

## Press release

### Helmholtz-Zentrum Dresden-Rossendorf

#### Simon Schmitt

03/28/2024

<http://idw-online.de/en/news831093>

Research results, Scientific conferences  
Biology, Information technology, Mathematics, Medicine  
transregional, national



## Künstliche Intelligenz macht hochauflösende Mikroskopie noch besser

Generative Künstliche Intelligenz (KI) ist von Anwendungen zur Texterstellung oder Bilderzeugung wie ChatGPT oder Stable Diffusion bekannt. Darüber hinaus zeigt sich ihr Nutzen in verschiedenen wissenschaftlichen Bereichen. In ihrer jüngsten Arbeit, die auf der International Conference on Learning Representations (ICLR) präsentiert wird, haben Forscherinnen und Forscher des Center for Advanced Systems Understanding (CASUS) am Helmholtz-Zentrum Dresden-Rossendorf in Zusammenarbeit mit Fachleuten des Imperial College London und des University College London einen neuen Open-Source-Algorithmus namens Conditional Variational Diffusion Model (CVDM) vorgestellt (doi: 10.48550/arXiv.2312.02246).

Dieses Modell basiert auf generativer KI und verbessert die Qualität von Bildern, indem es sie aus Zufallsdaten rekonstruiert. Darüber hinaus ist das CVDM rechnerisch weniger aufwendig als etablierte Diffusionsmodelle und lässt sich leicht an eine Vielzahl von Anwendungen anpassen.

Das Aufkommen von Big Data sowie neuer mathematischer und datenwissenschaftlicher Methoden nutzen Forscherinnen und Forscher, um bisher ungeklärte Phänomene aus Biologie, Medizin oder den Umweltwissenschaften zu entschlüsseln. Meist widmen sie sich dabei inversen Problemen. Bei solchen Problemen geht es darum, die ursächlichen Faktoren zu ermitteln, die zu bestimmten Beobachtungen führen. Ein Beispiel: Man hat eine Graustufenversion eines Bildes und möchte die Farben wiederherstellen. Hier gibt es in der Regel mehrere valide Lösungsvorschläge, da unter anderem hellblau und hellrot im Graustufenbild identisch aussehen. Die Lösung dieses inversen Problems kann somit das Bild mit dem hellblauen oder das mit dem hellroten Hemd sein.

Die Analyse von mikroskopischen Aufnahmen kann ebenfalls ein typisches inverses Problem darstellen. „Die Beobachtung ist die mikroskopische Aufnahme. Mit Hilfe einiger Berechnungen lässt sich sehr viel mehr über die untersuchte Probe herausfinden, als man beim Blick auf diese Aufnahme zunächst vermutet“, sagt Gabriel della Maggiora, Doktorand am CASUS und Hauptautor der ICLR-Studie. Das Ergebnis können Bilder in höherer Auflösung oder besserer Qualität sein. Jedoch ist der Weg von den Beobachtungen, also den mikroskopischen Aufnahmen, zu den „Superbildern“ in der Regel nicht offensichtlich. Hier spielt auch eine Rolle, dass die Beobachtungsdaten oft unvollständig, ungenau oder durch Rauschen beeinträchtigt sind. All dies trägt zur Komplexität der Lösung inverser Probleme bei und macht sie zu spannenden mathematischen Herausforderungen.

Leistungsfähigkeit generativer KI-Modelle wie Sora

Eines der leistungsfähigsten Werkzeuge zur Lösung inverser Probleme ist die generative KI. Ganz allgemein lernen generative KI-Modelle die zugrunde liegende Datenverteilung eines bestimmten Trainingsdatensatzes. Eine typische Anwendung ist beispielsweise die Bilderzeugung. Nach der Trainingsphase erzeugen generative KI-Modelle dann komplett neue Bilder, die mit den Trainingsdaten vereinbar sind.

Unter den verschiedenen generativen KI-Varianten hat in jüngster Zeit die Gruppe der Diffusionsmodelle an Popularität gewonnen. Bei Diffusionsmodellen beginnt ein stufenweiser Datenerstellungsprozess mit einem Grundrauschen. Das ist ein Konzept aus der Informationstheorie, das die Auswirkung der zahlreichen in der Natur vorkommenden Zufallsprozesse imitiert. Diffusionsmodelle zur Bilderzeugung haben gelernt, welche Pixelanordnungen in den Bildern des Trainingsdatensatzes üblich oder unüblich sind. Sie erzeugen nun Bit für Bit das gewünschte neue Bild, bis sich eine Pixelanordnung findet, die bestmöglich mit der zugrunde liegenden Struktur der Trainingsdaten übereinstimmt. Ein gutes Beispiel für die Leistungsfähigkeit von Diffusionsmodellen ist das Text-zu-Video-Modell Sora des US-Softwareunternehmens OpenAI. Durch eine eingebaute Diffusionskomponente ist Sora in der Lage, Videos zu erzeugen, die realistischer erscheinen als alles, was KI-Modelle zuvor erstellt haben.

