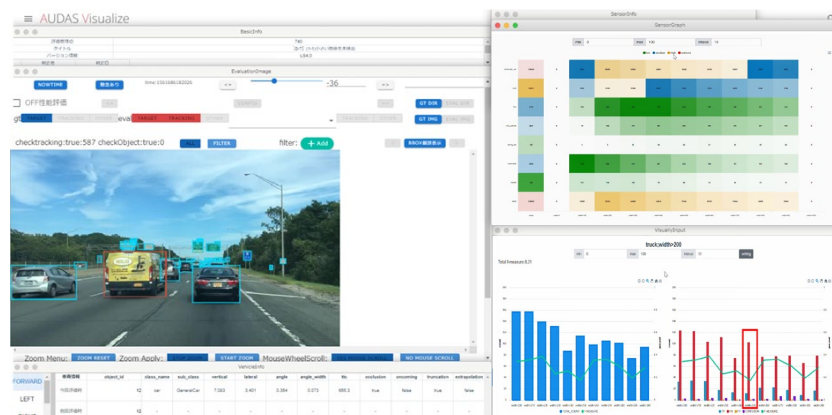


Improving the efficiency of deep neural network-based object recognition evaluation with a visualization tool!



Woven by Toyota



Mr. Kashiba



Ms. Suzuki

◆What type of work do you do?

I was responsible for evaluating the performance of an object recognition algorithm that utilized the first object recognition DNN*¹ introduced by our company. It is used in the Advanced Drive system of Toyota/Lexus Teammate, which is featured in the LS and MIRAI models launched in 2021. The algorithm development team also performed evaluations, but we were responsible for providing a more objective assessment of this new technology, incorporating the customer's perspective as well. This work helped us ensure the high level of safety required for the product.

*¹ DNN: deep neural network

◆Tell us about the challenges you face in your work

It was necessary to comprehensively assess the various situations that customers may encounter while driving to ensure objectivity during the evaluation process. This required us to systematically organize and understand the environments in which customers use their vehicles, and how they are operated. Simply combining all of the scenarios that customers might encounter would have resulted in an enormous number of evaluation items. Therefore, we needed to identify combinations that would effectively reveal weaknesses in the algorithm and enable efficient evaluation within the limited timeframe available before the product was released to the market.

◆Could you explain what you mean by “customer-perspective object recognition”?

The object recognition algorithm is required to detect preceding vehicles on all roads and in all operating environments where Advanced Drive is active. Although vehicles with Advanced Drive are currently sold in Japan and North America, the road environments in these regions are very different. In addition to differences in road condition, there is also considerable variation in weather conditions depending on the season and time of day, as well as in vehicle types, traffic signals and road signs, and even driving practices and traffic regulations. These factors make it essential to identify evaluation criteria from the perspective of customers who use Advanced Drive. The research and data preparation alone required a significant amount of time. Furthermore, to systematically organize the evaluation process and demonstrate comprehensive coverage of all relevant evaluation perspectives, we needed to study both internal and industry standards, which added considerably to the effort.

◆What are the requirements for solving issues?

Given the limited time allocated for evaluation, we needed to develop processes and tools that would allow us to accelerate the evaluation cycle. Specifically, we identified the following requirements: establish tools to support the process from evaluation data preparation to analysis, enhanced tool-chain processing performance, and more efficient sharing of project status and technical challenges.

●Establishing Tools for Evaluation Data and Analysis

Performance evaluation consists of a series of steps, including selecting evaluation data from test vehicle driving data, converting it into a format that can be processed by the recognition software, running the software, and analyzing its performance. At the start of the evaluation program, many of these tasks were performed manually. Because of the heavy manual aspect of the work, variations in output quality among operators and rework caused by human error became significant issues. As a result, there was a strong need to establish tools that would enable the process to be carried out in a more systematic and standardized manner.

●Enhancing Toolchain Processing Performance

DNN models are refined repeatedly throughout the development process, so evaluations must be completed within a short timeframe. In addition, as development progressed, the volume of evaluation data continued to grow. This created a need for greater processing capacity to handle more data within the same timeframe, as well as an integrated toolchain capable of supporting repeated evaluation cycles.

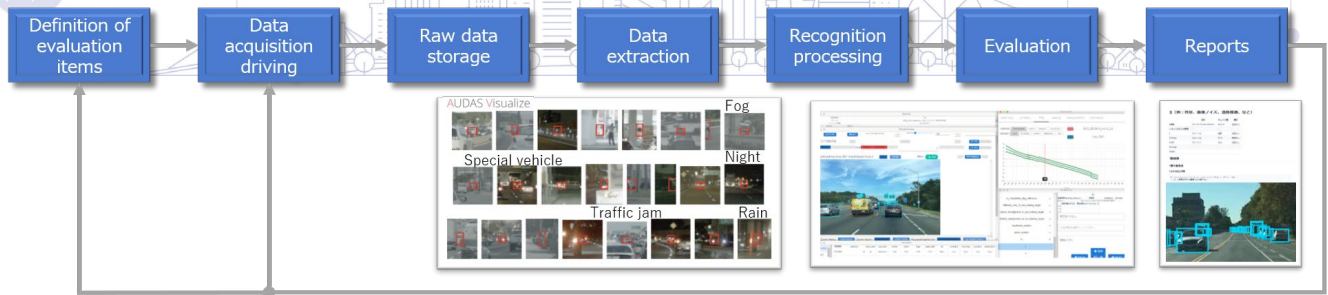
●Improving Progress and Issue Sharing

When sharing information both inside and outside our recognition development team, we needed a way to visualize progress across all evaluation items, efficiently share issues, and exchange relevant data.

