

# Issues and activities on Safe AI for Autonomous Driving

**Naoki Mitsumoto<sup>1</sup>**

naoki.mitsumoto.j6a@jp.denso.com

<sup>1</sup>DENSO CORPORATION, AI R&I Div.

Keywords: Safe AI, Autonomous Driving, Quality assurance

## ABSTRACT

Deep Neural Network has been becoming key technology for realizing autonomous drive. Such AI is expected to be utilized in open real world, safety and quality of AI is becoming very important issue in automotive industry. However they are not issues that can be handled by a single company, cooperative efforts among industries and academies are required.

## 1 INTRODUCTION

We have been currently addressing on R&D on AI for the core of an autonomous drive, especially has been focusing on perception and planning for tasks such as object detection, scene recognition, and path planning (Figure 1). There are five keys to be worked for making AI practical for autonomous drive application (Figure 2).

In particular, quality assurance of AI is one of the important issues for automotive industry from the view point of safety, and has many challenging tasks to be solved both technical and theoretical.

## 2 ISSUES TO REALIZE SAFE-AI

In the field of image recognition, Deep Neural Network has already exceeded human ability in specific tasks,

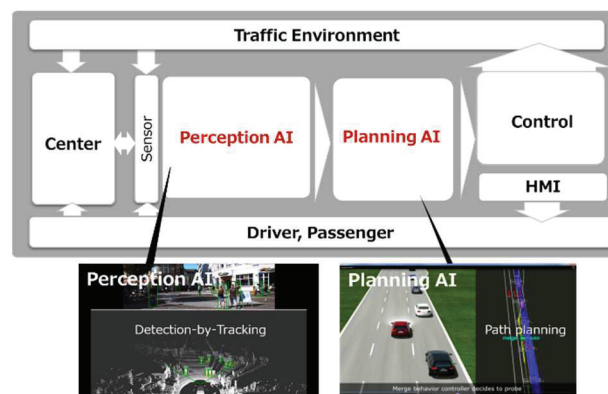


Fig. 1 Architecture for Autonomous Drive

such as image recognition, object detection or semantic segmentation, but on the other hand, way to assure quality of Deep Neural Network is not trivial, and there are many concerns for managing and utilizing DNN safely in the Advanced Driver Assistant System or Autonomous Driving System. The concerns are;

- DNN is known to exhibit unexpected behaviors when given previously unseen data.
- We cannot prepare data that covers every possible

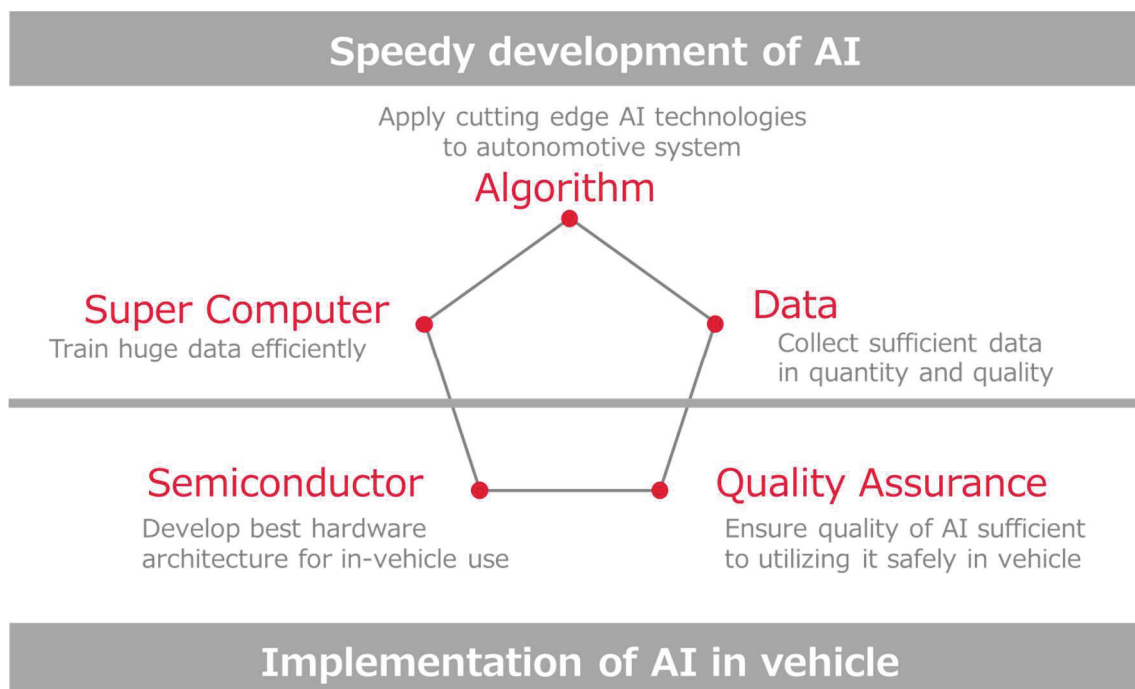


Fig. 2 Five keys for making AI practical use for autonomous drive

events of the world we live in.

- DNN can be fooled, possible to intentionally induce errors with very small noise, which is known as adversarial example.

