Neurocomputing 222 (2017) 127–136

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Automatic multi-fault recognition in TFDS based on convolutional neural network

Junhua Sun*, Zhongwen Xiao, Yanxia Xie

School of Instrumentation Science and Opto-electronics Engineering, Beihang University, Beijing 100191, China

ARTICLE INFO

Communicated by H.R. Karimi

Keywords:
Freight train faults
Multi-fault recognition
Image recognition
Convolutional neural network
Region detection

ABSTRACT

In recent years, more than 300 sets of Trouble of Running Freight Train Detection Systems (TFDSs) have been installed on railway to monitor the safety of running freight trains in China. However, TFDS is simply responsible for capturing, transmitting, and storing images, and fails to recognize faults automatically. Motivated by the success of convolutional neural networks (CNNs) in the tasks of image recognition, this paper makes essential use of CNN models and proposes an automatic fault recognition system (AFRS) for recognizing four typical faults in TFDS simultaneously. AFRS is a two-stage system: In the first stage, a coarse-to-fine scheme based on CNN model is adopted to detect the target regions of side frame keys (SFKs) and shaft bolts (SBs) simultaneously. In the second stage, we establish another CNN model for multi-fault determination to recognize four typical faults emerged in the target regions of SFKs and SBs. The experimental results show that this system has an excellent performance of multi-fault recognition in TFDS. High recognition accuracy rates, low false ratios and low omission ratios are obtained for all the four typical faults, demonstrating the high recognition ability and robustness against various low quality imaging situations.

1. Introduction

In recent years, with the development of China railway, railway safety has gained increasing attention. Trouble of Running Freight Train Detection System (TFDS) plays a crucial role in both transportation safety and efficiency of freight trains. TFDS is installed on railway to monitor the safety of running freight trains. The working process is as follows. First, it automatically captures images of some vital parts of freight trains. Then, the images are transmitted to the data server in monitor room for analysis by indoor inspectors to find out the faults of relevant parts. By observing the TFDS images, the indoor inspectors will inform the outdoor inspectors to confirm and handle the problem if the fault exists [1]. The TFDS effectively reduces manual labor intensity in fault detection of freight trains. However, this system is simply responsible for capturing, transmitting, and storing images, and fails to recognize faults automatically due to some difficulties such as the diversity and complexity of faults and some low quality images caused by bad weather, shooting angle change, illumination change, and so on. Currently, the work of fault detection is still completed by human eyes of indoor inspectors to observe images. This work still needs a huge amount of workload. Moreover, fatigue of indoor inspectors will easily cause missing and error detection.

To reduce the labor intensity of inspectors, the research on automatic fault recognition of TFDS has attracted much attention, but this task is still challenging. So, further studies of automatic fault recognition of TFDS are still necessary. A deep-learning-based algorithm called convolutional neural network (CNN) has been widely used in the tasks of computer vision and image processing [2,3]. Motivated by the success of convolutional neural networks (CNNs) in image recognition tasks, this paper makes essential use of CNN models and proposes an automatic fault recognition system (AFRS) for recognizing four typical faults in TFDS simultaneously. The whole system is showed in Fig. 1. AFRS is a two-stage system. In the first stage, a coarse-to-fine scheme is adopted to detect the target regions of side frame keys (SFKs) and shaft bolts (SBs) simultaneously. A CNN-based detection model is trained for coarse detection, and then with some prior information of geometrical and spatial position relationship, target regions are accurately detected. In the second stage, we establish another CNN model for multi-fault determination, which recognizes the target regions detected in the first stage. The CNN-based model adaptively represents image features by training so as to achieve a high recognition rate in our task. Unlike models based on hand-engineered feature extractor, the powerful CNN-based models automatically extract inherent and generalizable features of target objects, so the CNN-based models are robust enough to provide excellent performance with various outdoor imaging situations. By introducing the CNNs to

* Corresponding author.
E-mail address: sjh@buaa.edu.cn (J. Sun).

http://dx.doi.org/10.1016/j.neucom.2016.10.018
Received 6 May 2016; Received in revised form 25 August 2016; Accepted 14 October 2016
Available online 18 October 2016
0925-2312/

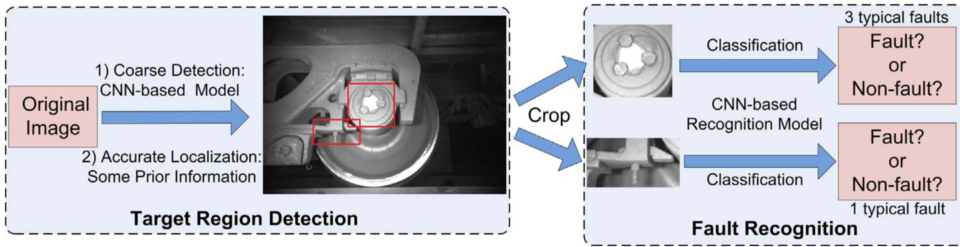

Fig. 1. Automatic fault recognition system (AFRS). The first stage is target region detection. In this stage, target regions of SFKs and SBs are accurately detected. The second stage is fault recognition for four typical faults emerged in the target regions of SFKs and SBs. In this stage, the detected regions of SFKs and SBs are parallel recognized separately.

