# AUTONOMOUS VEHICLES

## Solution Brief

Revision 0.1

habana

# AI in AV

**Autonomous driving is on the verge of transforming the transportation industry.**

## LEVELS OF DRIVING AUTOMATION

| | LEVEL 0 | LEVEL 1 | LEVEL 2 | LEVEL 3 | LEVEL 4 | LEVEL 5 |
|---|---|---|---|---|---|---|
| What does the human in the driver's seat have to do? | **You are driving** whenever these driver support features are engaged – even if your feet are off the pedals and you are not steering | | | **You are not driving** when these automated driving features are engaged – even if you are seated in the "driver's seat" | | |
| | **You must constantly supervise these support features: you must steer, break or accelerate as needed to maintain safety** | | | When the feature requests, you must drive | These automated driving features will not require you to take over driving | |
| | These are driver support features | | | These are automated driving features | | |
| What do these features do? | These features are limited to providing warnings and momentary assistance | These features provide steering OR brake/acceleration support to the driver | These features provide steering AND brake/acceleration support to the driver | These features can drive the vehicle under limited conditions and will not operate unless all required conditions are met | | This feature can drive the vehicle under all conditions |
| Example features | Automatic emergency braking<br>Blind spot warning<br>Lane departure warning | Lane centering<br>OR<br>Adaptive cruise control | Lane centering<br>AND<br>Adaptive cruise control at the same time | Traffic jam chauffeur | Local driverless taxi<br>Pedals/steering wheel may or may not be installed | Same as level 4, but feature can drive everywhere in all conditions |

*Figure 1: SAE's six levels of driving automation*

Autonomous driving is on the verge of transforming the transportation industry. The global autonomous vehicle (AV) market was valued at $820.29 billion in 2021 and is projected to reach $1,475.47 billion in 2026 at a compound annual growth rate (CAGR) of 12.2% [1]. This growth is mainly driven by the rise in the development of smart cities aiming to reduce traffic congestion, increase the efficiency and safety of the transportation system, and improve the quality and productivity of the time spent in cars.

The U.S. Department of Transportation uses the six levels of driving automation proposed by the Society of Automotive Engineers (SAE). These levels range from no driving automation (level 0) to full driving automation (level 5) and have become a standard for self-driving vehicles [2].

Artificial Intelligence has taken the automotive industry by storm to drive the development of level-4 and level-5 self-driving cars. AI and specifically deep learning solutions are widely used in all components of an autonomous driving system including perception, localization and mapping, decision making, and execution.

However, the implementation of AI in the automotive industry is not solely limited to self-driving capabilities. Many AI solutions are built around advanced driver-assistance systems (ADAS), the combination of edge computing and IoT devices, voice recognition, and more. And, of course, the common fields like car maintenance, supply chain, and marketing are also involved.

Habana offers hardware and software solutions that shorten the time of AI training cycles . This helps accelerate the pace of innovation and time to market, which is crucial for competitive industries like autonomous driving. We are also proud that our solutions reduce the development and validation costs letting the researchers explore more within their budget.

II.

# Applications of AI in AV

**AI-based algorithms are present in all levels of automation.**

Voice-enabled virtual assistants help execute commands related to driving or the entertainment system in the car. Computer vision systems help perform lane departure warning, distance warning, and emergency brake assistance in driver assistance (level 1) and partial automation (level 2) systems. It also helps autonomous driving under certain conditions in conditional automation (level 3) systems. And, of course, the fully automated driving on levels 4 and 5, where the vehicle independently controls complete journeys on the highway as well as in city traffic would not be possible without the recent advancements in AI.

Autonomous vehicles utilize the data provided by their sensors to make short- and long-term decisions. A typical workflow of an AV includes sensing, perception, localization and mapping, prediction, planning and control, and driver monitoring . The recent advancements in deep learning have facilitated developments in automatic perception, prediction, and planning algorithms. Large semantic segmentation tasks running over multiple cameras and point clouds help AVs detect and recognize road lanes, other vehicles on the road, pedestrians, traffic signs, lights, etc. Deep learning is also used for tracking those objects and predicting their behavior, which helps derive control decisions. Deep reinforcement learning algorithms are also widely used in planning and controlling components of an AV.

