

# AI Development Technologies for Road Management System Based on High-Reliability AI Video Recognition

Daisuke Yamaoka  
Masaki Miura

Shin Fujita

Hiroyuki Mizuno

Kenichi Ota

This article introduces technologies for achieving a high-reliability AI video recognition system using deep learning. In recent years, the use of outdoor video cameras for sensing various types of events has been gaining momentum. Fujitsu has been applying image recognition technology based on AI and deep learning in particular to the public infrastructure sector including road management in the form of vehicle and event detection. These efforts have led to the development of high-reliability AI video recognition systems that must provide high-accuracy and stable AI detection under diverse conditions expected of a natural environment. We here describe key AI development technologies for meeting this requirement.

## 1. Introduction

Recent years have seen a growing trend in the use of outdoor video cameras for sensing various types of events. Video contains diverse types of information, and the ability to extract various types of data from a single stream of video content is a major advantage. The use of video for sensing purposes is a market that is expected to grow significantly from here on, so Fujitsu is pursuing technology development by making use of its strengths in AI technology and enhancing it even further.

Fujitsu has come to apply image recognition technology based on AI and deep learning in particular to the public infrastructure sector including road management and has developed high-reliability AI video recognition systems as part of these efforts. One such system automatically detects abnormal events from the video of road-monitoring closed-circuit television (CCTV) cameras and notifies road management operators of those events so that a rapid response can be mounted [1].

A high-reliability system of this kind requires high-accuracy and stable video recognition under diverse conditions in a natural environment. Deep learning makes it possible to achieve a certain level of image recognition from a large number of images through trial and error without having to depend on any special

know-how. At the same time, in a field that requires operation under various types of natural environments, it is not easy to achieve stable and high recognition performance. As a reflection of this situation, we consider that AI video recognition development has the growth stages shown in **Figure 1**.

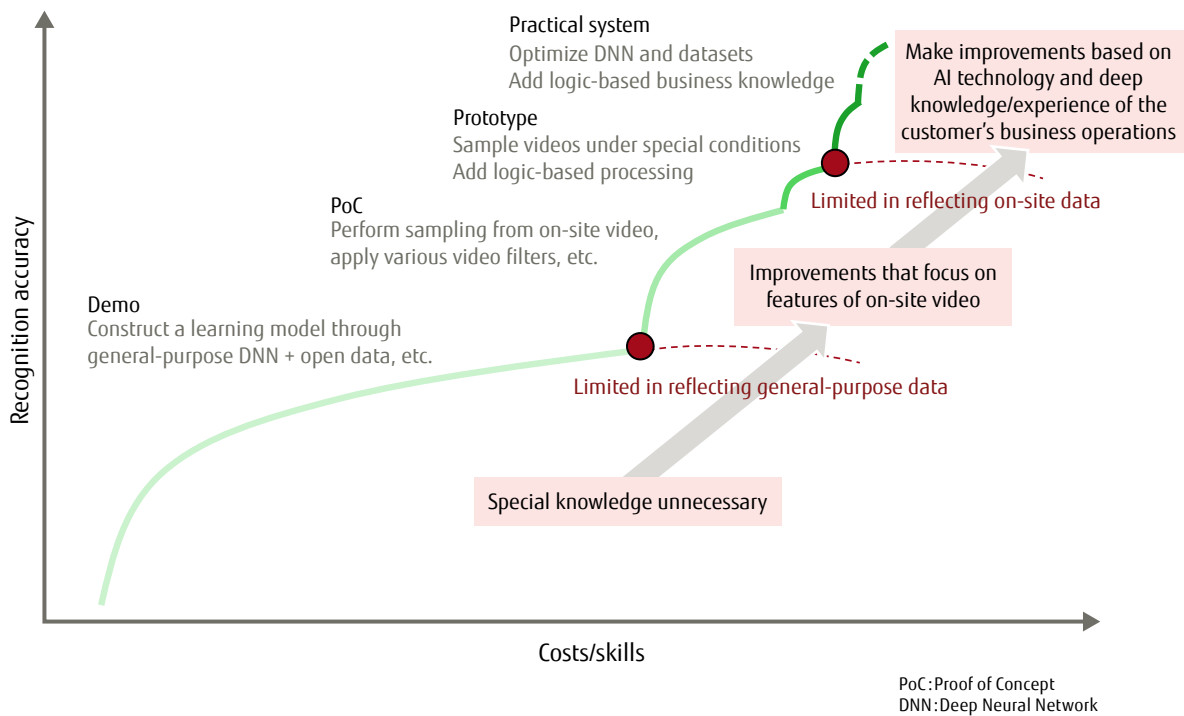
In this article, we describe AI development technologies for achieving a practical high-reliability system that can satisfy severe requirements in the field.

