

Generative Künstliche Intelligenz für blau-grüne Infrastrukturen

EIN NEUER «CO-PILOT» FÜR DIE SCHWAMMSTADT?

Angesichts rascher Fortschritte in der generativen künstlichen Intelligenz (KI) ging das Institut für Umwelt- und Verfahrenstechnik (UMTEC) der Frage nach, welches Potenzial und welche Tücken dabei bestehen, es als zukünftigen Co-Piloten einzusetzen, um die Zusammenarbeit von Entscheidungsträgern und die Sensibilisierung der Öffentlichkeit zu unterstützen.

Peter Marcus Bach & Veljko Prodanovic

Das Informationszeitalter des mittleren 20. Jahrhunderts revolutionierte die menschliche Gesellschaft und brachte eine neue Art des Lebens, der Arbeit und des Zugangs zum Wissen der Welt hervor. Wir stehen mitten an der Schwelle einer weiteren Veränderung und erleben den raschen Fortschritt und die Verbreitung der künstlichen Intelligenz (KI) in unserem Alltag – ein Zeitalter, das einige als das Intelligenz- und Wissenszeitalter bezeichnen. Bisher wurde KI in den Bereichen urbanen Wasser- und Umweltmanagements hauptsächlich für prädiktive Zwecke eingesetzt, indem maschinelles Lernen verwendet wurde, um die Planung und Verwaltung, den Betrieb und die Steuerung von Infrastrukturen zu unterstützen oder die Modellierung in datenarmen Umgebungen zu verbessern.

Mit der Einführung von grossen Sprachmodellen (LLMs), insbesondere mit der Veröffentlichung des Generativen Vortrainierten Transformers durch OpenAI und seiner zugänglichen Form, ChatGPT, auf dem Massenmarkt im November 2022, wurde Generative KI zu einem Begriff in aller Munde. Diese neue Form der KI ist nun in der Lage, neue Inhalte zu erstellen, von Texten über Bilder, Ton und andere visuelle Medien. Bildgenerierungs-

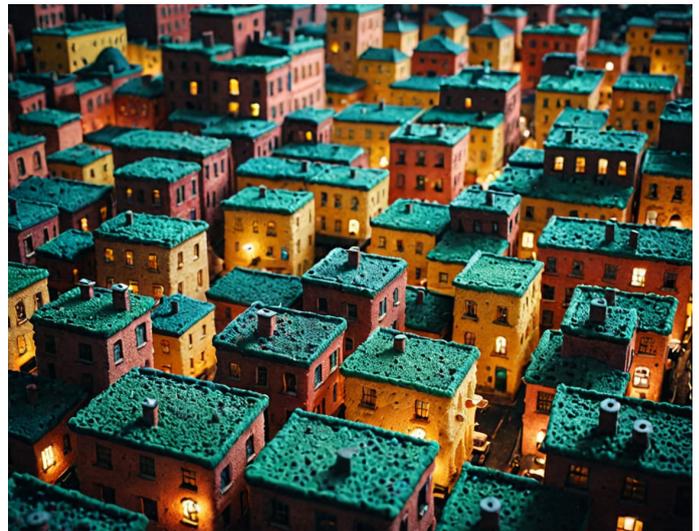
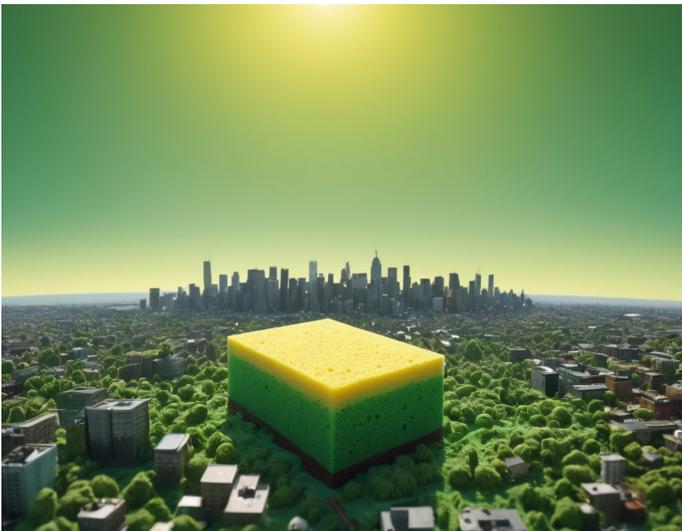
modelle werden auf grossen Datensätzen von Bildern trainiert und können, je nach Spezifikationen, fein abgestimmt werden, um spezifischere Arten von Bilddaten zu generieren. Im Gegensatz zu LLMs, die auf textbasierte Daten trainiert sind, um Wortsequenzen zu reproduzieren, versuchen Bildgenerierungsmodelle, einzelne Pixel innerhalb eines Bildbereichs zu definieren. Zwei vorherrschende Ansätze, die erforscht wurden, sind gegnerische Netzwerke und latente Diffusionsmodelle. Insbesondere Letztere wurden als überlegene Option demonstriert und umfassen einen Prozess der Umwandlung von Wahrscheinlichkeitsverteilungen in bedeutungsvolle Bilder, die durch Textvorgaben gesteuert werden, wobei die bekanntesten Beispiele DALL-E und Midjourney sind.

