Advanced Engineering Informatics 37 (2018) 79–87

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Full length article

Distress classification of class-imbalanced inspection data via correlation-maximizing weighted extreme learning machine

Keisuke Maeda[a,*], Sho Takahashi[b], Takahiro Ogawa[a], Miki Haseyama[a]

[a] *Graduate School of Information Science and Technology, Hokkaido University, Kita-14, Nishi-9, Sapporo-shi 060-0814, Japan*
[b] *Faculty of Engineering, Hokkaido University, Kita-13, Nishi-8, Sapporo-shi 060-8628, Japan*

ARTICLE INFO

Keywords:
Distress classification
Civil structures
Canonical correlation
Extreme learning machine
Class-imbalanced data
Maintenance inspection

ABSTRACT

This paper presents distress classification of class-imbalanced inspection data via correlation-maximizing weighted extreme learning machine (CMWELM). For distress classification, it is necessary to extract semantic features that can effectively distinguish multiple kinds of distress from a small amount of class-imbalanced data. In recent machine learning techniques such as general deep learning methods, since effective feature transformation from visual features to semantic features can be realized by using multiple hidden layers, a large amount of training data are required. However, since the amount of training data of civil structures becomes small, it becomes difficult to perform successful transformation by using these multiple hidden layers. On the other hand, CMWELM consists of two hidden layers. The first hidden layer performs feature transformation, which can directly extract the semantic features from visual features, and the second hidden layer performs classification with solving the class-imbalanced problem. Specifically, in the first hidden layer, the feature transformation is realized by using projections obtained by maximizing the canonical correlation between visual and text features as weight parameters of the hidden layer without designing multiple hidden layers. Furthermore, the second hidden layer enables successful training of our classifier by using weighting factors concerning the class-imbalanced problem. Consequently, CMWELM realizes accurate distress classification from a small amount of class-imbalanced data.

1. Introduction

In several countries, many civil structures such as bridges and tunnels have been constructed on the high economic growth period [1]. They will be deteriorated in the near future, whereas the number of inspectors of civil structures is limited. Since the burden on each inspector is increasing, supporting methods for maintenance inspection are required. This demand commonly exists for the maintenance of various kinds of infrastructures [2,3], and its importance is increasing year by year. From this background, development of computer vision methods for civil engineering applications has been increasing [4,5]. Especially, many researchers have studied supporting methods such as crack detection [6,7] and pothole detection [8,9]. These methods have attracted much attention since these distresses frequently occur in civil structures. However, there exist many kinds of distress [10,11] other than cracks and potholes. Therefore, for supporting maintenance inspection, a method for automatically classifying various distresses is necessary.

In order to classify many kinds of distress, construction of a classifier with ability to accurately discriminate them is the most common strategy for supporting inspectors. Generally, in order to realize such a classifier for successfully classifying multiple kinds of distress, it is desired to prepare training data in such a way that the number of training data becomes almost equal for each distress. However, since frequency of occurrence of each distress is different, that is, the training data become class-imbalanced data, general machine learning methods may not achieve sufficient performance of the distress classification. To overcome the above problem, we have proposed a distress classification method of class-imbalanced data [12]. Specifically, we have constructed a weighted extreme learning machine [13], which is a three layered neural network. Although our previously reported method realized accurate distress classification, the classification performance was limited since feature transformation through the previous networks did not have sufficient capability to distinguish many kinds of distress. Therefore, novel feature transformation is required to obtain semantic features from visual features, which can represent characteristics of distresses.

In recent image recognition fields, it has been reported that multi-

* Corresponding author.
E-mail addresses: maeda@lmd.ist.hokudai.ac.jp (K. Maeda), stakahashi@eng.hokudai.ac.jp (S. Takahashi), ogawa@lmd.ist.hokudai.ac.jp (T. Ogawa), miki@ist.hokudai.ac.jp (M. Haseyama).

https://doi.org/10.1016/j.aei.2018.04.014
Received 10 October 2017; Received in revised form 31 March 2018; Accepted 30 April 2018
1474-0346/

layered neural networks [14–18] are effective for various tasks such as image classification and object recognition. These methods can transform visual features to semantic features by using multiple hidden layers. However, these methods are not suitable for distress classification due to following reasons.

