blickrichtung oder den Blickwinkel auf einem Foto, beispielsweise von einem Haustier, anpassen“, erklärt Christian Theobalt, geschäftsführender Direktor des Max-Planck-Instituts für Informatik, Direktor des Saarbrücken Research Center for Visual Computing, Interaction, and Artificial Intelligence und Professor an der Universität des Saarlandes am Saarland Informatics Campus.

Möglich macht das die Künstliche Intelligenz, genauer gesagt die „Generative Adversarial Networks“, kurz GANs. „Wie der Name sagt, handelt es sich bei GANs um generative Modelle, also solche, die neue Inhalte wie Bilder synthetisieren können. ‚Adversarial‘ zeigt an, dass es sich um ein KI-Modell handelt, in dem zwei Netzwerke gegeneinander spielen“, erklärt der Erstautor des Papers, Xingang Pan, Postdoktorand am Max-Planck-Institut für Informatik und am Saarbrücker VIA-Center. In einem GAN arbeiten ein Generator, der Bilder erstellt, und ein Discriminator, der entscheiden muss, ob die Bilder echt sind oder vom Generator erstellt wurden, gegeneinander. Das System wird so lange trainiert, bis der Discriminator die Bilder des Generators nicht mehr von echten Bildern unterscheiden kann.

Die Einsatzmöglichkeiten von GANs sind vielfältig. Neben der offensichtlichen Anwendung des Bildgenerators sind GANs zum Beispiel gut darin, Bilder vorherzusagen: Die sogenannte Video-Frame-Prediction prognostiziert das nächste Bild eines Videos, was den Datenaufwand beim Videostreaming reduzieren kann. Die GANs können zudem niedrig aufgelöste Bilder hochskalieren und die Bildqualität verbessern, indem sie die Position der zusätzlichen Pixel der neuen Bilder vorhersagen.

„In unserem Fall erweist sich diese Eigenschaft von GANs als vorteilhaft, wenn in einem Bild zum Beispiel die Blickrichtung eines Hundes geändert werden soll. Das GAN berechnet dann im Grunde das ganze Bild neu und antizipiert, wo welches Pixel im Bild mit der neuen Blickrichtung landen muss. Ein Nebeneffekt davon ist, dass DragGAN auch Dinge berechnen kann, die vorher etwa durch die Kopfposition des Hundes verdeckt waren. Oder wenn der Nutzer die Zähne des Hundes darstellen will, kann er dem Hund auf dem Bild die Schnauze öffnen“, erklärt Xingang Pan. Auch

im professionellen Kontext könnte DragGAN genutzt werden. Beispielsweise könnten Modedesigner den Zuschnitt von Kleidern nachträglich in Fotos anpassen oder Fahrzeughersteller könnten mit wenigen Mausklicks verschiedene Design-Konfigurationen eines geplanten Fahrzeuges durchspielen.

DragGAN funktioniert zwar bei verschiedenen Objektkategorien wie Tieren, Autos, Menschen und Landschaften, die meisten Ergebnisse seien bisher mit GAN-generierten, synthetischen Bildern erzielt worden. „Die Anwendung auf beliebige vom Benutzer eingegebene Bilder ist immer noch ein schwieriges Problem, das wir untersuchen“, ergänzt Xingang Pan.

Das neue Tool der Saarbrücker Informatiker sorgte bereits wenige Tage nach Veröffentlichung des Preprints in der internationalen Tech-Community für Aufsehen und gilt vielen als der nächste große Schritt in der KI-gestützten Bildbearbeitung. Während Tools wie Midjourney dazu genutzt werden können, neue Bilder zu erstellen, vereinfacht DragGAN die Nachbearbeitung von Bildern enorm.

Die neue Methode wird federführend am Max-Planck-Institut für Informatik und dem darin angesiedelten, gemeinsam mit Google eröffneten „Saarbrücken Research Center for Visual Computing, Interaction and Artificial Intelligence (VIA)“ entwickelt, in Zusammenarbeit mit dem Massachusetts Institute of Technology (MIT) und der University of Pennsylvania.

Neben Professor Christian Theobalt und Xingang Pan waren an dem Paper unter dem Titel „Drag Your GAN: Interactive Pointbased Manipulation on the Generative Image Manifold“ beteiligt: Thomas Leimkühler (MPI INF), Lingjie Liu (MPI INF und University of Pennsylvania), Abhimitra Meka (Google) und Ayush Tewari (MIT CSAIL). Das Paper wurde von der ACM SIGGRAPH-Konferenz akzeptiert, der weltweit größten Fachkonferenz für Computergrafik und interaktive Technologien, die vom 6. bis 10. August 2023 in Los Angeles stattfinden wird.