Die raschen Fortschritte und die Akzeptanz von generativer KI werfen jedoch auch Bedenken hinsichtlich potenzieller und nützlicher Anwendungen auf, insbesondere im Bereich der Umwelt- und Stadtplanung. Forscher experimentieren damit, wie solche Ansätze die bestehenden Arbeiten in der Stadtplanung und -gestaltung unterstützen können. Ohne Feinabstimmung sind sogenannte stabile Diffusionsmodelle (eine Art latenter Diffusionsmodelle) bereits in der Lage,

sinnvolle Konzepte in der Stadtplanung zu visualisieren. Abgesehen von reinen Visualisierungen wurden auch erste Versuche unternommen, solche Tools zur Unterstützung der Planung nachhaltiger gebauter Umgebungen zu demonstrieren. Ebenso ist es im Bereich der urbanen Entwässerung, insbesondere angesichts der zunehmenden Aufmerksamkeit für Blau-Grüne Infrastrukturen (BGI), naturbasierter oder naturnaher Lösungen und klimaangepasster Schwammstädte, die eine Reihe von Ökosystemdienstleistungen bieten, unvermeidlich, dass die Nutzung und Nützlichkeit von generativer KI letztendlich hinterfragt werden.

Unterschiedliche Anwendungsbeispiele

Eine breite Potenzial-Erkundung der generativen KI (speziell der KI-Bildgenerierung) zeigt drei Anwendungs-Bereiche zur Unterstützung der Schaffung klimaangepasster und nachhaltiger, städtischer Umgebungen: Visuelle Analogien für Wissenstransfers und Denkanstösse, Visualisierungen von unerwünschten zukünftigen Szenarien und Visualisierungen zur Ideenfindung für Designs. In diesen drei Bereichen ist der grösste und unmittelbarste Einfluss der generativen KI erzielbar und am sinnvollsten – aber dennoch mit Vorsicht anzuwenden.



Visualisierung des Konzeptes Schwammstadt mit generativer KI (Bilder: zvg)

Schwammstädte und deren BGI tragen dazu bei, den natürlichen Wasserkreislauf wiederherzustellen, indem sie Regenwasser durch lokale Versickerung und Verdunstung zurückhalten. Eine häufige Art, dieses Konzept der breiten Öffentlichkeit zu kommunizieren, ist die Idee des «Schwammes», die den Regen aufsaugt. Durch gezielte Texteingaben können Experten solche Analogien, ohne weitere Informationen, visuell darstellen und kommunizieren. Zusätzlich können Experten durch gezielte und beschreibende Wortwahl in der Bildgenerierung lernen, wie die Darstellung und Kommunikation des Konzeptes an die Öffentlichkeit konkretisiert werden kann.

