Detection of Falling Objects in Tracks by Using Deep Learning※1

Taiki Yamamoto

Keywords Railroad, Detection of falling object, Deep learning

Abstract

If tools, bolts, and and other objects remain on railroad tracks after the completion of maintenance and improvement work, there is a serious threat that a dangerous accident can occur. In order to eliminate accidents caused by such fall objects, checking work for objects is carried out immediately after the work ends. For automating this checking work, we developed a detection technology for fallen objects on railroad tracks by using an image processing function of a camera.

For the detection of fallen objects, we adopted the method of deep learning that is noted for its high performance in resolving prediction problems and its attention to detail. By the deep learning, the presence of fallen objects displayed on the screen can be detected. By virtue of this expertise, the identification accuracy of fallen objects and debris is raised to 99.3%. If this technology is adopted for the dedicated inspection vehicle that is used for the checkout work for finding fallen objects and debris on the tracks after the maintenance and/or improvement work, it will contribute considerably to the automation of fallen object detection work for railroad facilities.

1 Preface

Maintenance and improvement work for railroad facilities is generally carried out at night after the last train and before the first train of the following day. If any tools, bolts, or fallen objects used for the work are carelessly left on the railroad tracks, a danger for a serious accident like derailment presents. For this reason, checking work is carried out by workers after the completion of night work in order to find any objects left behind.

Such checking work may be done by using a dedicated inspection vehicle. Such an inspection vehicle is equipped with an obstacle detection bar under the vehicle floor. If anything comes in contact with this bar, it is sensed as an obstacle. In many cases, however, most fallen objects on the track do not enter the minimum rolling stock gauge and the obstacle detection bar may fail to detect them. For the current checkout work, multiple individuals stand on the track, side-by-side in line and search for any fallen objects by illuminating their steps with a flashlight. Such a work is taxing on workers. Against this background, the market calls for a technology for automatically detecting fallen objects on railway tracks. This paper introduces an automatic detection technology for fallen objects on the railway tracks by using deep learning.

2 Fallen Object Detection Approach

For the detection of fallen objects, an area where a camera tilted downwards is installed on a vehicle roof for view. Fig. 1 shows an example of an

Fig. 1 Example of Input Image

A normal image is shown on the left and an abnormal image on the right. A normal image is devised to show ballast, sleepers, and rails only, and the abnormal image shows a fallen object.

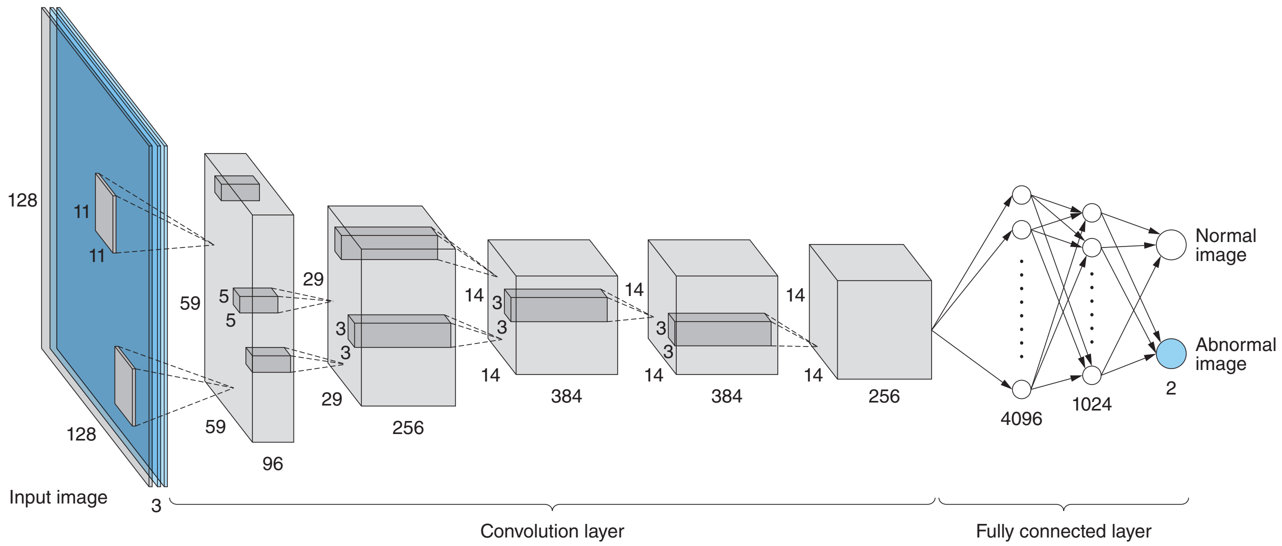

Fig. 2 Network Configuration Used for Auto-Detection Techniques

A network configuration is shown where a fully connected layer of AlexNet is made to adapt this approach by fine tuning.

input image. In this approach, an accurate image is taken, only of the ballast (gravels laid on roads and tracks), sleepers, and rails. An image where any fallen object is taken is regarded as an abnormal image. In this approach, a technology of deep learning is used only for two classes of recognition: a normal image and an abnormal image. Deep learning is one machine learning approach where a large-scale neural network having deep structures is employed. Recently, this approach has been recognized in the world for its outstanding performance in the field of identification and prediction of problems.

