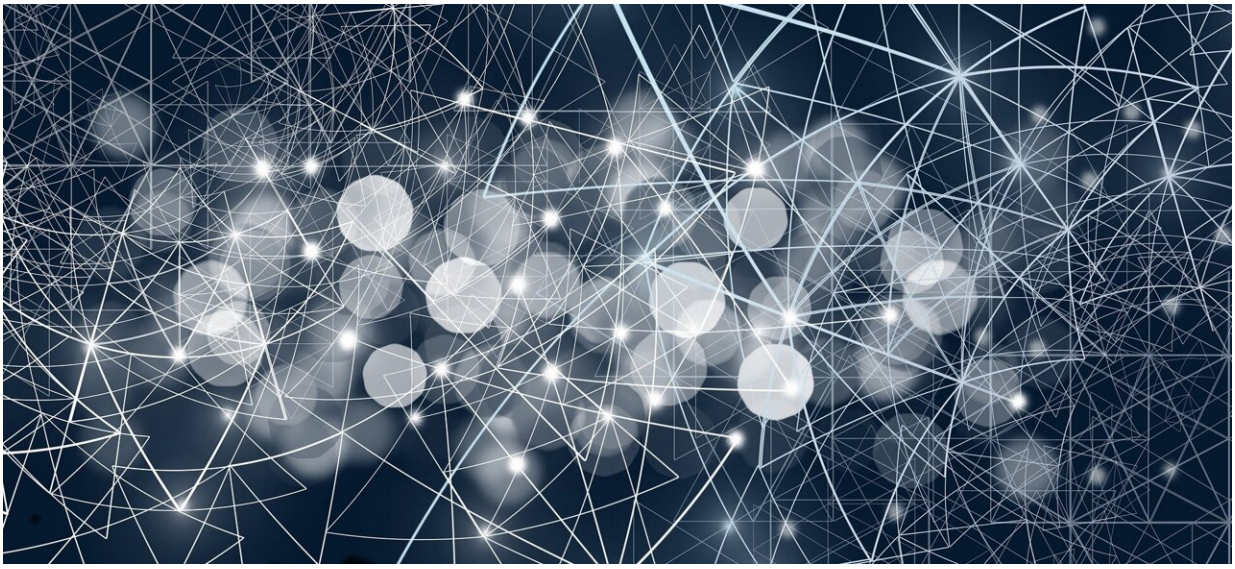
Scientists automate core box image recognition

June 30 2022



Credit: Pixabay/CC0 Public Domain

Skoltech researchers have trained a neural network to recognize rock samples in core box images efficiently. It sped the analysis process by up to 20 times and made it possible to automate the description of rock samples. The developed algorithm is used in the DeepCore system—a digital geological exploration service created by Digital Petroleum, a spin-off of Skoltech. The details of the method are described in the article published in *Computers & Geosciences*.

One of the routine tasks of geological research is the description of rock samples. In many cases, the extracted rock core is stacked in boxes. Scientists take photographs of boxes or columns during the core study. The description is compiled manually by filling out spreadsheets or geological journals. The standard analysis procedure involves manual extraction of columns from photographs of boxes in a graphical editor. This is a rather time-consuming process.

To automate this process, scientists have used machine learning methods. However, traditional computer vision algorithms perform poorly at this task due to the limited amount of data and large differences between images. For example, if the core column differs in color or texture from adjacent ones or ones photographed in different conditions. Such differences significantly affect the performance of machine learning algorithms, which require a large data set describing all possible variants. As a result, one has to spend time retraining the model.

To solve this problem, Skoltech scientists used deep convolutional neural networks—artificial neural networks that are similar in structure with the visual cortex of animals. To train the neural network, the scientists used augmentation that added modified copies of core boxes' photos to increase the amount of data. "Synthetic" images were created based on a modified CutMix algorithm. The CutMix algorithm creates a new image from a pair of existing ones by randomly cutting out a piece of one image and inserting it into another. Since the scientists were specifically interested in recognizing rock columns, they optimized this method based on a core image template, cutting and swapping pieces only from the areas where the core was located.

"Core boxes photographed in the same field may be visually very similar, but the rocks may differ. If rock from another box is virtually placed in the same box, the network can confuse the core area with the box boundaries due to the similarity in color. Augmentation helps the

network to focus on other characteristics besides color and shape, such as structure and texture," explains the first author of the work, Skoltech scientist Evgeny Baraboshkin.

In their study, the scientists described and tested the new method and compared the efficiency of the algorithm trained on "original" and mixed with augmented data. It turned out that due to augmentation, the algorithm is trained to detect rock columns efficiently and accurately in most of the new images. This automated approach speeds up the processing of one core box up to 20 times. In addition, the method made it possible to determine automatically the depths corresponding to each column. Previously this required measuring with a ruler.

"Interestingly, when we added augmented data into the usual data set, the neural network learned to recognize pieces of paper with inscriptions on the columns, although in the original dataset they were also labeled as core. The algorithm detected an error in the initial markup and avoided it in the future," Evgeny adds.

The scientists introduced the developed method as one of the stages of analysis into the DeepCore system, a software product they created for an automatic core description from images. After extracting columns from images, the program determines the layer boundaries and rock types. At the same time, users still have the possibility of maintenance. If necessary, the expert can add additional types of rocks or change layer boundaries. Since 2021, DeepCore has been used in the extractive industry, helping specialists reduce routine work time and automate analysis.

More information: Evgeny E. Baraboshkin et al, Core box image recognition and its improvement with a new augmentation technique, *Computers & Geosciences* (2022). DOI: 10.1016/j.cageo.2022.105099

Provided by Skolkovo Institute of Science and Technology