## 2. Issues in development of AI video recognition system using deep learning

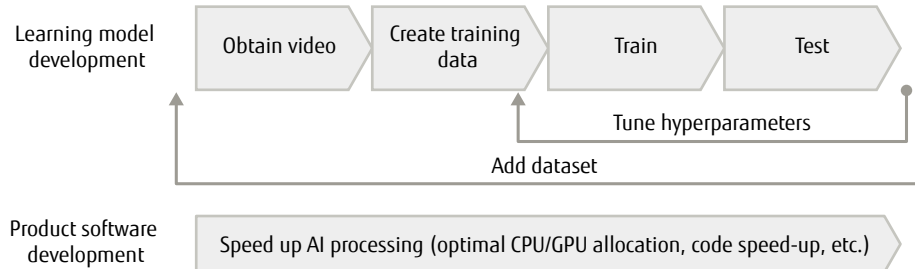
This section describes the issues involved in developing a high-reliability AI video recognition system using deep learning. In this article, “high-reliability” refers to stable and high-accuracy video recognition under diverse conditions that can be envisioned for a natural environment.

In general, AI development using deep learning follows the flow shown in **Figure 2**.

To begin with, a problem common to all types of image recognition technologies is how to go about preparing training data. The public infrastructure sector, in particular, requires video recognition that can operate under a variety of conditions in a natural environment, so variation in coverage is the first issue to contend with.



**Figure 1**  
Growth stages in AI video recognition development.



**Figure 2**  
AI development flow using deep learning.

The next issue concerns the accuracy of AI video recognition, that is, to what extent should accuracy be raised? In general, AI development repeats a sequence of development steps, namely, obtain video, create training data, train, and test. However, maximizing the repetition of these steps within a limited development period requires that the time of one cycle be shortened.

In addition, the computational complexity of AI processing is high, but increasing the scale of servers presents another problem in that such a system may not be feasible from the viewpoints of cost and installation space. Consequently, as a third issue, it is

imperative that AI processing be accelerated to minimize the number of servers.

The following section describes development technologies needed to resolve the above three issues.

### 3. What are AI development technologies for achieving a high-reliability system?

This section describes development technologies that are necessary for achieving a high-reliability AI video recognition system using deep learning.

### 3.1 Preparation of video dataset reflecting diverse environmental conditions

This subsection describes preparation of a video dataset as training data and the use of computer graphics (CG) for supplementing and increasing that data.

#### 1) Preparation of large-scale video dataset

Fujitsu first began applying AI to video recognition systems in the field of road management in 2016. However, the company has been researching AI itself for over 30 years and has come to possess a massive number of datasets of tagged road video from regions around the world. Furthermore, as a result of delivering AI video recognition systems for many customers in recent years, Fujitsu has added a huge amount of new video from sites within Japan and has been using this video as a dataset for development purposes. AI video recognition is deployed in a variety of natural environments, so video from the field is essential to development work. With this dataset for development purposes, we can cover a multitude of cases and quickly develop an AI video recognition system that can maintain a certain level of quality.

Acquiring video from the field must be done in a systematic manner by considering beforehand environmental conditions and patterns that can be expected to occur. We have established a technique for doing so

and have incorporated it into our development process. In this way, we have been developing the system with the goal of preventing any gaps in video coverage from the initial stage of development.

#### 2) Use of CG and AI-oriented quality improvement technology

Rarely occurring events may not exist in existing datasets or even in newly acquired video from the field. We are dealing with this problem by setting up a CG video generation environment and generating any type of video content based on envisioned scenarios.

