Application Examples and Issues of AI Technology

Toshihiro MATSUMOTO
Wataru HASEGAWA
Masayuki KISHIDA

Abstract

Artificial intelligence ("AI" hereinafter) has been evolving at a dramatic pace in recent years. Against the background of the technological innovation referred to as "deep learning", we have now entered the third AI boom period in which the performance of AI technology has vastly improved in comparison with past technology and we are now starting to see the possibilities of applications to a wide range of areas. In particular, there have been some very significant achievements in fields such as image processing and voice recognition.

FUJITSU TEN has been developing technology for environmental recognition in and around vehicles by using our sensor products, one of our strong points, as we strive to improve sensing performance and contribute new functions by applying AI technology for use in fields beyond just that of image recognition.

This paper describes the details of these efforts in the form of examples of AI application, as well as discussions of the difficulty and issues of AI construction.

1 Introduction

Artificial intelligence ("AI" hereinafter) has been evolving at a dramatic pace in recent years. Against a backdrop of the technological innovation referred to as "deep learning", we have now entered the third AI boom period in which the performance of AI technology has vastly improved in comparison with past technology and we are now starting to see the possibilities of applications to a wide range of areas.

In most cases, the abilities achieved by AI depend on the amount and quality of the data on which the corresponding AI is based. For this reason, AI is incorporated into the services of all Internet industries that retain large amounts of a variety of data. Moreover, the number of persons involved in the development of AI technology has been continuously increasing, thereby causing the technology to evolve at an ever-increasing pace due to reasons such as that the AI technical framework is open source, as well as the easy availability via the Internet of samples relating to AI used for purposes such as research and contests.

FUJITSU TEN has been developing technology for environmental recognition in and around vehicles by using our sensor products, one of our strong points, as we seek to improve sensing performance and contribute new functions by applying AI technology for use in fields beyond just that of image recognition. Nevertheless, there are numerous issues due to the technology lacking established methods and techniques. This paper provides a simple explanation of this current popular topic of AI technology, as well discussing AI application examples to illustrate the difficulty of AI construction and corresponding issues.

2 What is AI Technology?

AI refers to artificially created intelligence and intellectual abilities. Although there are many different approaches to creating AI, one method is "mechanical learning". The mechanical learning method consists of attempting to attain the learning function possessed by humans through machines, namely by the use of computers, circuits and similar technology. In this section, we will provide a simple description of the learning function of the human brain and explain methods of mechanical learning that can simulate this function.

2.1 Learning Function of the Brain

Nerve cells or neurons are a part of the brain related to the learning function (**Fig. 1**). Stimulus is input to each dendrite of a neuron, and stimulus in accordance with the amount and timing of the original stimulus is then output by the axons.

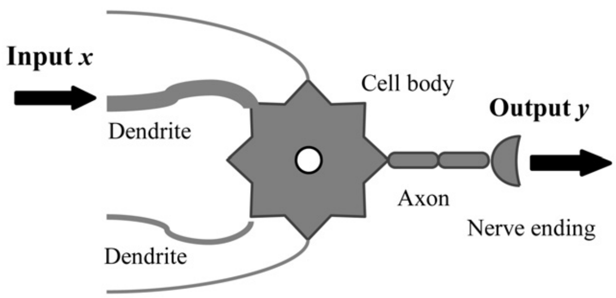


Fig. 1 Overview of a Neuron

Although the functions of single cells are comparatively simple, such that a single cell can rarely function by itself, configuring a network by the interconnection of mutual inputs and outputs between cells enables complex functioning such as generating outputs corresponding to multiple inputs (**Fig. 2**). The junction of an axon and dendrite between each cell is strengthened by the input of stimulus from the axon, transforming the junction so that the stimulus can pass through with greater ease. Conversely, junctions not receiving inputs of stimulus gradually become weaker, making it more difficult for stimulus to pass through them. Repeatedly inputting data of the same system causes these changes in junction strength to converge at a fixed value so that the relationship between outputs in relation to inputs of a network is firmly established. In other words, the relationship of network inputs and outputs can be considered to be "learned". The complex intertwining of these functions result in attainment of the learning function of the brain.

