AI Sensing
For a Safer Tomorrow

Artificial intelligence in advanced driver assistance systems

By: Dr. Florian Baumann, PhD
CTO, AI & Automotive, Dell EMC®

INCLUDING: “How Do Self-driving Cars Make Decisions?” from NVIDIA®
With the many recent reports about autonomous driving, it’s clear that Machine Learning (ML) has become a key practice in the automotive industry for the improvement of driver assistance and safety systems.

To the number and consequences of accidents involving vehicles, there have been developments in two key areas: physical safety measures in the car to protect the occupants and driver assistance systems. Currently the automotive sector uses ML in numerous applications, with a large proportion of algorithms focusing on driver assistance systems.

This article focuses on the important sub-area of camera-based driver assistance systems in combination with ML algorithms. Modern cameras, when combined with AI provide superhuman levels of perception and performance, all in a system that will never get drunk or drowsy. However, it can be difficult to apply camera sensor data in highly complex, variable environments.
The rise of Machine Learning

While it’s natural for the human brain to process limitless possible road traffic scenarios, weather, different lighting and traffic situations, and variations in different countries of the world, using camera sensor data to make the same determinations can be challenging. Developers must find methods that can handle these complexities as they develop driver assistance systems.

An important part of this is image processing methods for mathematical analysis of the image content. However, developers are always bumping into the limits of what’s possible. For example, no one so far has found a formula or hand coded algorithms to accurately identify people in scenario variations.

This is why ML has become such a crucial component in teaching the algorithm using training scenes, and then testing the system on unknown scenes to evaluate its generalization. These capabilities are not in some far-off future — ML algorithms are currently being used in real systems. The following sections provide a look back at the beginnings of machine learning in the auto industry and a look at the role it will play in the future of driver assistance systems.
Sensors, sensors everywhere

One of the first simple camera-based driver assistance systems includes the so-called night vision camera that provides drivers with infrared sensors enabling observation of the scenery in front of the vehicle in low-light scenarios. The inclusion of such a camera into the vehicle set the path for image processing algorithms that are now common in automobiles. As an example, lane keeping assistants utilize image processing algorithms to recognize road markings through visual recognition and warn drivers as they leave the lane. Further development of these systems enabled active intervention that, among other things, prevents the vehicle from leaving the lane. In addition, the algorithms merged camera data with radar data obtained from the transverse and longitudinal movement of the vehicle. Radar systems are also used to determine the distance to the vehicle in front.

Camera-based systems can be also be used for this purpose. Stereo camera systems enable the calculation of the distance to objects in the vehicle environment. Comparable to the human visual system, they take in the scene from two different perspectives (left and right eye). Based on the two pictures, triangulation is used to determine the distance of the objects to the camera — the depth of the objects. Over the last few decades numerous methods were presented to achieve such a depth map on the basis of a stereo image pair. The benchmarks of Middlebury and the Karlsruhe Institute of Technology for this purpose provide a good overview. As a result, such a depth map provides a robust input for more complex driver assistance systems and fully automated driving.
The learning process

At the turn of the millennium some scientists developed the first ML algorithms with sufficient complexity to display patterns and characteristics from larger data sets. These methods fall into the area of Supervised Learning and require, in addition to the algorithm itself, additional labeled training examples and a representation of characteristics.

Research and development of camera-based driver assistance systems requires training image data with positive examples, which contain the expected object, and negative examples that don’t contain it. Further, a feature representation is needed, also known as a mathematical representation, to display characteristics from images and extract the object as accurately and robust against changes in shape, weather, lighting and traffic situation.

Stuart Russell divides machine learning into two categories: Supervised Learning, also known as “learning with a teacher,” and Unsupervised Learning: “learning without a teacher.”

Supervised Learning: The algorithm learns with labeled training examples. It’s trying to learn all training examples, laws and patterns and extract the resulting knowledge so it can predict the label for new data. Typical methods are Support Vector Machines, Naive Bayes, neural networks, Random Forests and Boosting algorithms.

Unsupervised Learning: The algorithm learns with unlabeled training examples. The main goal is to find a structure that will describe training data in an optimal way and different subsets are categorized. Typical methods are the main component analysis and the Expectation-Maximization algorithm.