However, caution must be taken when using CG in that the required CG quality differs greatly according to the event that needs to be recognized. It is generally thought that required CG quality depends on the size of the recognition target, but its dependency on the complexity of environmental conditions is actually more dominant.

Given a simple operation test, CG quality is not an issue, but when using CG as final training and test data, a level of quality practically the same as on-site video is necessary as shown in **Figure 3**.

In addition, the purpose of the system is not only detection of vehicles but also recognition of other characteristics that might extend as far as the car model. These additional requirements will require realistic 3D models and rendering down to fine details. For



**Figure 3**  
Simulation video by CG.

example, which of the CG renderings in **Figure 4** to select will depend on the current objective. The control of optimal CG quality is being organized as know-how and used in development work.

Another problem is that there are many instances in which inputting on-site video in the case of a CG-based learning model does not improve recognition performance. If the reasons for this cannot be accurately analyzed, CG correction, training, and testing will have to be repeated on a trial-and-error basis. To avoid this and achieve efficient use of CG, we are developing and using CG analysis technology in collaboration with Fujitsu Laboratories. This technology automatically



**Figure 4**  
Control of CG quality.

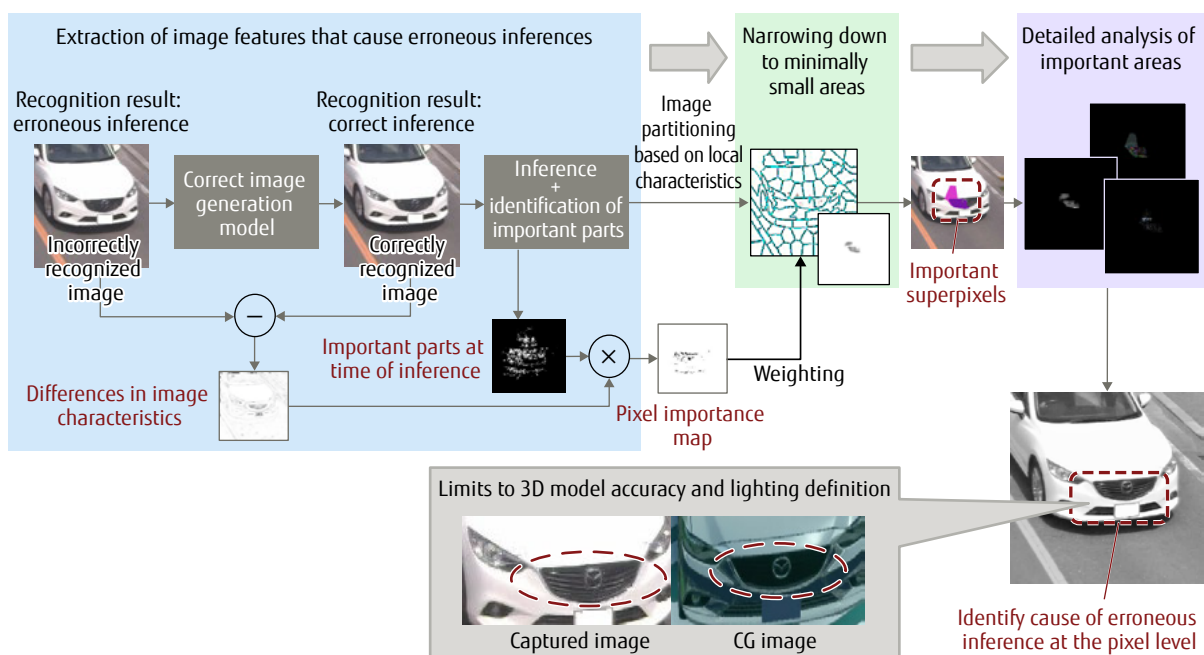
identifies parts of an image that cause errors in recognition at the pixel level (**Figure 5**) [2].

In the above way, we are achieving sufficient video coverage and quality by combining a massive on-site video dataset with proprietary CG analysis technology as a foundation for high-reliability AI video recognition.

### 3.2 Shortening of development cycle to enhance learning model

This subsection describes the importance of shortening the development cycle and the development technologies for doing so.