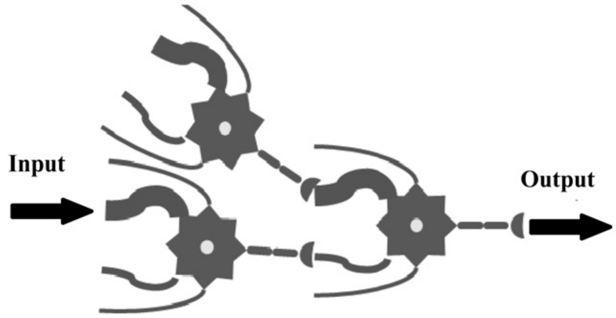


Fig. 2 Neuron Connectivity

2.2 Learning Function of AI

"Neural network" (NN) is a type of machine learning that has been receiving increasing attention in recent years. We will provide an example of an NN in relation to the AI learning function. An NN simplifies the network of neurons so that they it can be recreated with a computer by use of an artificial neuron model (**Fig. 3**) that simulates neurons. Artificial neurons function in the same way as actual neurons with junction strength corresponding to each input and output performed in accordance with inputs. These are interconnected to form a network that performs output in accordance with complex inputs in the same manner as an actual neuron network. Additionally, repeatedly inputting data of the same system results in strengths "w" of the junctions between each artificial neuron converging at a fixed value so that the relationship between outputs corresponding to the inputs is firmly established. In other words, an artificial neuron network possesses the same "learning" function as that of actual neurons.

Although there are various types of network configuration, we will configure one with a layered format as shown in **Fig. 4** to serve as an example. This network can be roughly divided into four masses. As these masses are referred to as "layers", the network shown in **Fig. 4** is called a "four-layer network". The first layer consists of the inputs themselves entering into the network whereas the inputs of the second layer are the outputs of the first layer, with the corresponding output results being passed on again to the next layer. Continuing to add layers in this manner creates a large amount of combinations of artificial neurons, thereby enabling the processing of complex inputs. There has been a trend in recent years to expand the number of layers in order to achieve even more complex learning, with learning using large-scale multi-layered networks becoming a commonplace.

