A Clear View into the Data Stream

In an effort to ensure that we make optimal use of the incredible amounts of data generated by satellites, researchers from nine nations have come together to work on AI4EO, a project spearheaded by TUM. The project's output is publicly available and, it is hoped, could help to tackle many social challenges – such as in megacities – and visualize changes in land use and the impact of climate change.

Gesamter Artikel (PDF, DE): www.tum.de/faszination-forschung-27

Durchblick im Datenstrom

D

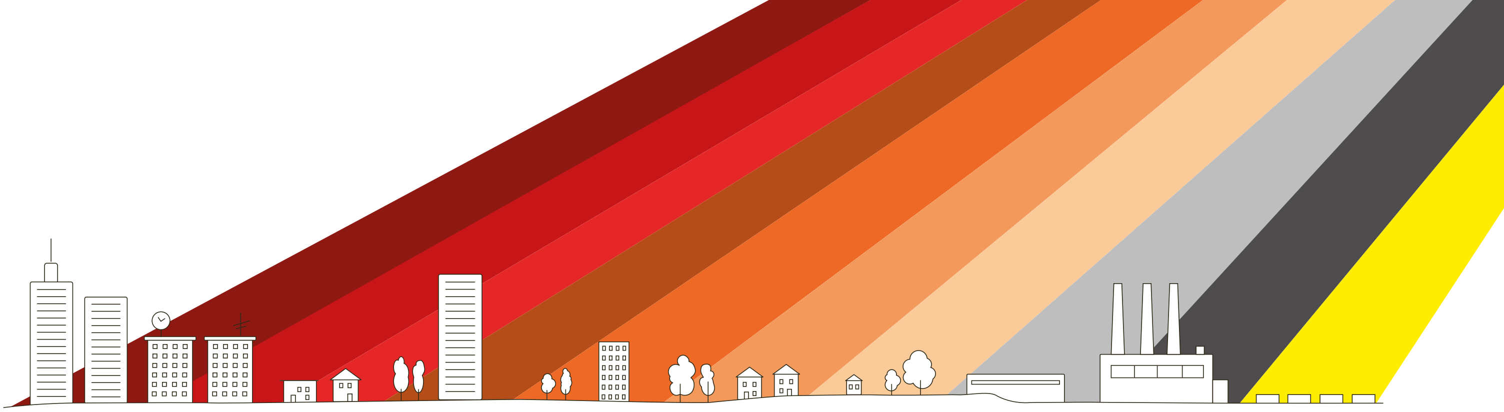
Im Projekt AI4EO entwickeln Forscherinnen und Forscher aus neun Nationen neuartige Methoden des Maschinellen Lernens und Lösungen für die Analyse von großen Datenmengen, um so den stetig anschwellenden Satellitendatenstrom zu analysieren. Xiaoxiang Zhu, Professorin für Data Science in der Erdbeobachtung an der TUM und am DLR, leitet das Projekt. □

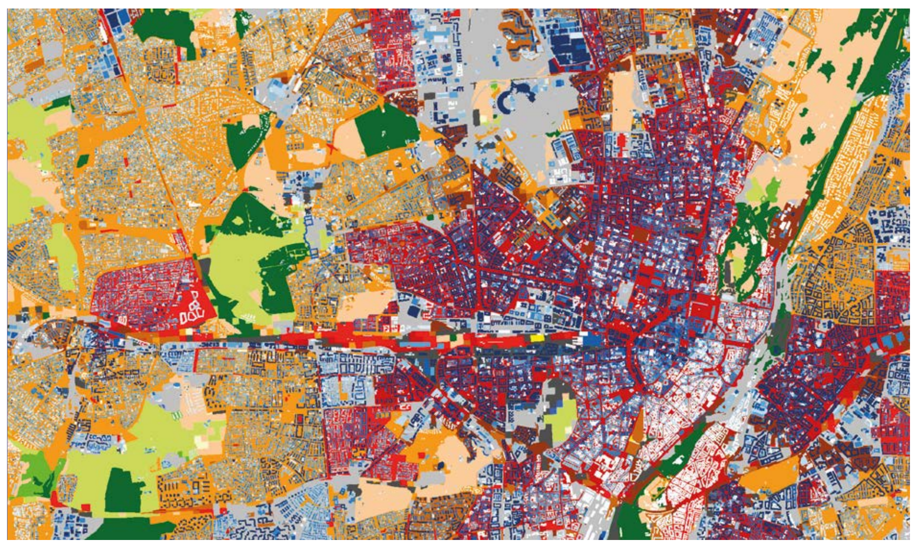
Graphics and picture credit: ediundsepp

Xiaoxiang Zhu is Professor of Data Science in Earth Observation at TUM as a joint appointment together with DLR.

