



Prof. Dr. Xiaoxiang Zhu

forscht am DLR-Institut für Methodik der Fernerkundung und leitet dort die Abteilung Earth Observation Data Science. Sie unterrichtet zudem an der Technischen Universität München und wurde dort 2015 im Alter von 30 Jahren eine der jüngsten Professorinnen in Deutschland. Ihre Forschungsarbeit wurde bereits mehrfach ausgezeichnet: So erhielt sie 2018 den Early Career Award der Nationalen Akademie der Wissenschaften Leopoldina für ihre herausragenden Leistungen in der satellitengestützten Erdbeobachtung zur Erfassung der weltweiten Urbanisierung sowie von Naturgefahren.

WIE AUS DATENBERGEN WISSEN WIRD

Prof. Dr. Xiaoxiang Zhu versteht sich als Brückenbauerin. Sie erschließt das, was Satelliten zur Erde senden, für die Wissenschaftler, kommerziellen Nutzer und Entscheidungsträger, die aus diesem riesigen Datenberg konkrete Informationen haben wollen. Die 34-jährige Wissenschaftlerin vom DLR-Institut für Methodik der Fernerkundung muss dafür viele Bereiche miteinander verschmelzen: die Fernerkundung der Erde, die Verarbeitung von Daten aus den unterschiedlichsten Quellen und die Informatik. Um die Welt im Wandel zu verstehen und Geoinformationen zu gewinnen, nutzt sie künstliche Intelligenz (KI). In folgendem Beitrag erzählt sie, was damit möglich ist.

Gespräch zur DLR-Forschung an künstlicher Intelligenz

Wo steckt denn in Ihrer Forschung das Intelligente? Was kann KI, was ein großes Team aus Wissenschaftlern nicht kann?

■ Meinen Sie damit, Satellitendaten manuell auszuwerten? (lacht) Das ist nicht machbar ...

Weil die Datenmengen zu groß sind oder weil es zu schwierig ist, die Daten sinnvoll zu kombinieren?

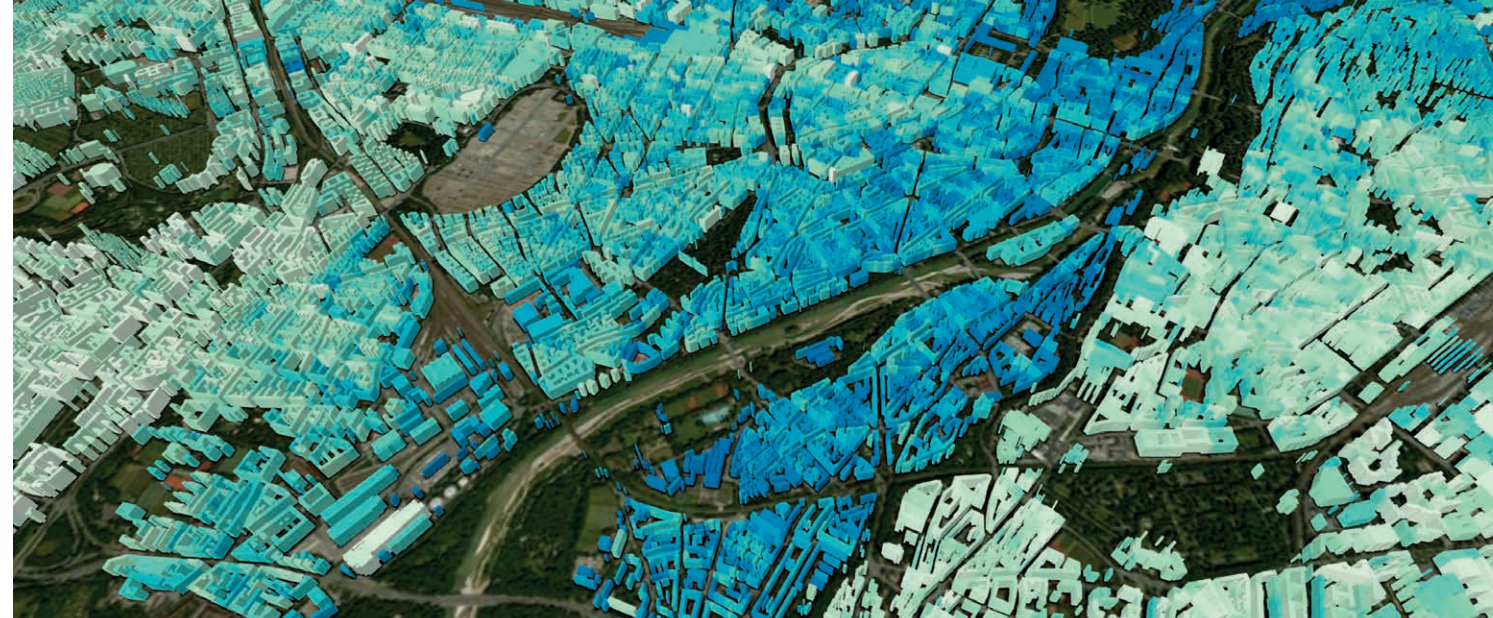
■ Beides. Durch Neuentwicklungen wie das Copernicus-Satellitenprogramm gibt es so viele frei zugängliche Daten von Erdbeobachtungssatelliten, dass die klassischen Verfahren nicht mehr ausreichen, um effizient Informationen daraus abzuleiten. Heute sind das schon 20 Petabyte, in drei Jahren werden es 60 Petabyte sein. Wir brauchen dafür datengetriebene Analysen – und das kann die künstliche Intelligenz.