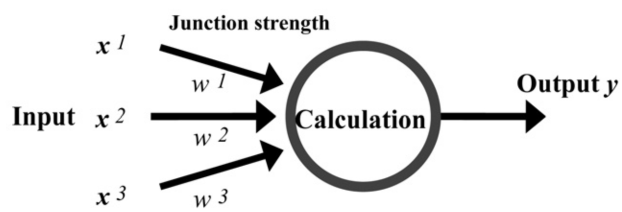


Fig. 3 Example of an Artificial Neuron Model

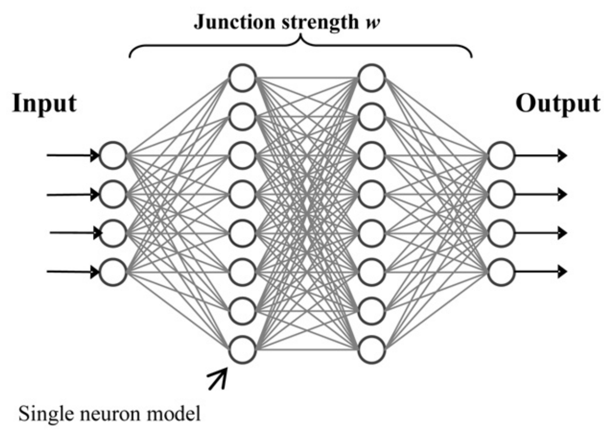


Fig. 4 Overview of Neural Network

In this case, because "learning" by repetitive data input can be performed, an NN using artificial neurons is capable of simulating the functions of an actual neuron network. However, giving consideration to the engineering perspective, establishing an intentionally desired input/output relationship makes a network easier to use than learning an input/output relationship in an arbitrary manner. By doing so, mechanical learning can provide intentional "learning" in which the desired input/output relationship is achieved. Intentional learning in this case indicates the establishment of the desired input/output relationship by learning achieved through repetitively "teaching" using outputs that are appropriate in relation to the corresponding inputs. Here, the fact that changes to the values of junction strength "w" also result in changes in the input/output relationship of the network is utilized. **Fig. 5** shows how "w" is adjusted by intentional learning. This is incorrect however, because the output corresponding to the input of an image of a dog is a cat. For this reason, "w" is adjusted so that the output corresponding to this input is also a dog. Repeating this type of adjustment enables intentional learning and the creation of an intentionally desired input/output relationship.

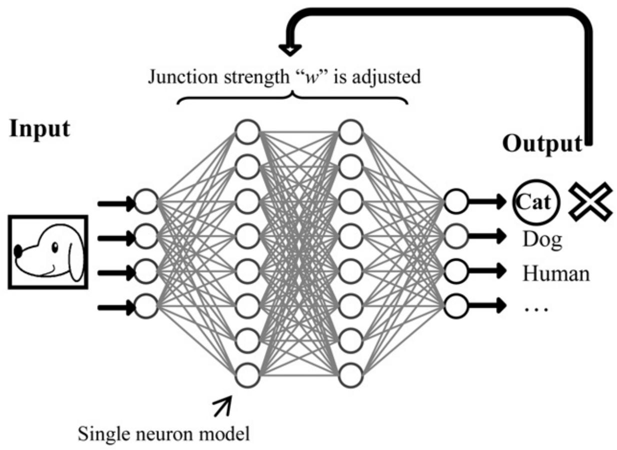


Fig. 5 Adjustment of "w" during Learning.

3 Application Examples and Issues of AI Technology

In AI technology, "network structure" (configuration of the network) and "learning data" (repetitively input data for learning) are the two elements that most contribute to recognition capability. In this section, we use some of the technology that uses AI as an example to discuss the issues of AI application arising from differences in recognition capability resulting from differences in network structure and learning data.

3.1 Example of Changes in Recognition Capability According to Differences in Network Structure (Detection of Overhead Objects)

A millimeter-wave radar detects objects by radiating a beam at a target object and receiving the waves that are reflected back from that object. One of the issues with this type of radar is determining whether a distant reflected object is located on the ground, or is at an elevated position such as an overpass or an information sign (**Fig. 6**).

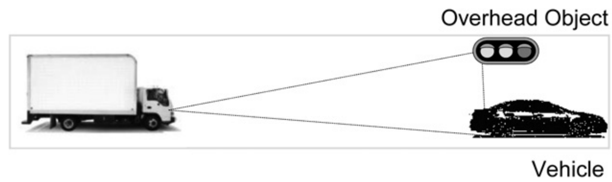


Fig. 6 Determination of Overhead Objects by Millimeter-Wave Radar

A radar elements closely related to this issue is the characteristic that the more the height of the reflected object increases, the greater the change in the reflection level due to the distance. Focusing on this characteristic, we attempted to use AI for determining an overhead object by using the distance to the reflected object and corresponding changes in the reflection level. Radar output data was used as learning data for AI, and accuracy of the learning data was divided for use into overhead and underside objects from the vehicle driving record.

Table 1 shows examples of the two types of networks used as an example of the changes in recognition capability according to the differences in AI networks.

Table 1 Identification Results of Each Network

Number of Network Layers	Identification Rate
Large-scale network (22 layers)	52.5%
Small-scale network (5 layers)	97.5%

Although a large-scale network configuration having a large number of layers generally tends to have a higher recognition capability, the results in this case indicate that the small-scale network achieved a higher recognition capability than the large-scale network. In regards to network configuration, this shows us that the small-scale network configuration used in this case was better suited for data processing.