There are also other categories such as partially supervised learning, in which a mixture of training data with and without labels is used.

Another aspect of ML algorithms is generative vs. discriminatory models:

- Generative methods learn the combined probability distribution $P(x,y)$ and are based on a probability model estimated from training data.
- Discriminatory methods learn the conditional probability distribution $P(y|x)$ and build a classifier that can differentiate between two or more classes.
So much data, so little time

As the technology for autonomous driving has evolved over recent years, the requirements for driver assistance systems have risen sharply. An autonomously driven vehicle requires a variety of cameras, multiple laser scanners and radar sensors, ultrasonic sensors, Differential Global Positioning System (DGPS) and an inertial measurement unit (IMU) in order to gather enough information about the traffic situation. All sensor data is becoming higher resolution, and vehicles must be able to operate in all countries of this world under any conditions to drive autonomously and safely. This generates a huge amount of data, and traditional algorithms and methods from the field of machine learning are reaching their limits.

To put this data growth into perspective, according to SAE International, developing level 2 Advanced Driver Assistance Systems (ADAS) — lane centering and adaptive cruise control — requires sensor data gathered over the course of approximately 200,000 km, which equates to 4-10 PB of sensor data.

Level 5, or full autonomy, requires sensor data gathered over 200 million km, which equates to 2+ exabytes.

This means that over the course of a few years, organizations racing to be the first to full autonomy will face a huge influx of sensor data just to get to market with a fully autonomous model. Once publicly available, and with the arrival of 5G, these numbers will only continue to grow exponentially. To keep up with this growth, companies need a core storage system that provides simple scalability to handle so much data, data management capabilities to manage data into the multiple exabyte range, and the performance to keep up with the demands of new AI modalities.

So much data, so little time
Deep Learning

To process so much data, automotive industry scientists are turning to Deep Learning, a new form of learning based on artificial neural networks to expand the machine learning repertoire. Neural networks are like the human brain structure: Neurons are connected to each other, forming an interconnected network. Training the network works in a more hierarchical approach than traditional ML training. The algorithm learns important characteristics independently and extracts them gradually. Teaching with training examples is the same, but the definition of a suitable characteristic representation is omitted. This means that pre-existing, detailed knowledge of the data isn’t necessary, and the representation of the characteristics will be refined and tested over multiple iterations. The implemented, hierarchical process of Deep Learning learns an implicit representation of data and forms them over several layers. The exemplary training of the first layer serves the recognition of different objects on the basis of pixels. A second layer can be used for higher-quality structures such as lines before a third layer connects the lines to whole objects.

The number of layers is arbitrary. This method is ideally suited to large and complex datasets without manual creation of a characteristic representation for analysis. However, there are still challenges. Deep Learning is non-linear and very complex. It is not yet possible to understand the exact effect a specific input parameter will have on the output from the network.

Orienting new approaches, therefore, mostly occurs on known and proven network structures and parameterizations. Training the network is achieved by optimizing the output layer based on all parameters and links performed throughout the network. Since many parameters are to be optimized, the training process requires an immense amount of computing resources and thus the use of GPUs, which can process data in parallel.

Deep Learning has led to a real paradigm shift in the field of machine learning. Despite the technical and legal hurdles Deep Learning is entering into series development of driver assistance systems. Today’s success of Deep Learning can be attributed to the interaction of the following points:

- Availability of a wide range of sensor data that can support complex algorithm training
- Novel hardware, such as GPUs, which enables the learning of large networks
- Progress in research into new methods
- Energy efficient supercomputing at the edge for real-time inferencing

Ensuring a robust series production in the automobile requires a combination of proven machine learning methods that have already been used in practice. Algorithms and novel Deep Learning methods can be an important — and perhaps even a decisive — factor.
An array of deep neural networks power autonomous vehicle perception, helping cars make sense of their environment.

Self-driving cars see the world using sensors. But how do they make sense of all that data?

The key is perception, the industry’s term for the ability, while driving, to process and identify road data — from street signs to pedestrians to surrounding traffic. With the power of AI, driverless vehicles can recognize and react to their environment in real time, allowing them to safely navigate.