Hintergrund: Max-Planck-Institut für Informatik

Das Max-Planck-Institut für Informatik in Saarbrücken zählt zu den weltweit führenden Forschungsinstituten in der Informatik. Seit der Gründung des Instituts im Jahr 1990 konzentriert es sich auf die Erforschung der mathematischen Grundlagen der Informationstechnologie in den Bereichen Algorithmen und Komplexität sowie Logik der Programmierung. Wissenschaftler am Institut entwickeln auch Algorithmen für verschiedene Anwendungsbereiche wie Programmverifikation, Datenbanken und Informationssysteme und Bioinformatik. Die Forschung im Bereich Visual Computing, also Computergrafik und Computer Vision, an der Grenze zu künstlicher Intelligenz und maschinellem Lernen, ist ebenfalls ein wichtiger Forschungsschwerpunkt am Institut. Mit Publikationen auf höchstem Niveau und der Ausbildung exzellenter Nachwuchswissenschaftler trägt das MPI für Informatik wesentlich dazu bei, die Grundlagenforschung in der Informatik voranzutreiben.

Hintergrund: Saarbrücken Research Center on Visual Computing, Interaction and Artificial Intelligence (VIA)

Das „Saarbrücken Research Center for Visual Computing, Interaction and Artificial Intelligence (VIA)“ ist eine strategische Forschungspartnerschaft zwischen dem MPI für Informatik und Google und betreibt Grundlagenforschung in zukunftsweisenden Bereichen der Computergrafik, Computer Vision und Mensch-Maschine-Interaktion an der Schnittstelle von künstlicher Intelligenz und maschinellem Lernen. Das Zentrum arbeitet eng mit der Universität des Saarlandes und den zahlreichen, international renommierten Informatik-Forschungseinrichtungen am Saarland Informatics Campus zusammen.

Hintergrund: Saarland Informatics Campus

900 Wissenschaftlerinnen und Wissenschaftler (darunter 400 Promovierende) und rund 2100 Studierende aus mehr als 80 Nationen machen den Saarland Informatics Campus (SIC) zu einem der führenden Standorte für Informatik in Deutschland und Europa. Vier weltweit angesehene Forschungsinstitute, nämlich das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI), das Max-Planck-Institut für Informatik, das Max-Planck-Institut für Softwaresysteme,

das Zentrum für Bioinformatik sowie die Universität des Saarlandes mit drei vernetzten Fachbereichen (Informatik, Mathematik, Sprachwissenschaft & Sprachtechnologie) und 24 Studiengänge decken das gesamte Themenspektrum der Informatik ab.

Redaktion:

Philipp Zapf-Schramm
Saarland Informatics Campus
Telefon: +49 681 302-70741
E-Mail: pzapf@cs.uni-saarland.de

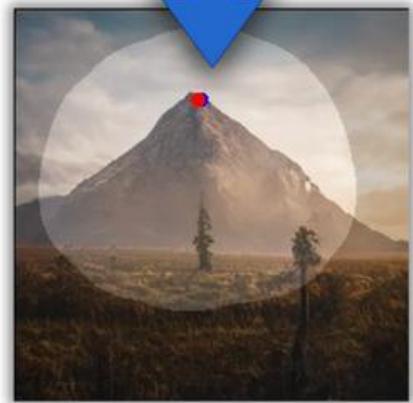
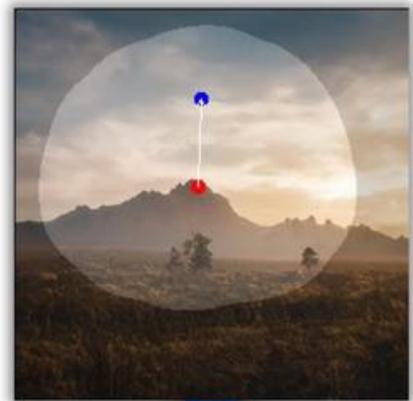
wissenschaftliche Ansprechpartner:

Prof. Dr. Christian Theobalt
Max-Planck-Institut für Informatik
Tel.: +49 681 9325 4500
E-Mail: theobalt@mpi-inf.mpg.de

Originalpublikation:

Preprint: Xingang Pan, Ayush Tewari, Thomas Leimkühler, Lingjie Liu, Abhimitra Meka, and Christian Theobalt. 2023. Drag Your GAN: Interactive Pointbased Manipulation on the Generative Image Manifold. In Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Proceedings (SIGGRAPH '23 Conference Proceedings), August 6–10, 2023, Los Angeles, CA, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3588432.3591500>
<https://arxiv.org/pdf/2305.10973.pdf>

URL zur Pressemitteilung: <https://vcai.mpi-inf.mpg.de/projects/DragGAN/> - Projektwebsite mit Videodemo



Verschiedene Bildbearbeitungen, die mithilfe der Methode DragGAN durchgeführt wurden.
MPI-INF
MPI-INF



Prof. Dr. Christian Theobalt, geschäftsführender Direktor des Max-Planck-Instituts für Informatik und Professor der Universität des Saarlandes
Oliver Dietze
MPI-INF