Welche Informationen können Sie denn aus der Erdbeobachtung mit künstlicher Intelligenz gewinnen?

■ Zurzeit arbeite ich mit meinem Team und anderen Kollegen aus dem Erdbeobachtungszentrum des DLR unter anderem an der globalen Kartierung von Städten. Schon jetzt wohnen mehr Menschen in Städten als im ländlichen Raum. Als Informationsbasis auf globaler Ebene haben wir nur ganz wenige Informationen – oftmals nur die Unterscheidung, wo es Gebäude gibt und wo nicht. Viele der Migrationsströme vom Land in die Stadt enden in Slums oder in formellen Siedlungen, und wenn man dort wirklich etwas verbessern möchte, braucht man zuerst bessere Informationen. Wir fusionieren zum Beispiel optische Daten aus dem Copernicus-Programm mit denen unserer deutschen Radarsatelliten TerraSAR-X und TanDEM-X und generieren 3D-Modelle von allen Gebäuden. Wir können dann auch auf die Funktion der Gebäude schließen und zum Beispiel die Bevölkerungsdichte ableiten.

Sie haben ja schon erwähnt, dass unterschiedliche Daten anfallen. Sie verwenden aber auch Social-Media-Daten. Wie bringen Sie das zusammen?

■ Optische Daten liefern uns die Umrisse von Gebäuden, Radardaten helfen, die Höhe der Gebäude abzuleiten. Wenn man diese zwei Informationen kombiniert, bekommt man ein wertvolles 3D-Modell. Und hier verwenden wir als Methodik natürlich Deep Learning und andere Verfahren des maschinellen Lernens. Wenn wir zum Semantischen kommen, also zum Beispiel zwischen Wohnhäusern und kommerziellen Gebäuden unterscheiden wollen, dann ist es noch komplizierter. Was wir dann zusätzlich verwenden können, sind Social-Media-Bilder: Die sind vom Boden aus aufgenommen, und man sieht nicht nur das Dach, sondern auch die Fassade. Dann können wir noch Textnachrichten von Social Media verwenden. Wir wissen beispielsweise, dass in einem Wohngebäude morgens und abends viele Tweets versendet werden, in einem Bürogebäude hingegen vor allem tagsüber. Man kann auch die Inhalte der Tweet-Texte analysieren, denn in Büro- und Wohngebäuden werden unterschiedliche Themen diskutiert.



Bilder: ©Xiaoxiang Zhu, DLR/TU München

Dieses dreidimensionale Modell von München wurde mit nur fünf TanDEM-X-Aufnahmen erstellt. Die Grundflächen der Gebäude wurden aus optischen Daten rekonstruiert und die Höhen (Farben) aus der Punktwolke von TomoSAR (Tomographic synthetic aperture radar). Wenn man beispielsweise global Probleme urbaner Siedlungen lösen möchte, braucht man genaue Informationen über die Gebäude. Diese bekommen die Wissenschaftler durch intelligentes Kombinieren unterschiedlicher Datenquellen.

Wenn man mit solchen persönlichen Daten arbeitet – wie ist dann der Datenschutz gewährleistet? Beschäftigen sich die Wissenschaftler auch damit?

