



Neuronale Puppenspieler

Künstliche Intelligenz rekonstruiert Bewegungsabläufe von Mensch und Tier

Stellen wir uns vor, wir sind auf Safari: Wir beobachten eine Giraffe, die frisst. Kurz sind wir abgelenkt, dann sehen wir, wie das Tier den Kopf senkt und sich setzt. Doch was ist in der Zwischenzeit passiert? Computerwissenschaftlern des Exzellenzclusters „Centre for the Advanced Study of Collective Behaviour“ der Universität Konstanz ist es gelungen, die statistisch wahrscheinlichen Zwischenschritte anhand von Kodierungen von Pose und Erscheinungsbild des Tiers aufzuzeigen.

Ein Kernproblem in der Computer Vision ist, dass Bilder unglaublich komplex sind. Die Giraffe kann sehr viele verschiedene Posen einnehmen. Bei einem Safariausflug ist es meist weniger problematisch, wenn man einen Bewegungsablauf verpasst hat. Doch in der Erforschung kollektiven Verhaltens kann das essentiell sein. Hier helfen Computerwissenschaftler nun mit einem neuen Modell, Neural Puppeteer (Deutsch: Neuronale Puppenspieler) genannt, weiter.

- **Videos sind verfügbar unter:** <https://www.campus.uni-konstanz.de/wissenschaft/neuronale-puppenspieler>

Vorhersage von Silhouetten auf der Basis von 3D-Punkten

„Eine Idee in der Computer Vision ist es, Kodierungen zu finden, die den sehr komplexen Raum von Bildern mit möglichst wenig Parametern beschreiben können“, erklärt Bastian Goldlücke, Professor für Computer Vision der Universität Konstanz. Eine Darstellung, die man bislang häufig benutzt, sind 2D-Skelette. In einem neuen Paper, das in den Proceedings der 16th Asian Conference on Computer Vision erschien, stellen Bastian Goldlücke und die Doktoranden Urs Waldmann und Simon Giebenhain ein 3D-Modell vor, das eine Repräsentation von Bewegungsabläufen anhand weniger Schlüsselpunkte ermöglicht. Die 3D-Ansicht ist plastischer und präziser als die bisherigen Skelett-Modelle.

„Die Idee war es, texturunabhängig 3D-Punkte vorherzusagen und auch tracken zu können“, erklärt Doktorand Urs Waldmann. „Dafür haben wir eine KI gebaut, die aus 3D-Schlüsselpunkten Bilder der Silhouetten aus beliebiger Kameraperspektive vorhersagt.“ Durch Umkehrung des Prozesses ist es ebenfalls möglich, aus Bildern der Silhouette die Skelettpunkte zu bestimmen. Anhand der gesetzten Schlüsselpunkte ist die KI in der Lage, statistisch wahrscheinliche Zwischenschritte zu ermitteln. Die individuelle Silhouette kann bei der Anwendung von Bedeutung

sein: Denn wenn man lediglich mit Skelettpunkten arbeitet, wisse man nicht, ob man ein ziemlich voluminöses Tier vor sich habe oder eines, das kurz vor dem Verhungern sei.

Anwendungsfälle für dieses entwickelte Modell gibt es insbesondere in der Biologie: „Wir haben im Exzellenzcluster Centre for the Advanced Study of Collective Behaviour gesehen, dass viele verschiedene Tierarten getrackt werden und dazu Posen vorhergesagt werden sollen“, sagt Waldmann.

Langfristiges Ziel: Anwendung auf möglichst viele Tierdaten in freier Natur

Angefangen hat das Team zunächst mit der Vorhersage von Silhouetten von Mensch, Taube, Giraffe und Kuh. Der Mensch sei in der Computerwissenschaft ein oft verwendeter Showcase, erklärt Waldmann. Mit Tauben arbeiteten Kolleginnen und Kollegen aus dem Cluster. Sie seien aufgrund der feinen Krallen eine Herausforderung gewesen. Zur Kuh gab es gute Modelldaten und bei der Giraffe war der extrem lange Hals eine Herausforderung, der sich Waldmann stellen wollte. Mit wenigen Schlüsselpunkten, 19 bis maximal 33 waren es, erstellten sie die Silhouetten.

Nun warten die Computerwissenschaftler auf die Anwendung: Trainingsdaten etwa von Insekten und Vögeln werden in Zukunft im Imaging Hangar, dem größten Labor zur Erforschung von Kollektivverhalten der Universität Konstanz, erhoben werden. Im Imaging Hangar seien Aspekte wie Beleuchtung oder Hintergrund einfacher zu kontrollieren als in der Natur. Doch langfristig soll das Modell für möglichst viele Tierarten in freier Natur trainiert werden, um neue Erkenntnisse zum Verhalten von Tieren zu entschlüsseln.

Faktenübersicht

- Giebenhain, S., Waldmann, U., Johannsen, O., Goldluecke, B. (2023). Neural Puppeteer: Keypoint-Based Neural Rendering of Dynamic Shapes. In: Wang, L., Gall, J., Chin, T.J., Sato, I., Chellappa, R. (eds) Computer Vision. ACCV 2022. Lecture Notes in Computer Science, vol 13844. Springer, Cham. https://doi.org/10.1007/978-3-031-26316-3_15
- In ihrer Publikation mit dem Titel Neural Puppeteer stellen sie ein Modell zur Repräsentation von Bewegungsabläufen von Individuen anhand weniger Schlüsselpunkte vor
- Bastian Goldlücke ist Professor im Bereich Computer Vision an der Universität Konstanz
- Urs Waldmann promoviert am Exzellenzcluster Centre for the Advanced Study of Collective Behaviour

Hinweis an die Redaktionen:

Ein Bild kann im Folgenden heruntergeladen werden: https://www.uni-konstanz.de/fileadmin/pi/fileserver/2023_EXSTRA/neuronale_puppenspieler.png

Bildunterschrift: Anhand von wenigen 3D-Schlüsselpunkten rekonstruiert die KI Bewegungsabläufe von Mensch und Tier.

Bild: Urs Waldmann

Kontakt:

Universität Konstanz
Kommunikation und Marketing
Telefon: + 49 7531 88-3603
E-Mail: kum@uni-konstanz.de

- *uni.kn*

—

—