They accomplish this using an array of algorithms known as deep neural networks, or DNNs.

Rather than requiring a manually written set of rules for the car to follow, such as “stop if you see red,” DNNs enable vehicles to learn how to navigate the world on their own using sensor data.

These mathematical models are inspired by the human brain — they learn by experience. If a DNN is shown multiple images of stop signs in varying conditions, it can learn to identify stop signs on its own.

But just one algorithm can’t do the job on its own. An entire set of DNNs, each dedicated to a specific task, is necessary for safe autonomous driving.
These networks are diverse, covering everything from reading signs to identifying intersections to detecting driving paths. They’re also redundant, with overlapping capabilities to minimize the chances of a failure.

There’s no set number of DNNs required for autonomous driving. And new capabilities arise frequently, so the list is constantly growing and changing.

To actually drive the car, the signals generated by the individual DNNs must be processed in real time. This requires a centralized, high-performance compute platform, such as NVIDIA DRIVE AGX.

Following are some of the core DNNs that NVIDIA uses for autonomous vehicle perception.

**Pathfinders**

DNNs that help the car determine where it can drive and safely plan the path ahead:

- **OpenRoadNet** identifies all of the drivable space around the vehicle, regardless of whether it’s in the car’s lane or in neighboring lanes.
- **PathNet** highlights the driveable path ahead of the vehicle, even if there are no lane markers.
- **LaneNet** detects lane lines and other markers that define the car’s path.
- **MapNet** also identifies lanes as well as landmarks that can be used to create and update high-definition maps.

**Object Detection and Classification**

DNNs that detect potential obstacles, as well as traffic lights and signs:

- **DriveNet** perceives other cars on the road, pedestrians, traffic lights and signs, but doesn’t read the color of the light or type of sign.
- **LightNet** classifies the state of a traffic light — red, yellow or green.
- **SignNet** discerns the type of sign — stop, yield, one way, etc.
- **WaitNet** detects conditions where the vehicle must stop and wait, such as intersections.

The list goes on

DNNs that can detect the status of the parts of the vehicle and cockpit, as well as facilitate maneuvers like parking:

- **ClearSightNet** monitors how well the vehicle’s cameras can see, detecting conditions that limit sight such as rain, fog and direct sunlight.
- **ParkNet** identifies spots available for parking.

These networks are just a sample of the DNNs that make up the redundant and diverse DRIVE Software perception layer.

To learn more about how NVIDIA approaches autonomous driving software, check out the DRIVE Labs video series.

Stay up to date with NVIDIA’s latest innovations at the NVIDIA Automotive Blog.
Dell EMC Isilon with NVIDIA DGX
AI infrastructure that scales

This integrated solution pairs Dell EMC’s scale-out all-flash F800 storage with the NVIDIA DGX-1, to address the challenges associated with scaling AI workloads in the data center, from design to deployment to support.

Dell EMC Isilon: Fuels AI with data

Dell EMC Isilon F800 All-Flash Scale-out NAS delivers the analytics performance and extreme concurrency at petabyte scale to consistently feed the most data hungry AI algorithms. Additionally, Isilon’s enterprise features for data management, data security, data compliance, and data protection help your AI solutions conform to regulatory and enterprise security policy requirements.

NVIDIA DGX-1: The essential tool for AI research and development

NVIDIA DGX-1™ is the world’s first integrated AI system purpose-built for enterprise. Using NVIDIA NVLink technology to Integrate eight of the world’s fastest data center accelerators – the NVIDIA V100 Tensor Core GPU – the DGX-1 delivers over 1 petaFLOPS of AI performance. NVIDIA DGX-1 is powered by the DGX software stack with optimized deep learning, machine learning and HPC software from NGC, delivering a plug-in, power-up deployment experience for the fastest path to AI-powered insights, effortless productivity for data science teams, and revolutionary performance that scales with NVIDIA DGX POD™.

Want to learn more?
Checkout the Dell EMC Isilon with NVIDIA DGX-1 Reference Architecture
Learn more about Dell EMC Isilon
Learn more about NVIDIA DGX-1