The current AI technology is not capable of performing general-purpose processing, but rather it is specialized only to perform some single-purpose tasks. For this reason, there are many cases where simply using a network structure generally believed to have a high recognition capability does not result in an improved recognition capability depending on the input data, such that the desired objective is not obtained. In cases where the input data and desired output have already been determined, it is necessary to give consideration to the data type and characteristics needed to perform proper learning, and then select and create a network configuration that is suitable for the corresponding processing.

3.2 Example of Changes in Recognition Capability According to Differences in Learning Data (Empty Parking Space Detection Technology)

The field of image recognition is one where current AI technology most excels and also where fundamental elements of learning methods have been established. Nonetheless, it is necessary to correctly create and adjust the input and learning data in order to achieve the desired output even in the image recognition field.

One target application of AI technology we discuss in this section is the "empty parking space detection system" that detects empty parking spaces from the images of a vehicle's front camera. Using the technology of this system for detecting the positions of parked vehicles as an example, we will use two types of learning data to show the changes in AI recognition capability according to differences in learning data. In this case, although the detection results of parked vehicles are displayed as rectangles superimposed in the position of vehicles in the images, the assumptions of the target application require the rectangles to have a uniform relationship with the positions of parked cars and for it to be possible to discern each individual parked car. Additionally, learning data must be created and adjusted into general-purpose data sets that include the images of parked vehicles in every direction, and in a manner that is compatible with system processing.

3.2.1 Learning Results Using General-Purpose Images

The results of learning and recognition using general-purpose data sets are shown in **Fig. 7**.

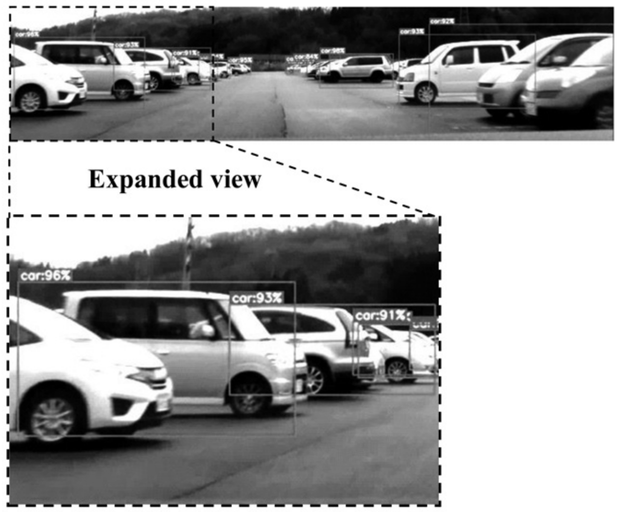


Fig. 7 Example of Judgment Results Using General-Purpose Data Learning

Although vehicle recognition is achieved, the rectangles indicating vehicle positions only roughly match the actual positions. As a result, these vehicle coordinates could not be used to clearly identify empty parking spaces. Additionally, these results could not be used for system processing because the displayed rectangles indicate regions that combine multiple vehicles.

3.2.2 Learning Results Using Partial Data

The operation of the empty parking space detection system is assumed to be limited to the environment within a parking lot. Although the detection of entire vehicles is difficult under such conditions, there is a good possibility of being able to constantly detect the front/rear face regions of parked vehicles. Accordingly, we created images limited to the front/rear face regions of parked vehicles, and performed learning and recognition accordingly. The recognition results are shown in **Fig. 8**.

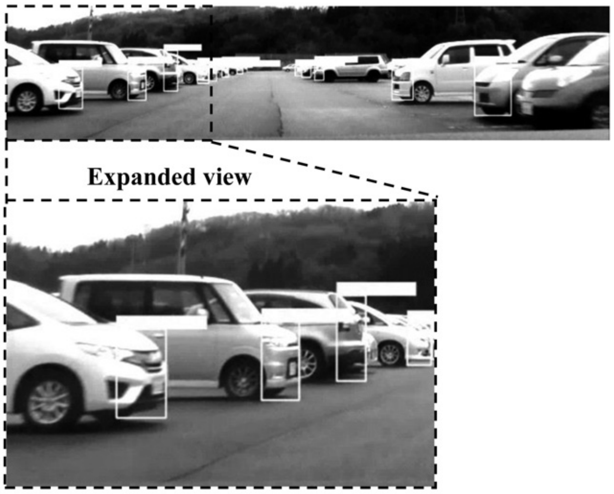


Fig. 8 Example of Judgment Results Using Learning from Partial Images

Compared to the results with learning using images of entire vehicles, we can see that the positions of the rectangles and those of vehicles are contained within certain ranges. Additionally, the rectangular regions are contained within the ranges of individual vehicles, confirming that the data can be used as judgment information.

