



Stop twice: Different road markings, such as in the UK (left) and South Korea, currently make it difficult to train AI systems.



The little difference

Text: Constantin Gillies

New situations are challenging for autonomous vehicles. As things stand today, if traffic signs or rules change the vehicles have to be completely retrained. The AI Delta Learning research project funded by the German Federal Ministry for Economic Affairs and Energy aims to solve this problem and thereby reduce the effort involved substantially.

Stop signs look similar in many countries—red, octagonal, with the word “STOP” in the middle. There are exceptions, however: In Japan the signs are triangular; in China the word “STOP” is replaced by a character; in Algeria by a raised hand. Non-local drivers have no problem with these little differences. After the first intersection at the latest, they know what the local stop sign looks like. The artificial intelligence (AI) in an autonomous vehicle, on the other hand, needs to be completely retrained to be able to process the small difference.

These ever-changing lessons take a lot of time, generate high costs, and slow down the development of autonomous driving as a whole. So the automotive industry is now taking a joint step forward: The AI Delta Learning project aims to find ways to selectively teach autonomous vehicles something new. Or to stick with our example: In the future, all you want to have to do is tell the autopilot: “Everything remains the same except the stop sign.”



“The objective is to reduce the effort required to be able to infer things from one driving situation to another without having to train each thing specifically.”

Dr. Joachim Schaper, Senior Manager AI and Big Data at Porsche Engineering



70,000

graphics processor hours are needed to train an autopilot.



100

experts from 18 partners are working on the AI Delta Learning project to find methods for knowledge transfer.



50%

less effort: AI Delta Learning could halve human labor input.

The importance of the task is demonstrated by the list of participants in the project, which is being funded by the German Federal Ministry for Economic Affairs and Energy: In addition to Porsche Engineering, partners in the project include BMW, CARIAD, and Mercedes-Benz, major suppliers such as Bosch, and nine universities, including the Technical University of Munich and the University of Stuttgart. “The objective is to reduce the effort required to be able to infer things from one driving situation to another without having to train each thing specifically,” explains Dr. Joachim Schaper, Senior Manager AI and Big Data at Porsche Engineering. “The cooperation is necessary because currently no provider can meet this challenge alone.” The project is part of the *AI family*, a flagship initiative of the German Association of the Automotive Industry aimed at advancing connected and autonomous driving.

Roughly 100 people at a total of 18 partners have been working on AI Delta Learning since January 2020. Workshops are held at which experts exchange views on which approaches are promising—and which have proved to be dead ends. “In the end, we hope to be able to deliver a catalog of methods that can be used to enable knowledge transfer in artificial intelligence,” says Mohsen Sefati, an expert in autonomous driving at Mercedes-Benz and head of the project.

In fact, the stop sign example conceals a fundamental weakness of all neural networks that interpret traffic events in autonomous vehicles. They are similar in structure to the human brain, but they differ in a number of crucial points. For example, neural networks can only acquire their abilities all at once, typically in a single large training session.

Domain changes demand huge efforts

Even trivial changes can require massive efforts in the development of autopilots. Here's an example: In many autonomous test vehicles, cameras with a resolution of two megapixels were previously installed. If they are now replaced by better models with eight megapixels, in principle hardly anything changes. A tree still looks like a tree, only that it is represented by more pixels. Yet the AI still needs millions of snapshots from traffic again to recognize the objects at the higher resolution.

The same is true if a camera or radar sensor on the vehicle is positioned just slightly differently. After that, a complete retraining is once again necessary.

Experts call this a domain change: Instead of driving on the right, you drive on the left; instead of bright sunshine, a snowstorm is raging. Human drivers usually find it easy to adapt. They intuitively recognize what is changing and transfer their knowledge to the changed situation. Neural networks are not yet able to do this. A system that has been trained on fair-weather drives, for example, is confused when it rains because it no longer recognizes its environment due to the reflections. The same applies to unknown weather conditions, to the change from left- to right-hand traffic or to different traffic light shapes. And if completely new objects such as e-scooters appear in traffic, the autopilot must first be familiarized with them.

Aim of the project: learning only the “delta”

In all these cases, it has so far not been possible to teach the algorithm only the change, that is what in science is called the “delta.” In order to become familiar with the new domain, it again needs a complete dataset in which the modification occurs. It's like a student having to go through the entire dictionary every time they learn a new word.

This kind of learning gobbles up enormous resources. “Today, it takes 70,000 graphics processor hours to train an autopilot,” explains Tobias Kalb, a doctoral student involved in the AI Delta Learning project for Porsche Engineering. In practice, numerous graphics processing units (GPUs) are used in parallel to train neural networks, but the effort is still considerable. In addition, a neural network needs annotated images, that is images from real traffic events in which important elements are marked, such as other vehicles, lane markings, or crash barriers. If a human performs this work manually, it takes an hour or more to annotate a snapshot from city traffic. Every pedestrian, every single zebra crossing, every construction site cone must be marked in the image. This labeling, as it is known, can be partially automated, but it requires large computing capacities.