Allerdings gibt es auch einen Nachteil. „Diffusionsmodelle sind seit Langem für ihren hohen Rechenaufwand im Trainingsprozess bekannt. Einige Wissenschaftler haben sie genau aus diesem Grund wieder aufgegeben“, sagt Dr. Artur Yakimovich, Leiter einer CASUS-Nachwuchsgruppe und korrespondierender Autor der ICLR-Publikation. „Aber neue Entwicklungen wie unser Conditional Variational Diffusion Model ermöglichen es, die Anzahl der ‚unproduktiven Durchläufe‘, die nicht zum endgültigen Modell führen, zu minimieren. Durch die erreichte Senkung des Rechenaufwands und somit des Stromverbrauchs hilft dieser Ansatz, das Training von Diffusionsmodellen klimafreundlicher zu gestalten.“

Cleveres Training lohnt sich – nicht nur im Sport

Ein wesentlicher Nachteil der Diffusionsmodelle sind die ‚unproduktiven Durchläufe‘. Einer der Gründe für deren Auftreten ist die empfindliche Reaktion des Modells auf die Wahl des vordefinierten Ablaufplans, der die Dynamik des Diffusionsprozesses steuert. Dieser Ablaufplan bestimmt, wie das Rauschen hinzugefügt wird: zu wenig oder zu viel, am falschen Ort oder zur falschen Zeit – es gibt viele mögliche Szenarien, die in einem misslungenen Training enden. Bislang wurde dieser Ablaufplan als Hyperparameter festgelegt, der für jede einzelne Anwendung neu eingestellt werden muss. Mit anderen Worten: Beim Entwurf des Modells ermitteln die Forscherinnen und Forscher den entsprechenden Ablaufplan in der Regel nach dem Prinzip „Versuch und Irrtum“. In dem auf der ICLR vorgestellten neuen Fachartikel beziehen die Autorinnen und Autoren den Ablaufplan bereits in der Trainingsphase mit ein, so dass ihr CVDM in der Lage ist, das optimale Training selbständig zu ermitteln. Das Modell lieferte dann bessere Ergebnisse als andere Modelle, die sich auf einen vordefinierten Ablaufplan stützen.

Unter anderem haben die Autorinnen und Autoren die Anwendbarkeit des CVDM auf ein typisches inverses Problem der Wissenschaft demonstriert: die hochauflösende Mikroskopie. Die hochauflösende Mikroskopie zielt darauf ab, die Beugungsgrenze zu überwinden – eine Grenze, die aufgrund der optischen Eigenschaften des mikroskopischen Systems die maximal mögliche Bildauflösung beschränkt. Um sie algorithmisch zu überqueren, rekonstruieren Datenwissenschaftlerinnen und -wissenschaftler Bilder mit höherer Auflösung, indem sie sowohl Unschärfe als auch Rauschen aus den aufgenommenen, (niedrig aufgelösten) Bildern eliminieren. In diesem Szenario erzielte der CVDM vergleichbare oder sogar überlegene Ergebnisse im Vergleich zu üblicherweise verwendeten Methoden.

„Natürlich gibt es verschiedene Methoden, um die Aussagekraft mikroskopischer Bilder zu erhöhen – einige von ihnen basieren ebenfalls auf generativen KI-Modellen“, sagt Yakimovich. „Aber wir sind davon überzeugt, dass unser Ansatz einige neue, einzigartige Eigenschaften aufweist, die in der Forschungsgemeinschaft Eindruck hinterlassen werden. Das wäre zum einen die hohe Flexibilität und Geschwindigkeit bei vergleichbarer oder sogar besserer Qualität im Vergleich zu anderen Diffusionsmodellansätzen. Darüber hinaus liefert unser CVDM Hinweise, wo genau es sich bei der Rekonstruktion nicht sicher ist. Diese äußerst hilfreiche Eigenschaft ebnet den Weg, um diese Unsicherheiten in neuen Experimenten und Simulationen anzugehen.“

Gabriel della Maggiora wird die Arbeit als Poster auf der jährlichen International Conference on Learning Representations (ICLR) am 8. Mai in Poster-Session 3 um 10:45 Uhr präsentieren. Ein kurzer, vorab aufgezeichneter Vortrag dazu steht auf der Website zur Verfügung. Die Konferenz wird in diesem Jahr zum ersten Mal seit 2017 wieder in

Europa organisiert, und zwar in Wien (Österreich). Sowohl für die Teilnahme vor Ort als auch für den Zugang per Videokonferenz ist ein kostenpflichtiger Zugang erforderlich. „Die ICLR nutzt ein doppelt anonymisiertes Peer-Review-Verfahren über das OpenReview-Portal“, erklärt Yakimovich. „Die Bewertungen von Fachleuten aus den jeweiligen Spezialgebieten beinhalten Noten. Nur Artikel mit hoher Punktzahl werden angenommen. Die Veröffentlichung unserer Arbeit ist daher gleichbedeutend mit einer hohen Wertschätzung durch die Community.“

contact for scientific information:

