

SocietyByte

BFH-Magazin für die Humane Digitale Transformation

Diffusion Models: Ein neuer Horizont in der Bilderzeugung

Von Souhir Ben Souissi (BFH Technik & Informatik) | 0 Kommentare



Als im vergangenen Jahr der Vorhang fiel, drehten sich die meisten Gespräche über generative KI in öffentlichen Foren und den Medien um die Texterstellung, wobei Namen wie chatGPT das Rampenlicht stahlen. Die Realität ist jedoch, dass die transformative Kraft der künstlichen Intelligenz, angeführt von Deep Learning, weit über den Bereich der Wörter hinausgeht. Ein herausragendes Beispiel ist die Bilderzeugung, die in der Welt der KI zu einem lebendigen Bereich der Erforschung und Innovation geworden ist.

Diese KI-gesteuerten Tools, die als generative Modelle bekannt sind, haben die Bilderzeugung auf ein neues Niveau gehoben und erzeugen Bilder, die nicht nur realistisch, sondern oft auch fesselnd in ihren Details sind. Sie sind sogar in den Bereich der Kunst und des Designs vorgedrungen und haben die Grenzen dessen, was mit Technologie möglich ist, verschoben.

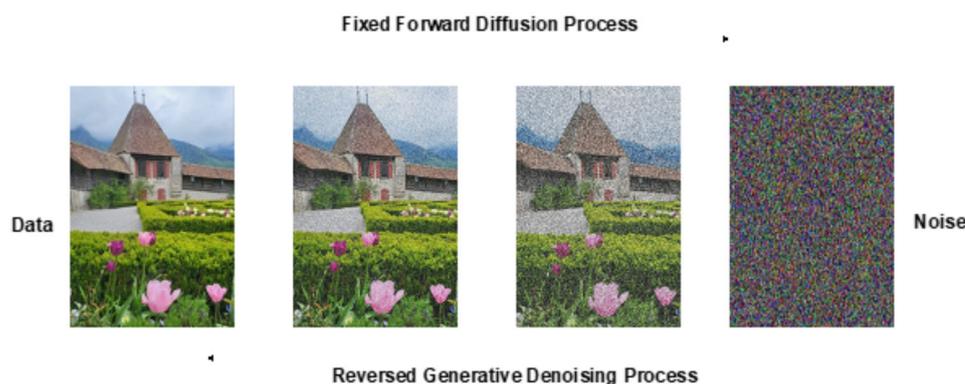
In der Reihe dieser bildgenerativen Modelle sind einige Namen besonders hervorzuheben: Variationale Autoencoder (VAE), flussbasierte Modelle und Generative Adversarial Networks. Aber unter diesen Spielern hat eines aufgrund seiner einzigartigen Fähigkeiten still und leise Wellen geschlagen – das Diffusionsmodell.

Wenn Sie sich für KI-Modelle interessieren, die in der Lage sind, einzigartige und realistische Bilder zu erzeugen, finden Sie hier eine Einführung in die Welt der Diffusionsmodelle und ihre Anwendungen in der Bilderzeugung.

Die Wissenschaft hinter Diffusionsmodellen

Diffusionsmodelle sind eine Art von KI, die neue Dinge, wie z. B. Bilder, erzeugt, indem sie einen Prozess nachahmt, der der zufälligen Bewegung winziger Teilchen in einer Flüssigkeit ähnelt. Das hört sich vielleicht komplex an, also lassen Sie es uns aufschlüsseln.

Stellen Sie sich vor, Sie haben ein klares Bild. Die Aufgabe des Diffusionsmodells besteht zunächst darin, dem Bild nach und nach zufälliges visuelles «Rauschen» hinzuzufügen, bis nur noch zufällige, unordentliche Flecken übrig bleiben, wie das Rauschen auf einem Fernsehbildschirm. Die Umwandlung von einem klaren Bild in dieses verrauschte Rauschen wird durch eine Reihe von Regeln gesteuert, die als stochastische Differentialgleichung (SDE) bekannt sind.



Um ein neues Bild zu erzeugen, wird dieser Prozess umgekehrt. Ausgehend von einer Stichprobe aus der Rauschverteilung wendet das Modell die Umkehrung der SDE an, um das Rauschen wieder in ein aussagekräftiges Bild zu verwandeln.