3.2.3 Comparison of Learning Results According to Differences in Learning Data

Fig. 9 shows a comparison of the number of vehicles detected and targets for AI using each type of learning data. The detection results here only indicate the number of vehicles that were detected when the differences between the positions of vehicles and the range of rectangles were within the specified range. Additionally, the targets were established in accordance with changes in circumstances due to the necessity to correctly judge the positions and number of vehicles given the characteristics of the system. It is possible to apply this to the system if the number of detected vehicles is above the target but not possible if the number is substantially below the target. Note that the number of vehicles within the range captured in the images decreases as the elapsed time progresses. The learning results using partial data are above or near the target across the entire graph, indicating that it can be used for the system, whereas the learning results using general-purpose data are substantially below the target for the zero to six seconds period, indicating that it cannot be used for the system. These results indicate that the selection of appropriate learning data is extremely important in order for AI to achieve the desired recognition capability. This also tells us that rather than only using standard processing methods and data, selecting and creating appropriate input and learning data that matches the objective results in differences in learning time and recognition rate, and allows us to obtain recognition capability that is closer to the desired output results.

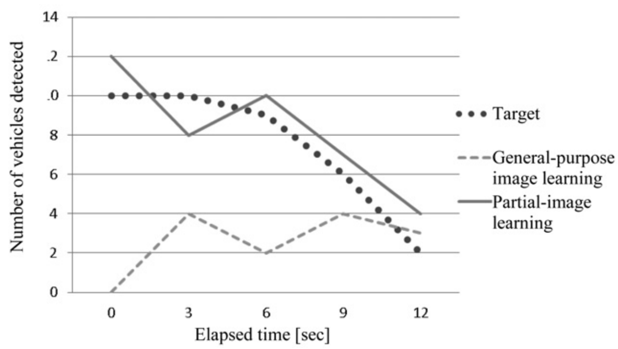


Fig. 9 Differences in Number of Vehicles Detected

4 Conclusion

In this paper, we used a part of a system using AI as an example to show changes in recognition capability according to differences in learning data and network configuration. It is extremely important to select network configuration and learning data that is appropriate for the corresponding system in order to provide AI with the desired recognition capability.

The research being conducted on the use of AI technology in various fields, and new network configurations and technology being developed almost daily engender great expectations for the future of this technology. Although there have been significant achievements in fields such as image processing and voice recognition, the AI used in these cases has been suitable for specialized tasks but it has not been capable of any other processing. In order for AI to be applied to various technologies, the data and knowledge possessed by the companies related to those technologies is very important. Use of this accumulated data and knowledge can enable the development of AI technology capable of even higher levels of performance. We plan to continue development of new technology using AI in order to achieve greater sophistication and functionality of sensing technology.

Acknowledgments

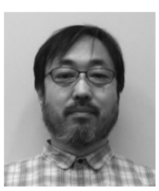
We would like to express our deepest gratitude to the members of the AI Platform Dept. of the FUJITSU Advanced System Development Div. for their ample assistance during the writing of this paper.

References

1) K. Simonyan, A. Zisserman: Very Deep Convolutional Networks for Large-Scale Image Recognition, ICLR, 2015.
2) Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, 86(11):2278-2324, November 1998.

Profiles of Writers

Toshihiro MATSUMOTO
VICT Engineering Group
Engineering R&D Dept



Wataru HASEGAWA
VICT Engineering Group
Engineering R&D Dept

Masayuki KISHIDA
VICT Engineering Group
Engineering R&D Dept