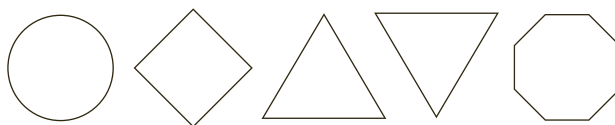
In addition, a neural network sometimes forgets what it has learned when it has to adapt to a new domain. “It lacks a real memory,” explains Kalb. He himself experienced this effect when using an AI module trained with US traffic scenes. It had seen many images of empty highways and vast horizons and could reliably identify the sky. When Kalb additionally trained the model with a German dataset, a problem arose. After the second run, the neural network had trouble identifying the sky in the American images. In the German images, it was it was often cloudy or buildings blocked the view.

Continual Learning

In Continual Learning, neural networks are gradually expanded to include new knowledge. In the process, they retain in memory what they have already learned. This means that they do not have to be trained again with the entire dataset each time.

First training run

- ① The neural network learns to distinguish shapes.



- ② The neural network can correctly classify colors.



- ③ The neural network distinguishes different labels and characters.



- ④ With these skills, it can now recognize a German stop sign.



Additional knowledge is learned

With the help of the previously learned knowledge about the shapes, colors, and fonts of the German signs, the neural network only has to learn the new characters after delta learning in order to reliably recognize a Japanese stop sign.



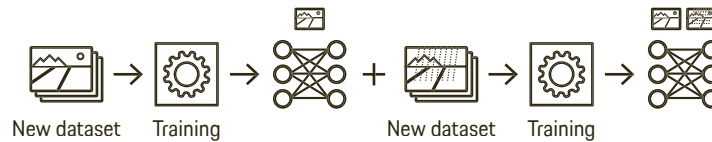
Five approaches to AI Delta Learning

❶ In **continual learning**, algorithms are developed that can be extended with new knowledge without loss of knowledge—without the need to retrain the entire dataset. Unlike traditional methods, not all data needs to be available at training time. Instead, additional data can be added to the training step by step at a later time. For example, a neural network can learn to recognize a Japanese stop sign without forgetting the German stop sign.

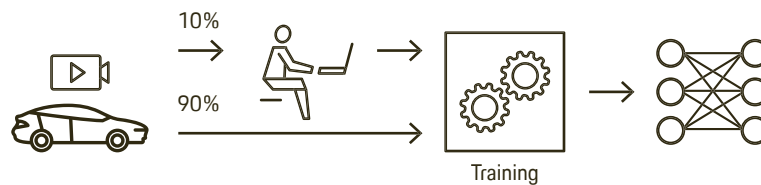
❷ In **semi-supervised learning**, labels are available for only a small portion of the data, which can be used to classify it into a category. Therefore, the algorithm trains with both unlabeled and labeled data. For example, a model trained with labeled data can be used to make predictions for some of the unlabeled data. These predictions can then be incorporated into the training data to train another model using this augmented dataset.

❸ **Unsupervised learning** is where an AI learns using data that has not previously been manually categorized. This allows data to be clustered, features to be extracted from it, or a new compressed representation of the input data to be learned without human assistance. In the AI Delta Learning project, unsupervised learning is used on to initialize neural networks and reduce the volume