Urbane Überschwemmungen und Überhitzung treten in vielen Städten weltweit zunehmend auf. Die Visualisierung der Auswirkungen solcher Katastrophen im Voraus kann helfen, die Dringlichkeit

von Schutzmassnahmen und effektiven Wasserbewirtschaftungsstrategien zu kommunizieren (insbesondere im Fall von Hitze, wo die mangelnde Verfügbarkeit von Wasser das Kühlpotenzial der Vegetation verringert und ihre Gesundheit verschlechtert). Als Basismotiv kann ein Foto eines innerstädtischen Gebietes verwendet werden, um diese beiden Szenarien durch «In-painting» (bei dem Teile des Bildes modifiziert werden) zu demonstrieren.

Der dritte Anwendungsfall testet die Fähigkeiten der generativen KI zur Unterstützung des Designs lokaler BGI-Lösungen. Unter Verwendung des Basismotivs eines versiegelten Platzes können solche Massnahmen konkret generiert und mit wenig Aufwand dargestellt werden, wie zum Beispiel: Grasflächen, Sträucher, ein Wasserelement und lokale Flora zur Förderung der Biodiversität. Dies wür-

de Planern und Designern ermöglichen, Konzepte und Ideen rund um BGI auf lokaler Ebene schnell zu visualisieren und Meinungen und Präferenzen von örtlichen Grundstücksbesitzern und der Gemeinde einzuholen, bevor sie sich in detaillierte architektonische Planungen vertiefen.

Methodik und Verfeinerungsbedarf

Die für diese Studie ausgewählte Bildgenerierungs-Engine ist das Stable-Diffusion XL-Framework (SDXL), eine Verbesserung des bereits bekannten Stable-Diffusion-Ansatzes (wie bei DALL-E und Midjourney) für eine zusätzliche, hochauflösende Bildgenerierung mit besonderer Aufmerksamkeit für Verfeinerungen. SDXL ist Open Source, weitgehend konfigurierbar und kann auf einem lokalen Rechner mit ordentlicher GPU-Ausstattung betrieben werden, was im Vergleich zu öffentlichen und Server-ba-

Visualisierung von Szenarien am Beispiel Luzern: Basismotiv ist der Ausblick von der St. Leodegarstrasse aus (oberste), darauf folgen visualisierte Auswirkungen bei heftigem Starkregenereignis, Hitze und Wasserknappheit. (Bilder zvg)



sierten Tools ein Vorteil für Anwendungen an sensiblen Daten und Projekten bedeutet. Die Generierung ist unterstützt von sogenannten «Checkpoint» Dateien, die Modellparameter beinhalten und selber an fallspezifischen Bildern erstellt werden können, um realistische und ortsrelevante Ergebnisse zu erzeugen.

Zusätzlich zu den Checkpoints können weitere spezialisierte Trainingsdateien erstellt werden, die sich nur mit gewissen visuellen Elementen befassen (zum Beispiel die Darstellung von europäischen Baumarten). Trotz seiner überzeugenden visuellen Qualität stiess der Prozess für die generierten Beispiele auf mehrere Herausforderungen. Einerseits hatte die generative KI Schwierigkeiten mit der Visualisierung von Bäumen und anderen grösseren, vertikalen Strukturen. Andererseits gab es auch, trotz des scheinbar fotorealistischen Erscheinungsbildes, noch einige Diskrepanzen im Modell wie zum Beispiel die Beleuchtung und Schatten, die für das geschulte Auge deutlich einen Mangel an Realismus hervorrufen können.

Zusätzlich sind sogenannte Konzeptüberschneidungen, bei denen die KI dazu neigt, minimale Informationen von einem Objekt oder einer Person in andere Elemente im Hintergrund zu kopieren oder falsch zu interpretieren, auch noch weit verbreitet. Obwohl dieses Problem durch schrittweise Generierung von Bildern oder durch die Erweiterung der Open Source Implementierung behoben werden kann, deutet dieses Ergebnis darauf hin, dass das Expertenwissen noch eine wichtige Rolle in Zukunft spielen wird, um täuschende Ergebnisse herauszufiltern. Ausserdem hat die KI eine Neigung zur Datenbasis, die für ihr Training verwendet wurde. Daher würden spezifische Trainingssets für die BGI-Stadtplanung ein raffinierteres Ergebnis liefern. Schliesslich erfordert die genaue Darstellung von Szenarien und BGI-Visualisierungen auch ein gutes Basismotiv und eine Komposition, um unrealistische Ergebnisse zu vermeiden (wie zum Beispiel Regenwolken an einem heissen Sommertag oder Menschen, die be-

quem bei einer Überschwemmung unter Wasser sitzen).