◆Tell us why you chose TTDC

It all began when they helped us build a test vehicle for collecting evaluation data. While working together on the development of components for the test vehicle, we were also discussing evaluation methods for the DNN. During those discussions, TTDC introduced us to its initiatives for evaluating image recognition software, as well as its toolchain. They not only possessed the necessary knowledge, but had also actually operated the process themselves and experienced its challenges firsthand. In addition, their proposed improvements were highly concrete, which gave us confidence that they had solutions to the issues we were facing. As a result, we asked for their support.

Enabling more efficient evaluations and shorter development cycles with logic that quantifies rainfall and snowfall from image data!



Constructed Recognition Performance Evaluation Process and Toolchain

While many companies specialize in AI development and evaluation, only a limited number have experience in developing and evaluating AI for automotive applications. It is even more difficult to find companies with experience developing and evaluating AI intended for production vehicles, rather than AI at the research stage. We decided to seek TTDC's support because it has hands-on experience with image recognition software development processes, as well as extensive experience in a wide range of automotive development and evaluation activities beyond AI evaluation alone.

◆ What were the results of engaging TTDC?

First and foremost, I believe the project helped improve the recognition performance of the system delivered to customers, thereby contributing to a higher level of safety. A major factor was the improvement in evaluation coverage achieved through faster evaluation cycles and the ability to assess a larger volume of data, enabling us to more comprehensively evaluate conditions representative of real-world customer use. In addition, I believe the evaluation process benefited in several tangible ways, including the systematic organization of evaluation perspectives, a reduction in issue outflow to downstream processes, and a lower reliance on manual effort in evaluation activities.

● Systematizing evaluation perspectives

To improve the validity and explainability of the evaluation items, it was necessary to establish a structured evaluation framework. Evaluation items are derived from a variety of perspectives, including customer-focused analysis, known weaknesses of conventional image recognition systems, the design philosophy of the recognition software, and international standards such as SOTIF^{*2}. By systematically organizing these evaluation perspectives using the scenario model of the PEGASUS Project as a reference and establishing a mechanism that links them to tools for managing the corresponding datasets and evaluation results, we were able to quickly retrieve and present the evaluation results we needed. In snowfall scenarios, it is often difficult to judge snowfall intensity based solely on visual inspection of images. To address this challenge, we worked with TTDC to develop an image similarity analysis tool that quantitatively compares images against reference images corresponding to known snowfall levels. This made it possible to efficiently identify and prepare evaluation data with the desired snowfall conditions.

AUDAS Visualize



Evaluation Data Extraction Tool Based on
Snowfall Intensity (Conceptual Images)

● Reducing issue outflow to downstream processes

We believe our evaluation activities helped reduce rework during vehicle-level evaluation of Advanced Drive. The primary purpose of vehicle-level testing is to evaluate vehicle behavior and the user interface. Because these tests require significant time and manpower to prepare and execute, it was important to address issues originating from the recognition software as early as possible, ideally before vehicle-level testing began. By providing early feedback to stakeholders on weaknesses identified in the DNN-based image recognition system, we were able to facilitate a reassessment of performance allocation. This helped ensure that vehicle-level evaluation was not initiated until the software had reached an appropriate level of maturity.

● Reducing manual effort in evaluation activities

Advanced Drive continues to receive updates even after the vehicle has been delivered to the customer. Whenever the recognition software is updated, previously collected evaluation data must be reused to verify that no performance regressions have been introduced. Establishing the toolchain enabled us to perform these verification activities with minimal effort.

^{*2} SOTIF: An ISO standard for automotive safety.

◆ What are your expectations for TTDC in the future?

To help solve challenges across the Toyota group, TTDC is engaged in a wide range of technology development activities, leveraging emerging technologies and incorporating knowledge and information from both inside and outside the group. When introducing new technologies, TTDC provides concrete examples that include the challenges and lessons learned during implementation. As a result, it is easier for us to visualize how those technologies can be applied to the projects and responsibilities we manage. We look forward to the continued sharing of case studies and successful practices from both within and outside the Group.



Back row (from left): Mr. Ichikawa (TTDC),
Mr. Sakaguchi (TTDC), and Mr. Yamaguchi (TTDC)
Front row (from left): Mr. Kashiba and Ms. Suzuki

All photographs and figures were provided courtesy of Woven by Toyota, Inc.