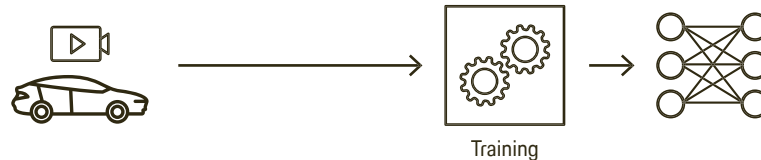
❶ Continual learning



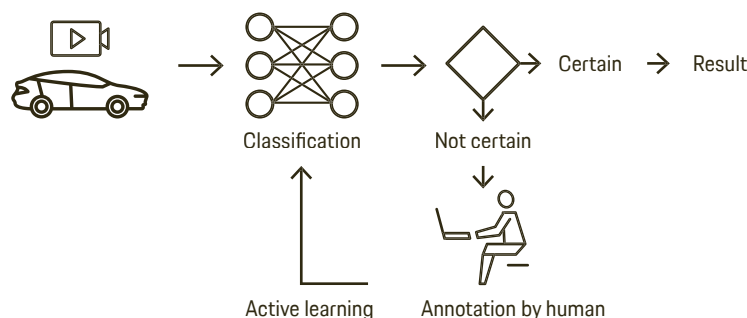
❷ Semi-supervised learning



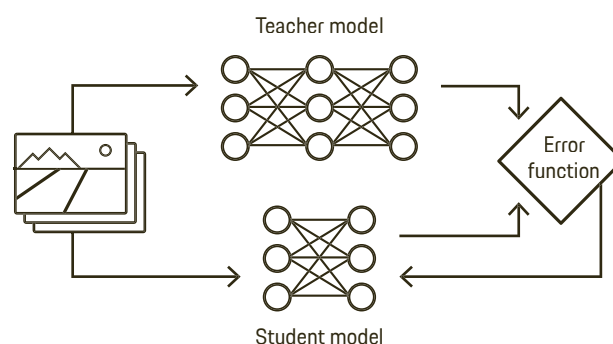
❸ Unsupervised learning



❹ Active learning



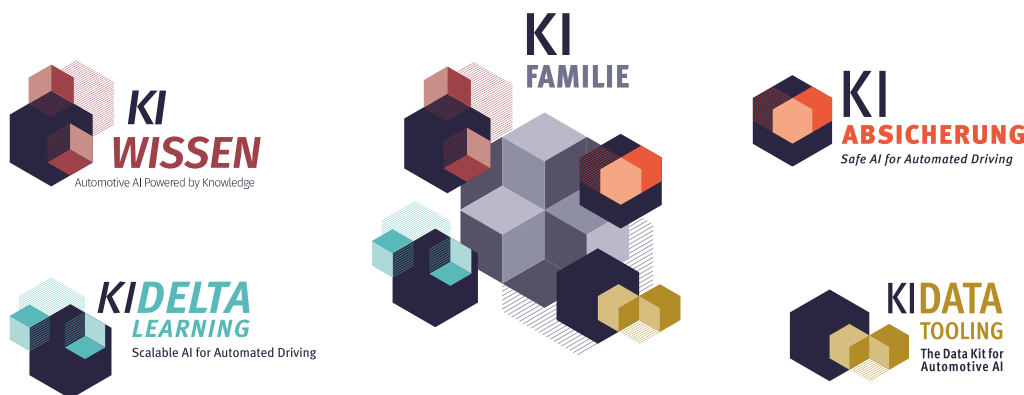
❺ Knowledge transfer



of annotated training data. It can also be used to adapt an already trained network to a new domain by trying to learn a unified representation of the data. For example, when making a domain switch from daytime to nighttime images, the features that the model learned for a car during the day should be equally applicable at night. So ideally they should be domain invariant.

❹ In **active learning**, algorithms self-select the training data for a neural network during training time—for example, those situations that have not occurred before. The selection is based, among other things, on degrees of uncertainty that estimate how certain a prediction of the neural network is. Through active learning, one can, for example, reduce the effort required for the manual annotation of video images because only the training data that is essential for learning later on has to be processed.

❺ **Knowledge transfer** (knowledge distillation) is the transfer of knowledge between neural networks—usually from a more complex model (teacher) to a smaller model (student). More complex models usually have a larger knowledge capacity and thus achieve higher prediction accuracies. Knowledge distillation compresses the knowledge contained in the complex network into a smaller network, with little loss of accuracy. Knowledge distillation is also used in Continual Learning to reduce knowledge loss.



"Until now, the model has been retrained with both datasets in such cases," Kalb explains. But this is time-consuming and reaches its limits at some point, for example when the datasets become too large to store. Kalb found a better solution through experimentation: "Sometimes very representative images are enough to refresh the knowledge." For example, instead of showing the model all American and German road scenes again, he selected a few dozen pictures with particularly typical highway distant views. That was enough to remind the algorithm what the sky looked like.

Two AIs train each other

It is precisely such optimization possibilities that AI Delta Learning aims to find. For a total of six application areas, the project partners are looking for methods to quickly and easily train the respective AI. This includes, among other things, a change in sensor technology or the adaptation to unknown weather conditions. Proven solutions are shared among the organizations involved in the project.

Another promising approach is for two perception AIs to train each other. First, a teacher model is built for this purpose: It receives training data in which a class of objects, for example signs, is marked. A second AI, the student model, also receives a dataset, but one in which other things are marked—trees, vehicles, roads. Then the teaching begins: The teacher system imparts its knowledge to the student as it learns new concepts. So it helps it recognize signs. After that, the student in turn becomes the teacher for the next system. This method, "knowledge distillation," could

The AI project family

AI Knowledge

Developing methods for incorporating knowledge into machine learning.

AI Delta Learning

Developing methods and tools to efficiently extend and transform existing AI modules of autonomous vehicles to meet the challenges of new domains or more complex scenarios.

AI Validation

Methods and measures for validating AI-based perceptual functions for automated driving.

AI Data Tooling

Processes, methods, and tools for efficient and systematic generation and refinement of training, testing, and validation data for AI.

save OEMs a lot of time in localizing their vehicles. If a model is to be introduced in a new market, all that needs to be done when training the autopilot is to use a different teacher model for the regional signs—everything else can stay the same.

Much of what the researchers are currently testing is still experimental. It is not yet possible to predict which method will ultimately allow a neural network to best adapt to new domains. "The solution will lie in a clever combination of several methods," expects expert Kalb. After a year of project work, those involved are optimistic. "We have made good progress," says project manager Sefati from Mercedes-Benz. He expects to be able to show the first methods for AI Delta Learning when the project ends at the end of 2022. That could yield huge benefits for the entire automotive industry. "There is high potential for savings while increasing quality if the training chain is highly automated," explains AI expert Schaper. He estimates that human labor input in the development of autonomous vehicles can be halved through AI delta learning. ◀

→ IN BRIEF

When changing the environment or the sensor technology, neural networks in vehicles today have to be trained again and again from scratch. The AI Delta Learning project aims to teach them only the differences after such a domain change, thus significantly reducing the effort involved.