This approach employs a network structure proposed by Alex et al. ("AlexNet" hereafter). Among many deep learning network structures, the AlexNet has proven track records in the field of image object recognition. Since the AlexNet has a network structure where images are identified into 1000 classes, the output of the fully connected layer also contains 1000 classes. Apart from this, the output of our approach involves only two classes: normal image or an abnormal image. For this reason, it is necessary to make fine-tuning beforehand in order to adapt the fully connected layer of the AlexNet to our approach. Fine-tuning is a process when a network finishing learning a certain problem, it is then made to learn the last part of the weight again so that it can then adapt it to another problem. The convolution layer of deep learning can be used as an extractor feature. In addition, it has a characteristic capable of extraction common to a variety of problems in the image recognition field. As such, fine tuning is believed to be effective. Still more, fine tuning offers an advantage featuring efficient learning based on a small amount of learning data. In this approach, the fully connected layer consisting of three layers can be made to learn newly with a vector input of 4096 dimensions obtained from the AlexNet convolution layer. Fig. 2 shows a network configuration used for auto-detection techniques. Using a network obtained from the aforementioned learning, 2-class recognition of normal image and abnormal image is carried out.

3 Accuracy Verification

We verified an accuracy of this technology through the photo-taking of normal and abnormal images in the daytime inside the railroad track. The number of data was 26,562 in total of both normal and abnormal images. 87.5% of total data volume for both normal and abnormal images was used as the learning data respectively. Each remainder of 12.5% was used as evaluation data. Fig. 3 shows a relationship between learning frequencies and identification accuracy. Around 300 times of learning frequencies, the identification accuracy came to convergence. The obtained identification accuracy was 99.7% for the learning data and 99.3% for the evaluation data. Fig. 4 shows an example of successful fallen object detection. Input images are shown on the left side while the right side shows images where the influence rate on the recognition is visualized.

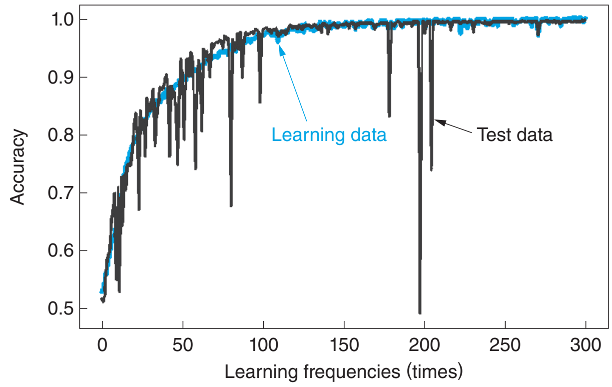

Fig. 3 Relationship between Learning Frequencies and Identification Accuracy

The axis of ordinates shows identification accuracy and that of abscissas shows learning frequencies. The blue line shows the result of identification of an image included in learning and the black line indicates the result of identification of an image not included in learning.

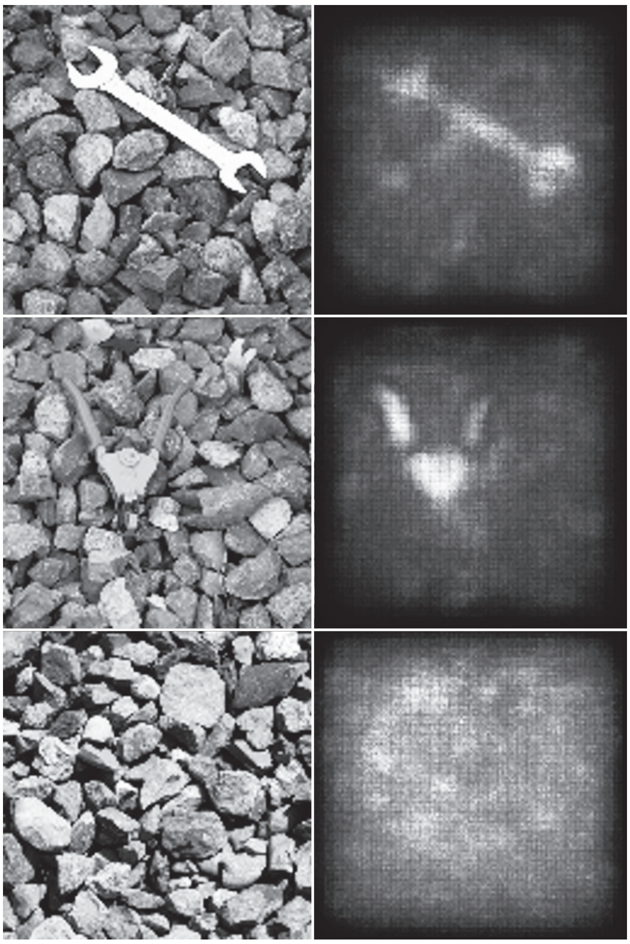

Fig. 4 Example of Successful Fallen Object Detection

An example of images of the detected fallen object is shown. Input images are shown on the left side while the right side shows images where the influence rate on recognition is visualized. The influence rate on recognition by deep learning is shown with grayscale values. The closer to white, the greater the influence rate.

4 Postscript

We introduced a method for detecting fallen objects on railway tracks by using deep learning. With this technology, we were able to show the possibility of automating the work of checking for fall objects. In the future, we aim to put it into practical use by verifying it on an actual vehicle assuming actual operation.

・All product and company names mentioned in this paper are the trademarks and/or service marks of their respective owners.

(Note)

※1. An approach to let a computer learn so that the computer can execute a human being's tasks such as voice recognition, image definition, and prediction.