The training/testing cycle needs to be repeated to raise the accuracy of AI video recognition. In this regard, various types of automatic training technologies have been proposed, but at present, they are incapable of achieving video recognition in complex natural environments. This has made it necessary to repeat the training/testing cycle in a manual manner by expert researchers and developers. Consequently, to maximize the number of cycles within a limited development period, it is essential that training/testing time be shortened.



**Figure 5**  
Technology for analyzing CG problem parts.

- 1) Automatic testing and efficient retraining for a ten-fold increase in speed

We are setting up a dedicated test environment for AI video recognition systems to both automate the testing of learning models themselves and automate system testing using a large number of video streams.

In this environment, large quantities of image and video data are delivered according to scenario and results are automatically judged to be good or not good (OK/NG) and compiled into a report. This essentially eliminates the human labor expended for testing regardless of the scale of testing.

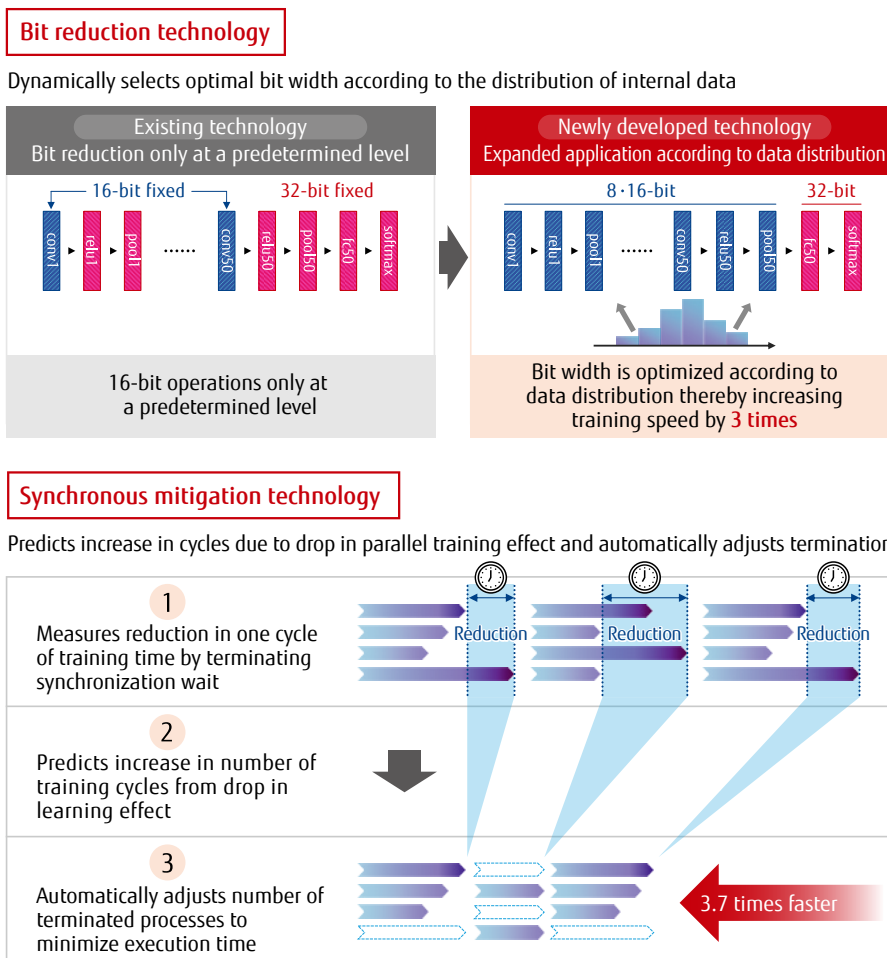
In addition to the above, we are applying a proprietary training-data extraction technique to the retraining process. By combining the minimization of training data for retraining purposes, the optimal selection of training methods, etc., we have succeeded

in shortening the retraining and testing time in one cycle to one-tenth that of simple repetition.

- 2) Content-Aware Computing technology for a ten-fold increase in training speed

We are applying technology for automatically increasing training speed to further reduce training time.