## Challenges

Streets are wild environments from the computer vision point of view. The variety of objects that can appear on the road is limitless ranging from expected objects like cars, bikes, and people to unexpected ones like animals or even pieces of furniture dropped in the middle of the road. The core perception algorithms need to evolve to detect these objects in geographically diverse locations with various lighting and weather conditions. This can only happen by continuously adding interesting and informative samples to the train data. This means lots of training iterations on enormous datasets.

The main challenge in autonomous driving is hyperscaling. It needs thousands of compute nodes for development and validation. AV fleets collect massive and rapidly growing datasets; Terabytes of data flow daily into the system from each vehicle. Along with the data scaling the compute infrastructure also needs to be dynamically scalable. This raises the need for instances that can handle large distributed machine learning training jobs.

III.

# Why is Gaudi a good fit for AV use cases?

As the AV Challenges section mentioned, the autonomous driving industry needs scalable compute infrastructures that can handle many large distributed ML training jobs. There is an increasing need for frequent retraining and updating of models to reach generalizability across different location conditions. Training deep 2D/3D semantic segmentation and object tracking models used for perception and localization require a large amount of processing that can be parallelized and thus accelerated. They can benefit specifically from accelerators that can handle data parallelism when the training dataset is huge or distributed and model parallelism when the models are large.

The two primary considerations that come into play in employing AI processing - whether for computer vision or NLP applications - are time-to-train models to the desired level of accuracy and cost-to-train. Habana's Gaudi Training Processors are expressly designed—in both hardware and software—to deliver high-efficiency cost- and time-to-train, making AI training more accessible to more organizations and for more applications. This helps to reduce development and validation costs and to enable rapid innovation and faster time to market.

Training with Gaudi clusters is available both in the cloud with AWS EC2 DL1 instances consisting of 8 Gaudis and on-premises with the Supermicro X12 Gaudi Training Server, also featuring 8 Gaudis.

The ideal equation for end users is to achieve desired AI price-performance, meaning that the cost and time to train each image or language sequence meets cost and time investment criteria. In other words, enabling more training at a low cost is the objective for data scientists and IT infrastructure management.

First-generation Gaudi, in fact, has proven delivery of up to 40% better price-performance than comparable GPU-based solutions—for both the EC2 DL2 instance and on-premise systems. And, there are customer cases that have proven even greater cost savings, which will be shared in part 4 of this brief.

In addition, Gaudi2, which launched in May, offers substantial performance advances that enable significantly faster training of models, while preserving cost-efficiency. Gaudi2 systems will be available in 2H 2022 for on-premises implementation.

Following are customer cases that prove Gaudi's efficiency.

# News and customer testimonials

*"As a world leader in automotive and driving assistance systems, training cutting edge Deep Learning models for mission-critical to Mobileye business and vision. As training such models is time consuming and costly, multiple teams across Mobileye have chosen to use Gaudi-accelerated training machines, either on Amazon EC2 DL1 instances or on-prem; **Those teams constantly see significant cost-savings relative to existing GPU-based instances across model types, enabling them to achieve much better Time-To-Market for existing models or training much larger and complex models aimed at exploiting the advantages of the Gaudi architecture.** We're excited to see Gaudi2's leap in performance, as our industry depends on the ability to push the boundaries with large-scale high performance deep learning training accelerators."*

**Gaby Hayon**
Executive Vice President of R&D
Mobileye

*"On our own models the increase in price performance met and even exceeded the published 40% mark."*

**Chaim Rand**
Machine Learning Algorithm Developer
Mobileye

*"We are consistently seeing cost-savings compared to existing GPU-based instances across model types, enabling us to achieve much better Time-to-Market for existing models or training much larger and complex models."*

**David Peer**
DevOps Tech Lead & Specialist
Mobileye

## Other Articles

https://towardsdatascience.com/training-on-aws-with-habana-gaudi-3126e183048

https://medium.com/intel-tech/mobileye-journey-towards-scaling-amazon-eks-to-thousands-of-nodes-leveraging-intel-xeon-scalable-a3ffc71cdbf8

V.

# References


[1] The Business Research Company, Autonomous Cars Global Market Report 2022, available at https://www.thebusinessresearchcompany.com/report/autonomous-cars-global-market-report

[2] SAE J3016: Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems