Human vision—a challenge for AI

October 22 2022

If the camera is comparable to the human eye in computer vision, the code is certainly the brain, according to Michael Felsberg. Behind the code, there is a lot of advanced mathematics. Credit: Thor Balkhed

Achieving diversity in human vision is one of the major challenges for AI research. In the vast majority of cases, we are better than machines at understanding the world around us. But machines are catching

up—slowly but surely.

"Within a single day we humans can go from driving a car to free diving, and continue to reading the newspaper and navigating a dense forest—all without a great deal of effort. For a robot, doing the same things would currently be impossible," says Michael Felsberg, professor at Linköping University and one of Sweden's foremost researchers in computer vision and [artificial intelligence](#) (AI).

That we humans can do all this, and much more, is largely due to vision. Estimates say that some 80% of our impressions reach us by way of our vision. It is the single most important sense for perceiving what happens around us. Michael Felsberg's research focuses mainly on what is called the artificial visual system, where the aim is to get computers to see as well as humans do.

"Biological systems simply work. Humans are remarkably skilled in general perception and analysis, skills we want to emulate in computers. Today we can build technical systems that are good at doing a particular task, such as [self-driving vehicles](#). But if in the future we want to be able to collaborate with robots, they must be able to see and understand exactly what we see," says Michael Felsberg.

Imitating human vision might seem easy at first glance. When AI research began, the feeling was that computer vision would be solved with a simple camera—maybe a project for the summer break. Now, almost 60 years later, general computer vision has developed into one of the most salient challenges in AI research.

The code is the brain

Michael Felsberg and his co-workers test many of the solutions they develop in the vision laboratory on Campus Valla in Linköping. For

instance, between the huge glass walls, autonomous drones and small self-driving cars equipped with advanced sensors and cameras are test-driven. But the actual brain in the computer vision is behind the lens.

"The camera is just a light sensor; it can't do anything else. The actual work is done by the code and the software behind the camera. It's the same with people: the eye registers the light and the brain does the work," says Michael Felsberg.

There have been many attempts to emulate the human brain—with varying results. Today, a method of machine learning called deep learning is usually used. Put simply, it means that the computer learns its models organized in neural networks from large amounts of data. The algorithms are fed with huge amounts of data, which are analyzed on several levels. This might sound complicated, and it is. The truth is that no one can say exactly what happens in every activation in a deep network.

Michael Felsberg draws parallels to the human brain:

"On a brain scan you can see which parts of the brain are active during different stimuli. But we still don't know what actually happens and how a thought is formed in the brain. Deep learning works in a somewhat similar way. We see that it works, but not how it works in detail," he says.
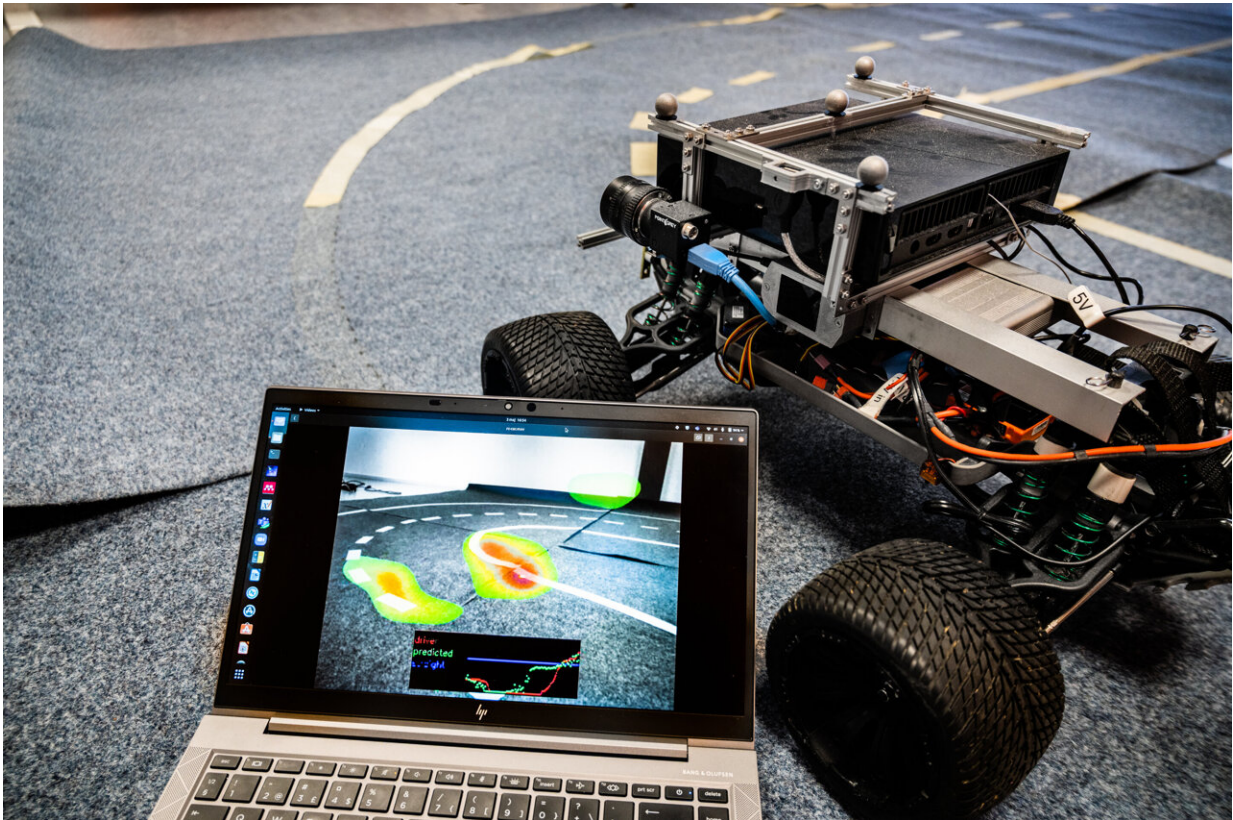
Autonomous vehicles and drones are some current areas of application for the research conducted by Michael Felsberg and his group. Credit: Thor Balkhed

The way forward

But why is it so difficult for a computer to see what we see? The answer lies in our ability to rapidly adapt to different situations, and the [feedback loop](#) between our perception of our surroundings and our constantly active cognitive ability.

Looking out through a dirty window pane is an everyday example of a situation where computers struggle but we humans manage swimmingly. We see immediately what's going on outside the window, despite our slightly obstructed vision. On the other hand, a computer will first auto-

focus on the dirt on the pane. But once it has found the right focus—on the scene outside—it still won't fully understand what is happening, because some of the view is blocked by the dirt.

Still, there are areas where computers already see better than humans—in particular when it comes to exact calculations and assessments of distances, temperatures and patterns. In these cases, computer vision can complement our own vision, rather than draw its own conclusions and act on them.

"A technical system works well as long as everything is as expected. But faced with something unexpected, it will have problems. We must work to make the systems more robust," says Michael Felsberg.

But developing software that can surpass the flexibility of human vision takes time. And according to Michael Felsberg, research must take time if it is to be robust. Science is a process, and every new research article adds another little piece to a massive puzzle. Breakthroughs that give research a huge leap forward are very rare.

"General situational awareness in a computer could possibly exist in our lifetime. But creating the link between cognition and general situational awareness in a computer is probably very far off in the future," says Michael Felsberg.

Once general computer vision exists, he believes there will be many different applications, e.g. social robots, safer autonomous vehicles and more efficient production. But AI is not uncontroversial. Many fields of use risk encroaching on individual privacy when large volumes of personal data are processed.

For this reason, Michael Felsberg and his research team are focusing on how AI can give better insight into how we can prevent additional

climate change:

"Climate change is one of humanity's greatest threats. Using advanced computer vision, we will be able to rapidly analyze large tracts of land, and their importance for the climate. What would take humans several years to map out manually could potentially be finished in a few weeks with the help of AI."

Provided by Linköping University