Content-Aware Computing technology [3] developed by Fujitsu Laboratories performs dynamic analysis of AI processing code. The results of this analysis are then used to combine bit reduction technology according to the distribution of data and synchronous mitigation technology that automatically minimizes wait times among parallel training processes to maximize the parallel-training effect without affecting training-model performance. This technology reduces computational complexity and enables training speed to be automatically increased by up to 10 times (Figure 6).



Adapted from FUJITSU JOURNAL

**Figure 6**  
Content-Aware Computing technology.

We plan to apply this technology to the sequential development of AI video recognition systems and to the field of road management as well.

In this way, we are achieving sufficient repetition of the development cycle within a limited development period and rapid improvement in recognition performance by automating the testing process, making retraining more efficient, and optimizing training automatically.

### 3.3 Accelerated AI processing for large-volume video recognition

Optimal distributed processing of high-load AI processing on CPUs and GPUs is essential to achieving simultaneous processing of large volumes of video on servers. In addition, performing distributed design in an optimal manner requires highly accurate analysis of data transfer bandwidth and bottlenecks in code.

With the above in mind, we are using a design support environment [4] developed by Fujitsu Laboratories for systems using high-performance accelerators such as GPUs and FPGAs. Compared with ordinary tools, this support environment features time-series analysis of processing load and data transfer bandwidth. In this way, we can design code with optimal distribution of load from the initial stage of development and maximize AI recognition processing speed over the entire system.

With this design support environment, there are cases in which learning speed can be increased by 40 times just in the initial stage of development. It is especially effective in the design of optimal load distribution in new code.

In the above way, we are achieving a system capable of large-volume AI video recognition by constructing a development environment that can easily accelerate AI processing without the need of expert AI developers.

### 3.4 Development support team

Up to this point, we have been describing development technologies, but here we would like to touch upon the development team. In contrast to existing logic-based IT, uncertainty is inherent in existing AI technology. For this reason, it is extremely important that all stakeholders both inside and outside the

company including customers have a common understanding as to what can and cannot be achieved by AI video recognition. In addition to technical aspects, project management must keep in mind that such a common understanding takes on more importance in AI-related development than in the development of an ordinary IT system.

## 4. Application example

This section describes an example of a system that applies the technologies described up to here.

We are proposing an AI detection system to the Ministry of Land, Infrastructure, Transport and Tourism. The aim of this system is to detect stranded vehicles or abnormal driving by AI video recognition and to notify road management operators of those events so that a rapid response can be mounted. An overview of this system is shown in **Figure 7** [5].

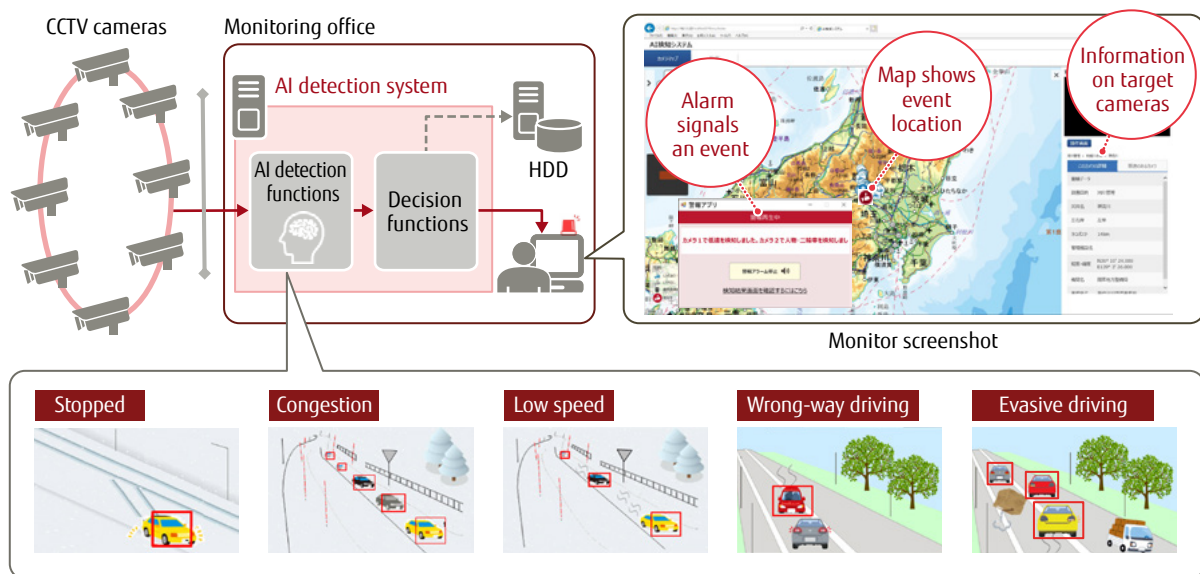
This system is now being used in actual operations after customer testing demonstrated its effectiveness [6]. In addition, many customers in the fields of road management and public infrastructure have taken delivery and begun use of various AI video recognition systems.