Potenzial, Tücken und Empfehlungen

Diese ersten Ergebnisse und die unterschiedlichen Anwendungsmöglichkeiten weisen auf einen Mehrwert der generativer KI in der Umweltechnik und der Städteplanung hin. Unter Verwendung von Basistrainingsdaten für den Bildgenerator können realistische, klare und deutliche Bilder innerhalb von Minuten auf einem lokalen Rechner erzeugt werden. Den Einsatz eines Open Source KI Modells ermöglicht auch dem Benutzer bessere Kontrolle über die Einstellungen und Generierung der Bilder, ist aber mit einem Lern- und Verständnisaufwand verbunden.

Die ersten Erfahrungen bieten wertvolle Erkenntnisse für den breiteren Einsatz der generativen KI in den Bereichen BGI und Schwammstädte. Potenzielle Bedenken hinsichtlich deren Einsatzes sind bekannt. Grosse Tech-Unternehmen betonen zunehmend die Notwendigkeit eines verantwortungsvollen Umgangs mit KI. Angesichts der Vielzahl von Anwendungsmöglichkeiten und vermeintlich endloser Potenziale zur Verbesserung der Kommunikation, des Designs und des Bewusstseins für die Auswirkungen des Klimawandels auf die Lebensqualität in städtischen Gebieten, sind umfassende Experimente von Experten mit diesem Werkzeug, eine erhöhte Überprüfung und transparente Kommunikation des generierten Inhalts gegenüber der Öffentlichkeit unbedingt notwendig, um Missverständnisse oder Fehlinterpretationen zu vermeiden. Generative KI sollte als Werkzeug oder «Co-Pilot» betrachtet werden, der von Experten gesteuert wird, die nach wie vor die Möglichkeit und die Mittel zur Gestaltung und Kontrolle seiner Entwicklung und Nutzung haben sollten.

Zukünftige Arbeiten zielen darauf ab, zu erforschen, wie diese Modelle gemeinsam mit Interessensgruppen angewendet werden können. Es besteht kein Zweifel dar-

an, dass diese und viele neue Anwendungen weiterhin entstehen werden, wenn neue Werkzeuge und Modelle aktualisiert werden. Neben diesen muss die Verantwortung und Rechenschaftspflicht der Benutzer sowie die transparente Kommunikation von KI-generierten Inhalten betont werden, um Missbrauch und Fehlinformationen zu vermeiden. ●

● Weitere Informationen unter:
www.umtec.ch
www.ost.ch

Konfiguration des generative KI-Modells für diese Studie

Für diese Studie wurde das Fooocus v2-Paket (<https://github.com/lllyasviel/Fooocus>) verwendet zusammen mit dem Juggernaut XL v6 und realvisXLV40_v40 Checkpoints. Für das In-Painting wurde ein 2-Schritte-Ansatz verwendet, um das Bild von Luzern im zweiten Anwendungsfall schrittweise zu modifizieren, wobei zuerst der Himmel und anschliessend die urbane Landschaft verändert wurde.

Zur Autorenschaft:

Dr. Peter Marcus Bach ist Wissenschaftlicher Mitarbeiter am Institut für Umwelt- und Verfahrenstechnik (UMTEC) an der Ostschweizer Fachhochschule in Rapperswil (OST) in den Themen: «Schwammstadt», und «naturbasierte Lösungen». Er arbeitet auch mit internationalen Partnern zusammen an diversen Themen im Bereich Klimaanpassung, Städtebau und Umwelt (weitere Aktivitäten unter www.petermbach.com). Dr. Veljko Prodanovic ist Wissenschaftlicher Mitarbeiter am Institut für Künstliche Intelligenz und am Institut für Multidisziplinäre Forschung in Belgrad, Serbien.



Ideenfindung und Design von Blau-Grünen Infrastrukturen am Beispiel des Sechseläutenplatzes in Zürich (oberste) als Basismotiv. (Bilder: zvg)