Machine learning and other artificial intelligence techniques help extract specific information from the great wealth of satellite data.

In recent years, satellite data has become one of the most important sources of data for earth observation. It reveals how fast megacities around the world are spreading, where arable land is disappearing and nature reserves are shrinking. The problem, however, is that satellites nowadays send so much data down to earth that it is impossible to maintain an overview without technological assistance. For instance, around 40 billion megabytes of data are currently stored at the Remote Sensing Data Center (DFD) at the German Aerospace Center (DLR). This data volume could fill the storage capacity of 625 million smartphones and is forecast to more than double by 2030. In the last year, researchers from nine countries have come together to launch AI4EO (Artificial Intelligence for Earth Observation), an international future lab directed by TUM aiming to facilitate effective use of this data in future. These researchers have pooled their expertise to develop novel machine learning methods and big data analytics solutions, thereby making it possible to analyze the ever-rising tide of satellite data. “Artificial intelligence enables us to identify unknown objects and relationships – but that’s not all,” emphasizes Prof. Xiaoxiang Zhu, data scientist at TUM and director of AI4EO. “The algorithms we have developed are also able to uncover unidentified phenomena. This will make it possible to use satellite data in an even more multifaceted manner in future.” ▷





With the help of AI, urban climate zone classes could be automatically extracted from satellite data. The pictures show the world's first global urban local climate zone map (below) and the local climate zone map of Munich overlaid on the 3D building model (right).

Graphics: ediundsepp (source: TUM/Zhu); Picture credits: TUM/Zhu

- Dense trees
- Scattered trees
- Bush, scrub
- Low plants
- Bare rock or paved
- Bare soil or sand
- Water

Recording three billion buildings worldwide

Examples of this type of valuable new information can be found in the machine learning methods and big data analytics solutions developed by Zhu and her colleagues, which have made it possible to create a complete map of all buildings around the world for the first time. Map services like Google Maps and OpenStreetMap also provide detailed maps. However, they often identify built-up spaces simply as "urban areas" – and only depict individual houses and buildings to a limited extent. "This data is enormously important for urban planning," says Zhu. One example is cities in which "informal settlements", or slums, are constantly growing. Experts estimate that there are around three billion buildings around the world. Yet only 16 percent of these buildings are individually depicted in OpenStreetMap, with information on building height only available for 0.5 percent. In the So2Sat project, the TUM team successfully harnessed intelligent algorithms to extract detailed building information and produce maps complete with individual buildings for locations around the world, including Cairo and other major African cities. "We make our datasets publicly available," underscores Zhu, "which means they can be used in different ways."

The urban microclimate

Researchers at the AI4EO future lab use an abundance of these datasets, which developers around the world can also utilize to design new satellite information services. One example is So2Sat LCZ42, a dataset in which Xiaoxiang Zhu's group combine satellite data on cities with particularly current information: the urban structures that shape the urban climate, namely densely packed high-rise buildings that cause the summer heat to build up, and green spaces that provide fresh air. The team has identified 17 such characteristic urban climate zone classes that are now automatically extracted from satellite data using artificial intelligence. This tool and the dataset are also particularly valuable for urban planners.

The TUM group in the AI4EO future lab has developed other solutions, such as a program that provides information updated daily on changing land use – areas that have been built on or lost due to flooding. Other outcomes include a machine learning pipeline that automatically removes intrusive clouds from satellite images and a complex analytical algorithm that compares satellite images to predict future events – such as streets that could soon be covered by rising waters. "In a film sequence taken from observation cameras, you quickly see what's changing," says Zhu. "However, we have to work with individual images that satellites take during several flyovers. Identifying a development from these images is much harder. But our algorithm can do it."

The algorithms identify far more events than the naked human eye. Like everything produced by the AI4EO future lab, these innovations and datasets are publicly available. Further tools are set to be added in the near future, among them a visual question answering (VQA) system based on natural language processing. It can examine satellite images directly and answer questions like "What changes have happened in this image?" in a way that people can understand. This system is the first of its kind, a sort of Alexa or Siri for earth observation, and will screen the wealth of earth observation data to evaluate the highly dynamic changes underway around the world.

■ *Tim Schröder*

Link
www.asg.ed.tum.de/sipeo www.ai4eo.de www.so2sat.eu