Woher weiß es, wie es das machen soll? Nun, während des Trainings lernt das Modell, wie es vom Rauschen zu einem aussagekräftigen Bild kommt, indem es ständig versucht, die Differenz zwischen seinen neu erstellten Bildern und den tatsächlichen Bildern aus seinem Trainingsdatensatz zu minimieren.

Kurz gesagt: Diffusionsmodelle erzeugen neue Bilder, indem sie zunächst klare Bilder in Rauschen verwandeln und dieses Rauschen dann geschickt wieder in klare, aussagekräftige Bilder zurückverwandeln.

Warum sind Diffusionsmodelle wichtig?

Diffusionsmodelle erweisen sich als bahnbrechend, da sie beeindruckende, qualitativ hochwertige Bilder erzeugen und gleichzeitig eine feinkörnige Kontrolle über den Erstellungsprozess ermöglichen. Im Gegensatz zu ihren Pendanten, den GANs (Generative Adversarial Networks) oder VAEs (Variational Autoencoders), verfügen Diffusionsmodelle über ein klares, gut strukturiertes Trainingsziel, das dazu beiträgt, häufige Probleme wie den Zusammenbruch des Modus zu vermeiden, mit denen andere generative Modelle oft zu kämpfen haben.

Was Diffusionsmodelle von anderen Modellen unterscheidet, ist ihr schrittweises Vorgehen, das die Erzeugung von Bildern mit verschiedenen Auflösungen ermöglicht. Diese Vielseitigkeit ist unglaublich wertvoll für Aufgaben wie die Verbesserung der Bildauflösung oder das Auffüllen fehlender Teile eines Bildes. Die mit diesen Modellen erzeugten Bilder sind nicht nur von außergewöhnlicher visueller Qualität, sondern bieten auch spannende Möglichkeiten in einer Reihe von Bereichen, darunter Videospiele, virtuelle Realität und sogar die Filmindustrie.

Beliebte Diffusionsmodelle

- Dall-E 2 von OpenAI[1] [#_ftn1]



- Googles Imagen[2] [#_ftn2]



- StabilityAI's Stabile Diffusion[3] [#_ftn3]



- Midjourney[4] [#_ftn4]



Zusammenfassend lässt sich sagen, dass sich Diffusionsmodelle mit ihren einzigartigen Fähigkeiten zur Bilderzeugung eine Nische im Bereich der künstlichen Intelligenz geschaffen haben. Wenn die Forschung weitergeht, können wir in naher Zukunft noch mehr beeindruckende und vielfältige Anwendungen dieser Modelle erwarten. Sie sind ein weiterer Schritt auf dem Weg des maschinellen Lernens und der künstlichen Intelligenz, der ein vielversprechendes Bild der Zukunft zeichnet.

Referenzen

[1] [[#_ftnref1](#)] <https://openai.com/dall-e-2> [<https://openai.com/dall-e-2>]

[2] [[#_ftnref2](#)] <https://imagen.research.google/>
[<https://imagen.research.google/>]

[3] [[#_ftnref3](#)] <https://stability.ai/stablediffusion> [<https://stability.ai/stablediffusion>]

[4] [#_ftnref4] <https://www.midjourney.com/>
[<https://www.midjourney.com/>]



AUTHOR: SOUHIR BEN SOUISSI



Dr. Souhir Ben Souissi ist Tenure-Track-Professorin für Data Engineering am Institut für Datenanwendungen und Sicherheit (IDAS) der BFH Technik & Informatik. Ihre Forschungsschwerpunkte liegen unter anderem auf den Themen Medizinische Entscheidungssysteme, Semantische Webtechnologien und Multikriterielle Entscheidungssysteme.

[Posts from Souhir Ben Souissi | Website](#)

[Create PDF](#)

Ähnliche Beiträge

[Generative AI – was ist das und was kann sie bereits?](#)

[Sprachmodelle und Natural Language Processing - ein Rückblick auf die Transform-Konferenz 2023](#)

[Was Künstliche Intelligenz heute schon im Gesundheitswesen leistet](#)

[Wenn ChatGPT bei der Prüfung hilft - eine Podcast-Episode über KI im Studium](#)

0

COMMENTS