## 5. Conclusion and future plans

This article introduced development technologies for a high-reliability AI video recognition system in the field of road management. We have established a development process from three viewpoints: a huge video dataset and CG analysis technology, automating and efficiency-raising technology for shortening the development cycle, and the acceleration of AI processing. We are also applying new AI development technologies on an ongoing basis.

Going forward, we will further expand the range of using monitoring cameras by expanding the application field of AI video recognition technologies. For example, we have already completed the development of river-overflow detection technology in response to recent flooding in Japan based on AI technology cultivated in the field of road management. We also plan to study the application of AI video recognition technologies to Smart City, Mobility as a Service (MaaS), and other fields.





**Figure 7**  
Overview of AI detection system.

-----  
All company and product names mentioned herein are trademarks or registered trademarks of their respective owners.

Analysis. (in Japanese).

[https://www.ktr.mlit.go.jp/ktr\\_content/content/000746665.pdf](https://www.ktr.mlit.go.jp/ktr_content/content/000746665.pdf)

## References and Notes

- [1] FUJITSU BLOG: How Fujitsu is using AI to make Japan's roads and rivers safer.  
<https://blog.global.fujitsu.com/fgb/2019-11-20/how-fujitsu-is-using-ai-to-make-japans-roads-and-rivers-safer/>
- [2] T. Kubota et al.: "A proposal of a method for analyzing causes of incorrect detection when detecting objects using Deep Learning," IEICE Tech. Rep., Vol. 119, No. 317, AI2019-30, pp. 1-6, 2019. (in Japanese).
- [3] Fujitsu Laboratories: Fujitsu Develops Technology to Automatically Adjust Computing Accuracy to Accelerate AI Processing by 10 Fold.  
<https://www.fujitsu.com/global/about/resources/news/press-releases/2019/1025-02.html>
- [4] Y. Tomita et al.: "Tool Environment Supporting FPGA Accelerator Development," IPSJ SIG Technical Report, System and LSI Design Methodology (SLDM), 2017-SLDM-181 (31), 1-6 (2017), ISSN 2188-8639. (in Japanese).
- [5] Fujitsu: "FUJITSU Public Sector Solution: AI Detection System" —Vehicle Detection System Using AI. (in Japanese).  
<https://pr.fujitsu.com/jp/news/2019/07/26-1.html>
- [6] Kanto Regional Development Bureau, Takasaki River and National Highway Office, Disaster Management Section: Detection of Stranded Vehicles by AI Image



**Daisuke Yamaoka**

Fujitsu Limited, Social Systems Unit  
Mr. Yamaoka is engaged in the development of social infrastructure systems using AI video recognition technology.



**Shin Fujita**

Fujitsu Limited, Social Systems Unit  
Mr. Fujita is engaged in the design and construction of social infrastructure systems using AI video recognition technology.



**Hiroyuki Mizuno**

Fujitsu Limited, Social Systems Unit  
Mr. Mizuno is engaged in the design and construction of social infrastructure systems using AI video recognition technology.



**Kenichi Ota**

Fujitsu Limited, Social Systems Unit  
Mr. Ota is engaged in the planning and development of social infrastructure systems using AI video recognition technology.



**Masaki Miura**

Fujitsu Limited, Social Systems Unit  
Mr. Miura is engaged in the development of social infrastructure systems using AI video recognition technology.

This article first appeared in Fujitsu Technical Review, one of Fujitsu's technical information media. Please check out the other articles.

Fujitsu Technical Review

<https://www.fujitsu.com/global/technicalreview/>

