

# Algorithms obtain the maximal amount of knowledge from images

January 5 2018

---



Professor Stefan Roth. Credit: Technische Universität Darmstadt

Digital images and videos contain much more information than computers currently extract from them. With the help of intelligent algorithms, a research team led by Professor Stefan Roth aims to obtain

the maximal amount of knowledge from images.

A typical street scene can be seen on the screen in Stefan Roth's office – but from the 'viewpoint' of a [computer](#). Cars tinted red pull in and out of parking spaces, purple pedestrians bustle about, green-marked plants indicate the verge. "For the computer, a video first of all only consists of pixels," explains computer science professor Stefan Roth. "We teach it to interpret the pixels," adds the head of the Visual Inference Lab at Technische Universität Darmstadt.

Roth's team teaches intelligent algorithms to detect cars, pedestrians, or even potentially dangerous objects in X-ray images from transportation security. The software developed by the scientists of TU Darmstadt also reconstructs the image information that may be hidden behind blurred or out-of-focus images. The research question that guides them:

How much information can be extracted from a digital image? The need for automatic image analysis is huge. Millions of digital cameras create an unprecedented flood of images. If computers could reliably interpret not only ordered road scenes such as on a motorway, but also traffic that may appear rather chaotic, for instance at a junction, "then fully autonomous driving would also be possible in busy inner cities," says Roth. "There are many other potential fields of application," adds the computer scientist. Intelligent image analysis systems could assist users in tedious tasks, such as baggage control at airports. Land use can be automatically classified in satellite images, for example to ascertain on which fields wheat grows.

But teaching computers to see is difficult. Decades ago, researchers tried to directly create programs that imitate human perception. But this was largely unsuccessful, at least so far. "Today's approaches are very data-driven," says Roth. Computers learn by means of a large quantity of examples. The basis are often so-called [artificial neural networks](#). These

are inspired by the structure of the brain: Nerve cells, referred to in technical language as neurons, are interconnected by [neural pathways](#). When photos of cars are shown to such a network, recurring patterns such as chassis, wheels, and headlights, reinforce certain neural pathways. If similar patterns appear on unknown photos, the same neurons become active via the intensified neural pathways as during training: The neural network has learned to recognise cars in images. Or pedestrians and plant pots.

The catch: During training one has to literally show the computer on each sample image where the car is, where the pedestrian is, and where the plant pot is. "This used to take us an hour and a half per image at the beginning," says Roth. Because computers only reliably recognise objects after tens of thousands of examples, that is not always practical. "For this reason, we first of all try to get by with less data and secondly, aim to access data sources that already contain some of the information," says Roth. Computer games, for instance, show deceptively realistic street scenes. On a photo of a real scene, the researchers first have to painstakingly separate the individual objects from each other by tracing their outlines. "In a computer game, however, the individual objects are already separated," explains Roth. Then one only has to tell the neural network where the cars and the road surface are.

To get by with less data, the researchers come up with more tricks. "Based on the information contained in the computer game, we can detect which object that is already known re-appears at a later point in time," explains Roth. This means that the object, for example a particular car, no longer needs to be re-annotated on each frame of a video sequence.

The success of the approach developed by the scientists of TU Darmstadt is made apparent by the computer-interpreted video of a busy shopping street. Even further down the street, distant pedestrians and

vehicles are detected.

Provided by Technische Universität Darmstadt

Citation: Algorithms obtain the maximal amount of knowledge from images (2018, January 5)  
retrieved 25 April 2024 from

<https://phys.org/news/2018-01-algorithms-maximal-amount-knowledge-images.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.