

RESEARCH NEWS

February 2017 || Page 1 | 3

CeBIT 2017: Analysis software for neural networks

Watching computers think

Neural networks are commonly used today to analyze complex data – for instance to find clues to illnesses in genetic information. Ultimately, though, no one knows how these networks actually work exactly. That is why Fraunhofer researchers developed software that enables them to look into these black boxes and analyze how they function. The researchers will present their software at CeBIT in Hannover from March 20 to 24, 2017 (Hall 6, Booth B 36).

- The software recognizes which parameters a neural network uses to make decisions.
- It is of interest for any area in which data such as text, images and signals are automatically processed and evaluated by neural networks.
- For example, the software can analyze complex production data and deduce which parameters are decisive for product quality or possible deviations.

Sorting photos on the computer used to be a tedious job. Today, you simply click on face recognition and instantly get a selection of photos of your daughter or son. Computers have gotten very good at analyzing large volumes of data and searching for certain structures, such as faces in images. This is made possible by neural networks, which have developed into an established and sophisticated IT analysis method (see box, "How neural networks function").

The problem is that it isn't just researchers who currently don't know exactly how neural networks function step by step, or why they reach one result or another. Neural networks are, in a sense, black boxes – computer programs that people feed values into and that reliably return results. If you want to teach a neural network, for instance, to recognize cats, then you instruct the system by feeding it thousands of cat pictures. Just like a small child that slowly learns to distinguish cats from dogs, the neural network, too, learns automatically. "In many cases, though, researchers are less interested in the result and far more interested in what the neural network actually does – how it reaches decisions," says Dr. Wojciech Samek, head of the Machine Learning Group at Fraunhofer Heinrich Hertz Institute HHI in Berlin. So Samek and his team, in collaboration with colleagues from TU Berlin, developed a method that makes it possible to watch a neural network think.

Contact

Janis Eitner | Fraunhofer-Gesellschaft, Munich | Communications | Phone +49 89 1205-1333 | presse@zv.fraunhofer.de
Anne Rommel | Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute HHI | Phone +49 30 31002-353 | anne.rommel@hhi.fraunhofer.de
Einsteinufer 37 | 10587 Berlin | www.hhi.fraunhofer.de

Machine learning enables customized cancer treatments

This is important, for instance, in detecting diseases. We already have the capability today to feed patients' genetic data into computers – or neural networks – which then analyze the probability of a patient having a certain genetic disorder. "But it would be much more interesting to know precisely which characteristics the program bases its decisions on," says Samek. It could be certain genetic defects the patient has – and these, in turn, could be a possible target for a cancer treatment that is tailored to individual patients.

Neural networks in reverse

The researchers' method allows them to watch the work of the neural networks in reverse: they work through the program backwards, starting from the result. "We can see exactly where a certain group of neurons made a certain decision, and how strongly this decision impacted the result," says Samek. The researchers have already impressively demonstrated – multiple times – that the method works. For instance, they compared two programs that are publicly available on the Internet and that are both capable of recognizing horses in images. The result was surprising. The first program actually recognized the horses' bodies. The second one, however, focused on the copyright symbols on the photos, which pointed to forums for horse lovers, or riding and breeding associations, enabling the program to achieve a high success rate even though it had never learned what horses look like.

Applications in big data

"So you can see how important it is to understand exactly how such a network functions," says Samek. This knowledge is also of particular interest to industry. "It is conceivable, for instance, that the operating data of a complex production plant could be analyzed to deduce which parameters impact product quality or cause it to fluctuate," he says. The invention is also interesting for many other applications that involve the neural analysis of large or complex data volumes. "In another experiment, we were able to show which parameters a network uses to decide whether a face appears young or old."

According to Samek, for a long time banks have even been using neural networks to analyze bank customers' creditworthiness. To do this, large volumes of customer data are collected and evaluated by a neural network. "If we knew how the network reaches its decision, we could reduce the data volume right from the start by selecting the relevant parameters," he says. This would certainly be in the customers' interests, too. At the CeBIT trade fair in Hannover from March 20 to 24, 2017, Samek's team of researchers will demonstrate how they use their software to analyze the black boxes of neural networks – and how these networks can deduce a person's age or sex from their face, or recognize animals.

RESEARCH NEWS

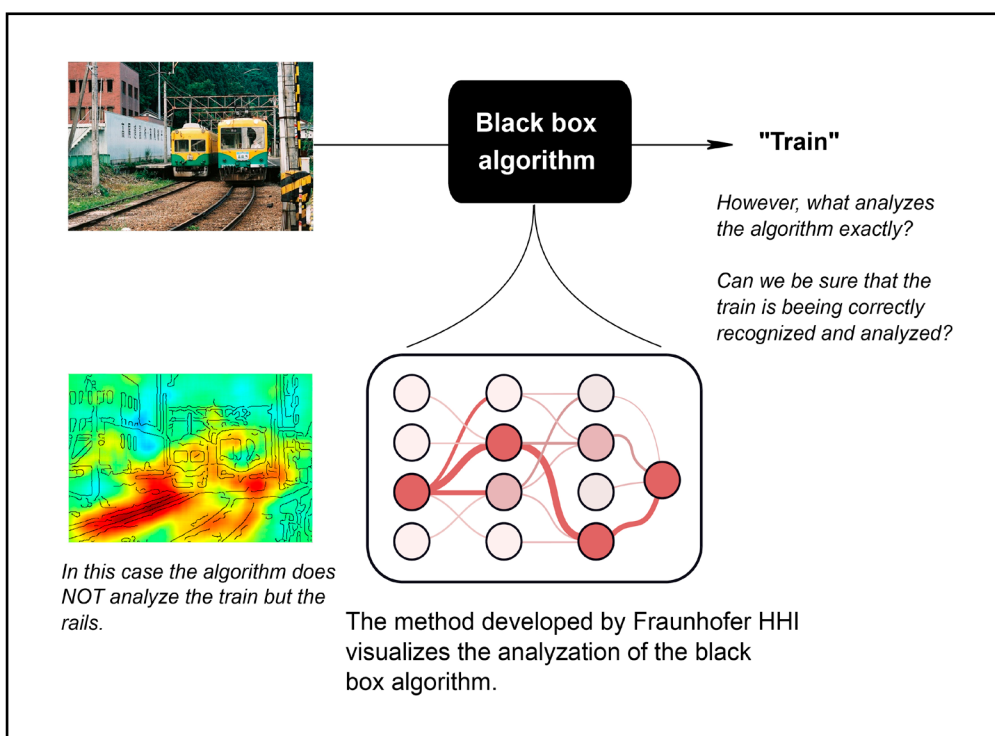
February 2017 || Page 2 | 3

How neural networks function

Neural networks function, as their name suggests, based on the same principle as the brain. Just as various areas of nerve cells, the neurons, cooperate in our brain, mathematical units work together in the artificial neural network. A photo that, to a computer, is initially only an unordered cloud of pixels, is systematically analyzed by these groups of artificial neurons. Some neurons recognize edges while others recognize corners. Bit by bit, the image becomes clear to the computer. Ultimately, the outcome of this neural computing operation is a probability value – the computer calculates, for example, how likely it is that a face actually does correspond to that of the person being sought.

RESEARCH NEWS

February 2017 || Page 3 | 3



Fraunhofer HHI's analysis software uses algorithms to visualize complex learning processes (schematic diagram). © Fraunhofer HHI | Picture in color and printing quality: www.fraunhofer.de/en/press