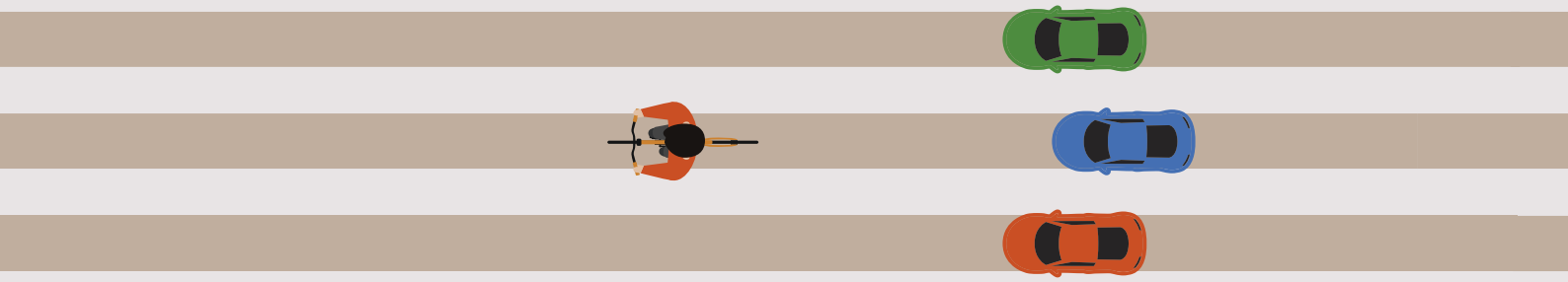
SEEING MORE WITH LESS

By Bennie Mols Image iStock



In a new NWO Vici project, Jan van Gemert aims to develop computer vision systems that achieve state-of-the-art accuracy with a hundred times less data.

In a short essay called 'The Bitter Lesson', Turing Award winner Richard Sutton stated that success in AI does not come from explicitly incorporating human knowledge, but from building systems that can learn and discover on their own, fuelled by ever-increasing computing power and the abundance of data. Examples abound, Sutton writes: in games like computer chess and go, but also in machine translation and computer vision.

Jan van Gemert, head of the computer vision lab at Delft University of Technology, refers to Sutton's essay to characterise the current trend in AI: 'It is all about more data and more compute. At companies like Meta or Google, a single person can use more GPUs than our entire department has. I am no longer even trying to compete with that. In my research, I'm deliberately going against that trend.'

Using prior knowledge

Van Gemert received an NWO Vici grant for his 'Project dAIta: Data Efficient AI Foundation Models'. 'My hypothesis is that you need much less training data in computer vision when you use prior knowledge about physical properties of the world', Van Gemert explains. 'And using less data not only means less computation but also fewer biased outcomes, less privacy and copyright issues, and a decreased dependency on Big Tech companies.'

Van Gemert started the Vici project in July 2025. Three PhD students will each work on a different subtheme, followed by the hiring of a postdoc to integrate their research. Van Gemert: 'Each PhD student will explore a different aspect of prior world knowledge: one will focus on depth, another on lighting conditions, and the third on the compositional structure of objects.' Neural networks recognise objects by colour differences, such as a bicycle in front of a red, green, or blue car. They memorise those colour differences by heart, which demands many data samples per colour. 'But if we use estimated depth differences instead, one example is enough', states Van Gemert, explaining the first idea of using prior knowledge. 'This will also work for rare colours like pink cars that are not in the data.'

In the second idea, different lighting conditions are explored. For a camera, a bicycle looks different under cloudy or sunny skies, but again, it's still the same bicycle. With the third idea, Van Gemert will allow the system to learn that a blue and a red variant of the same bicycle differ only in colour, but not in shape. For current systems, those two bikes are totally different. When they have learned to recognise a blue bike, they have to start again from scratch to recognise a red bike.

Tracking improvement

Van Gemert: 'I hope that the Vici project will allow computer vision systems with a hundred times less data to achieve the same accuracy as state-of-the-art systems.' He wants to do that with a novel approach. Many AI systems are currently built by trial and error. Then at the end of the run people say, "look, it works". In my opinion, too little research is done on where exactly the improvement comes from. I'm more into understanding, rather than trial and error.' He points to the whiteboard behind him, on which words like *Dlaczego? Pse? Neden? De ce?* are written. 'The question I ask my students most often is "why?", so I asked each of them to write the why-question in their own language on the whiteboard.'

More information
jvgemert.github.io