■ Sicher! Wir erhalten die Daten natürlich ohne Nutzerinformationen und Nutzernamen, und auch nur von den Nutzern, die der wissenschaftlichen Verwertung ihrer veröffentlichten Daten zugestimmt haben. Zudem ist unser Ziel, global konsistente Informationen abzuleiten – da arbeiten wir nicht mit hoher Auflösung.

Wie filtert man denn genau jene Tweets heraus, die verwertbare Informationen bieten?

■ Da sind wir beim Thema Data-Mining. Wir entwickeln Tools, mit denen wir in einem Gebiet zufällig ausgewählt ein Prozent der Tweets abfragen können. Dann müssen wir Algorithmen finden, um diesen unstrukturierten Tweets eine Struktur zu geben. Anschließend folgt das Information-Mining: Was sind die relevanten Informationen, welche inhaltlichen Aussagen kann ich ableiten? Schließlich muss ich diese Social-Media-Daten mit meinen Satellitendaten kombinieren.

Kommen wir noch einmal auf den Nutzen zurück: Wer muss wissen, wie hoch Gebäude sind oder wie sie genutzt werden?

■ Ein Beispiel: Im indischen Mumbai gab es in den letzten Dekaden 50.000 Feuer. Rund zwei Drittel davon entstanden durch falsche und illegale Verlegung und Unterdimensionierung von Stromkabeln. Wenn ich diese Gefahr ausschalten will, brauche ich Informationen: Wie viele Gebäude habe ich hier? Wie viele Leute wohnen darin? Mit solchen Basisinformationen können Wasserverteilung, Ausbildung oder auch das Gesundheitswesen besser organisiert werden. Was wir mit künstlicher Intelligenz und Big Data schaffen, hilft den Menschen in Entwicklungsregionen.

Glauben Sie, dass die KI auch Ihre Arbeit oder die Ihrer Kollegen verändern wird? Werden zukünftig weniger Wissenschaftler benötigt?

■ Wenn ein Algorithmus nur gelernt hat, was eine Katze und was ein Hund ist, wird er nicht wissen, was ein Auto oder ein Flugzeug ist. Das heißt, die Menschen sind in der Entwicklung und im Training der

Algorithmen sehr wichtig. Der Bedarf an Informatikern und Datenwissenschaftlern ist sehr groß, und diese Leute zu bekommen, ist super schwierig. Internetfirmen und die Autoindustrie sind bei der Suche starke Konkurrenten. Dann muss man diese Experten auch erst einmal für die Erdbeobachtung begeistern. Daher investieren wir auch in die Lehre. Es gibt in der Erdbeobachtung so viele wertvolle Daten, dieses Potenzial nutzen wir noch nicht einmal ansatzweise. Und es gibt noch viele innovative Anwendungen. Das werden wir auch den Data Scientists und Informatikern nahebringen. Dann werden sie den Marktwert der Erdbeobachtung besser erkennen und das Interesse wird steigen.

Wie weit sind Sie bei der Verknüpfung der künstlichen Intelligenz mit der Beobachtung des globalen Wandels?

■ Beim maschinellen Lernen in der Erdbeobachtung sind wir schon seit Jahrzehnten unterwegs; die datengetriebene Analyse von Satellitendaten ist seit 30 Jahren ein klassisches Thema am Erdbeobachtungszentrum des DLR. Aber jetzt ist es ein heißes Thema geworden: Erstens durch die große Menge an Satellitendaten, die heute verfügbar ist. Zweitens gibt es jetzt so viel Rechenpower, dass man große Modelle trainieren kann. Und drittens gibt es Durchbrüche im maschinellen Lernen durch Deep Learning. An der Verwendung von Daten aus den sozialen Medien forschen wir allerdings erst seit zwei, drei Jahren.

Wenn Sie in 30 oder 40 Jahren auf die Entwicklung der künstlichen Intelligenz zurückblicken – was würden Sie dann gerne erlebt haben?

■ Worüber ich mich sehr freuen würde, wäre, wenn wir diesen Hype der KI zu einer Bedeutung für die KI in der Erdbeobachtung gebracht hätten. Wenn dann durch die Verfahren, die wir entwickelt haben, ein Mehrwert geschaffen worden ist. Wenn für gesellschaftliche Probleme, für die es keine Lösung gab, durch unsere Geoinformationen Lösungen gefunden worden sind.



Das Interview führten **Manuela Braun** und **Peter Poete**, strategische Kommunikation Raumfahrt beziehungsweise Think Tank im DLR