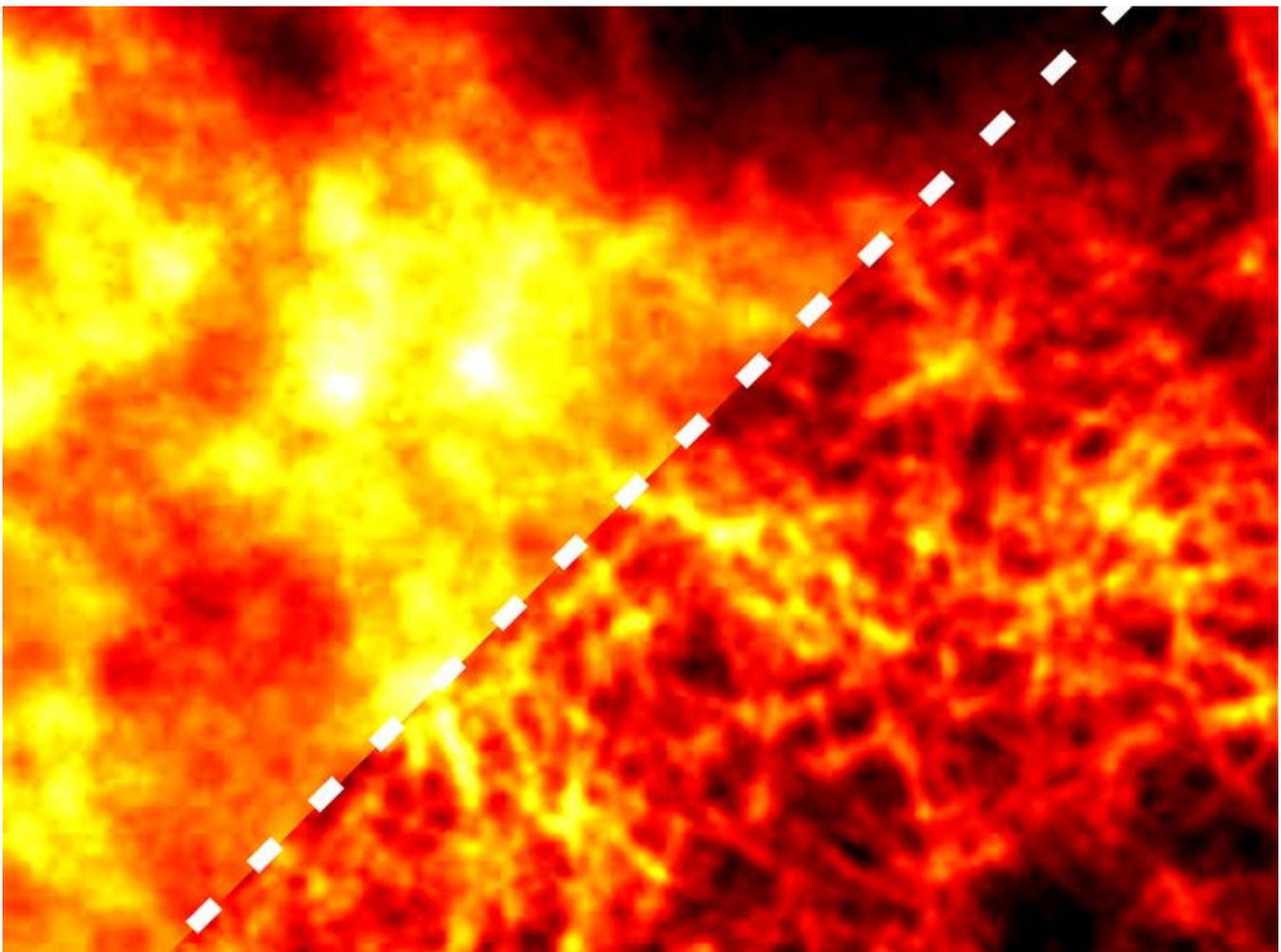
Dr. Artur Yakimovich | Leiter Young Investigator Group  
Center for Advanced Systems Understanding (CASUS) am HZDR  
E-Mail: a.yakimovich@hzdr.de

Original publication:

G. della Maggiora, L. A. Croquevielle, N. Deshpande, H. Horsley, T. Heinis, A. Yakimovich, Conditional Variational Diffusion Models, Proceedings of the Twelfth International Conference on Learning Representations, 2024 (doi: 10.48550/arXiv.2312.02246)

URL for press release:

<https://www.casus.science/news/casus-news/artificial-intelligence-boosts-super-resolution-microscopy/>



Ein Fluoreszenzmikroskopiebild aus dem BioSR-Datensatz für die hochauflösende Mikroskopie (oben & links) wird mit demselben Bild verglichen, das mit dem CVDM rekonstruiert wurde (unten & rechts). Das Bild zeigt markierte F-Actin-Zytoskelettproteine.

© A. Yakimovich/CASUS, bearbeitetes Bild des BioSR-Datensatzes von Chang Qiao & Di Li (lizensiert unter CC BY 4.0)