In order to deal with these problems, it is necessary to challenge four important issues; (1) policy, (2) quality assurance, (3) social acceptance and (4) fundamental technologies (Figure 3). Safe AI starts from building a policy. Way of thinking and mechanism of quality assurance can be constructed based on the policy; process, assessment and audit can be materialized. Risks and benefits of AI need to be accepted by society, which requires standardization and consensus. And it is important to develop the fundamental technologies that supports them.

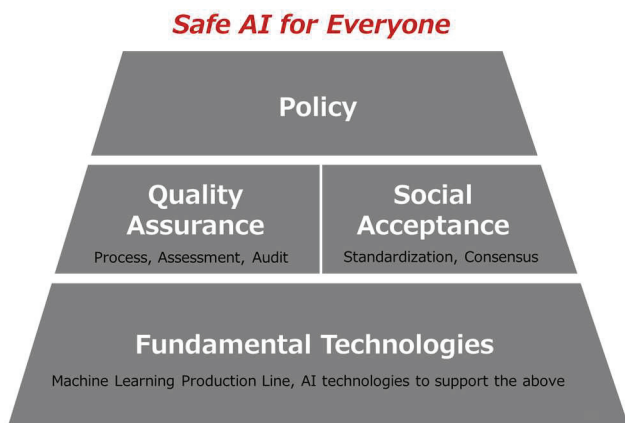


Fig. 3 Four important issues for realizing safe AI

### 3 TECHNOLOGICAL CONCEPT FOR SAFE-AI

How can we ensure the quality and safety of AI from the technological view point? We believe that it is extremely important to firmly manage and improve the quality of each step of machine learning process. To that end, we think it is important to realize a machine learning production line that efficiently manages large-scale data, and to build a group of advance AI technologies that support that production line. Figure 4 shows our concept.

In order to build a high-quality data set, it is necessary to have a technique for identifying important data to be trained and a technique for analyzing the coverage of training and evaluation data. To build a high-quality model, we need a technology to identify the weaknesses of DNN model, a technique to mitigate the weaknesses, a technique to explain/understand the behavior of DNN, and a technique to deal with the uncertainty of DNN. We are currently conducting research and development to realize this concept.

### 4 CONCLUSIONS

Quality and safety of Deep Neural Network are very important issue for autonomous drive, and we're challenging this from the perspectives of quality assurance, social acceptance, and technological development. However they are not issues that can be handled by a single company, cooperative efforts among industries and academies are required.

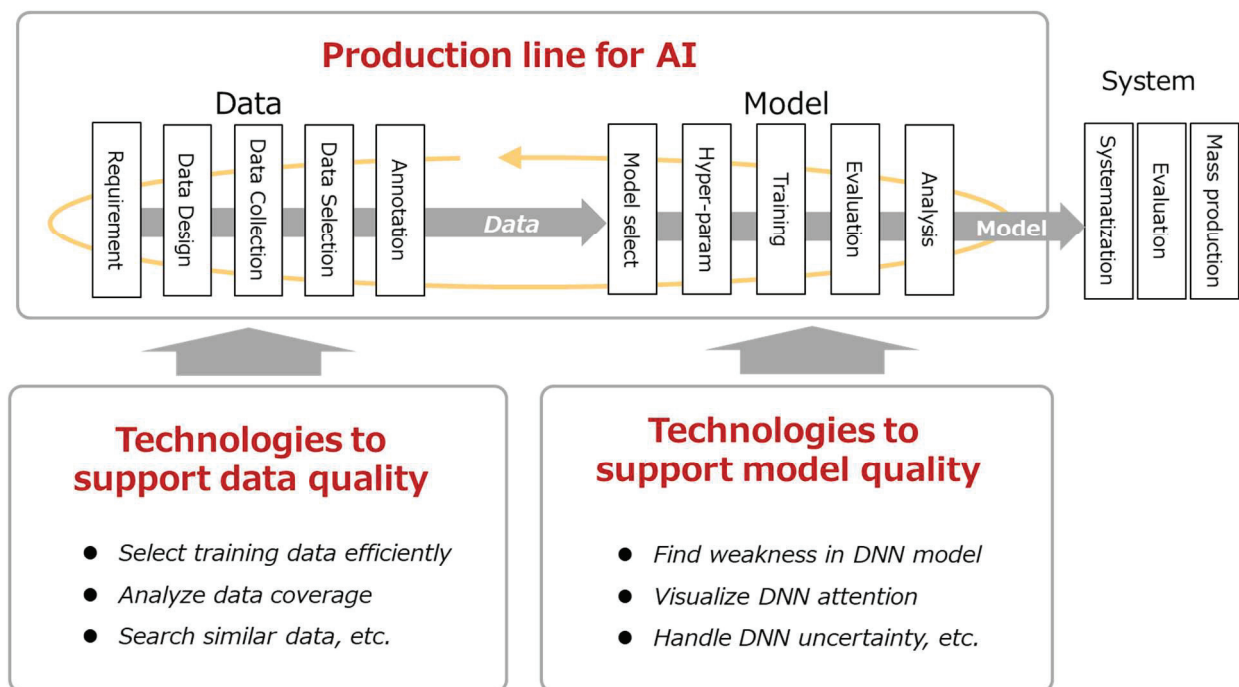


Fig. 4 Technological concept for realizing safe-AI