AFRS, the AFRS detects and recognizes multiple typical faults efficiently and simultaneously.

The rest of this paper is organized as follows. Section 2 presents previous research on fault recognition in TFDS and image recognition with CNNs. The whole framework of AFRS is introduced in Section 3, including the two-stage system. In Section 4, our dataset is presented, and the experimental results and performance analysis are demonstrated. Finally, Section 5 concludes this paper.

2. Related works

2.1. Fault recognition in TFDS

There are some significant previous studies about the fault recognition in TFDS. The process of the fault recognition is generally consists of two main steps: target region detection and fault determination. The target region detection is the first step of the automatic fault recognition and plays an important role in the fault determination. As for research on the region detection and localization, the methods can be grouped into three categories: template matching [4,5], image segmentation [6], and edge detection [7,9]. Meanwhile, with respect to the research on fault determination, the methods mainly consist of two categories: methods based on hand-engineered features [5–8] and learning-based algorithms [10,11].

Previous studies on fault recognition in TFDS are mainly based on hand-engineered features. Ming Wang [5] combined the discrete point sampling image, grayscale image, and edge image for template matching to detect the target regions, and then used the theory of Otsu for region segmentation to recognize two typical faults of middleware in the bottom of freight trains. Lirong Jiang [6] proposed a method of area threshold for obtaining regions of center plate bolts, and then determined faults of the center plate bolts based on a closed rectangular area descriptor. Yan Chen et al. [7] first detected the region of four center plate bolts indirectly with the theory of histogram equalization, edge detection, and hough transform, and then determined whether the bolts are missing. Guodong Sun et al. [8] proposed a shape context descriptor to recognize the fault of side frame keys in TFDS. The approaches mentioned before are mainly based on hand-engineered features such as gray projection, region segmentation, edge detection, shape context and so on. Although results of these methods based on hand-engineered features are also encouraging, they require a considerable amount of prior knowledge and engineering skills to design a good feature extractor that transforms the pixel values of an image into a suitable internal representation. In particular, a certain feature extractor is basically designed for a specific task. This means that a recognition task to a certain fault can be solved quite well with a designed method based on a hand-engineered feature extractor, but this designed method is likely loss efficacy to another type of fault.

Except for the above methods based on the hand-engineered features, several learning-based algorithms were also proposed in the problem of fault recognition. For example, Yuan Jiang et al. [10] trained a cascade framework with the adaboost algorithm to recognize sleeper springs, and Maosheng Wang [11] applied the method of decision tree to recognize the fault of side frame keys. These learning-based algorithms that upgrade the performance are all based on conventional machine-learning techniques. In fact, it is often impossible to simply adopt one machine-learning model to a specific recognition task, and it usually requires model merging, meanwhile, the careful tuning of each model is overwhelmingly labor-intensive. Moreover, conventional machine-learning techniques are limited in their ability to process natural data in their raw form. Accordingly it is hard to adopt a model based on conventional machine-learning techniques for multi-fault recognition.

For the methods based on hand-engineered features and conventional machine-learning techniques, it is really difficult to design a unified method to detect and recognize various faults simultaneously. Consequently, until recently, automatic fault recognition in TFDS is still a challenging mission.

2.2. Image recognition with CNNs

In recent years, deep learning has made major advances in solving problems such as speech recognition, visual object recognition, object detection and many other domains like drug discovery and genomics [2]. The key aspect of a deep-learning-based algorithm is that feature extractions and representations are not designed by human engineers, but they are learned from a huge number of data. Unlike conventional machine-learning techniques, deep learning methods use a general-purpose learning procedure and learn the feature representation from a huge number of data. Deep networks actually implements functions of higher complexity, so that they are able to deal with more difficult problems [12]. Convolutional neural networks (CNNs) use a very similar architecture to some time-delay neural networks [13,14], since the tiling of neuron outputs can be done in timed stages, in a manner useful for analysis of images. The CNN is fed with raw data and uses a multilayer architecture to automatically discover the general-purpose feature representations needed for detection or classification [15].

The CNNs have turned out to be exceedingly excellent at discovering intricate structures in images and are therefore applicable to the tasks of computer vision and image processing, such as image classification [16–18], object detection [19–21], critical point detection [22], scene analysis [23] and so on. LeCun [24] first proposed a classic 5-layer CNN model named LeNet-5, which was used to recognize the checks in American banks at that time. In classification tasks, recognition systems based on CNNs have made amazing achievements, especially large scale image classification. The Computer Vision Group at Microsoft Research Asia [25] developed a computer vision system based on a very deep CNN model. Surprisingly, it achieved perfect recognition performance surpassing human-level performance on ImageNet classification.

3. Proposed framework

In this section, we show the whole framework of the proposed AFRS. The proposed system consists of two CNN-based models: One is used for region detection and the other for fault determination. We first introduce the layers of a basic CNN architecture. Second, we describe a CNN-based target region detection method, which is a coarse-to-fine