1. Convolutional neural network (CNN)-based methods [14–16] require a large amount of training data to train multiple hidden layers. When preparing a large amount of training data is difficult, fine-tuned CNNs are used instead of CNNs trained from scratch. However, the effective learning of fine-tuned CNNs is difficult when it is applied to unexpected tasks, that is, in a case in which technical data such as inspection data of civil structures with a property different from that of pre-trained data are used [19,20].
2. As another approach, multi-layer extreme learning machine (MLELM) [17] and multi-layer weighted extreme learning machine (MLWELM) [18], which are extended versions of extreme learning machine (ELM) [21], train multiple hidden layers by using an weighted ELM-based auto encoder (AE) [18]. Since almost all parameters of the WELM-based AE are set to random values, we can train these classifiers from a small amount of training data. However, since the WELM-based AE is an unsupervised learning method, it has difficulty in transforming from visual features to semantic features representing characteristics of distresses, which have several variations such as cracks and corrosion.

In order to realize effective feature transformation, we focus on relationships between distress images and text data recorded by inspectors. In maintenance inspection, inspectors take images of distress parts, and record text data, which describes categories and materials of structures and so on. It is reported that using not only distress images but also text data is effective for distress classification [22]. For example, corrosion occurs in metal parts of civil structures, but the others occur in concrete structures. Thus, distress images and text data have a correlation. Yeh et al. reported that canonical variates obtained by calculating the canonical correlation between heterogeneous features have better classification performance than original features if the heterogeneous features have semantic relevancy [23]. Thus, it is expected that visual information can be directly transformed to semantic information based on projections using the canonical correlation. Then we can calculate new features that are suitable for representing distresses without preparing a large amount of training data. Therefore, it is expected that effective feature transformation can be realized by considering the canonical correlation between visual features and text features obtained from distress images and text data, respectively. Moreover, since the distress images also include unnecessary regions such as backgrounds, the visual features contain many noisy information. On the other hand, the text features are calculated based on actual maintenance inspection recorded by inspectors, so they are very high quality features. In order to compute features with the higher representation ability using both of these features, it is necessary to project these features to a comparable feature space. Therefore, we focused on CCA, which is one of the most general methods which make it possible to project two different features to the comparable feature space. Furthermore, since the projection matrix maximizing the canonical correlation obtained by CCA can be integrated into the neural network, we can perform end-to-end learning of both feature transformation and classification. This is the reason why we chose CCA.

In this paper, we present a correlation-maximizing weighted extreme learning machine (CMWELM). CMWELM is an improved version of our previously reported method [12]. CMWELM consists of two hidden layers. The first hidden layer directly transforms visual features to semantic features, and the second hidden layer performs classification with solving the class-imbalanced problem. Specifically, the first hidden layer realizes the feature transformation by using projections maximizing the canonical correlation between visual features and text features as weight parameters of the hidden layer. Thus, since we determine the parameters by considering the canonical correlation, that is, we can set the parameters more accurately compared to MLELM, the extraction of semantic features becomes feasible without using multiple hidden layers. This is the main contribution of this paper, and the construction of the first hidden layer can solve the problem mentioned in the previous paragraph. Furthermore, the second hidden layer enables successful training of our classifier by using weight factors considering the class-imbalanced problem based on WELM. ELM-based classification methods need not to prepare a large number of training data since the number of parameters to be tuned is small. Thus, construction of the second hidden layer can solve the class-imbalanced problem. Consequently, CMWELM realizes accurate distress classification of class-imbalanced data.

This paper is organized as follows. The summaries of related work in a area of the computing in civil engineering are explained in Section 2. The inspection data including distress images and text data used in our study are shown in Section 3. The proposed method based on CMWELM is presented in Section 4. Experimental results are shown to verify its superiority in Section 5. Finally, conclusions are shown in Section 6.

2. Related work in area of computing in civil engineering

In this section, we explain the summaries of related work in the area of the computing in civil engineering. In order to realize the automatic inspection, various projects have been conducted by utilizing robotic technologies [2,3]. Although a manual inspection process often occurs human errors due to its subjectivity, robotic technologies can overcome these disadvantages and provide the high quality inspection. A survey of the novel developed robotic tunnel inspection systems has been explained in [2]. Furthermore, in [3], the usage of a teleoperated robotic platform, which is used for assessment of the condition of power line infrastructures, has been provided. Koch et al. have presented the current state of practice in visual condition assessment of civil infrastructures [4]. Furthermore, they have categorized several state-of-the-art computer vision methodologies, which are used to automate the process of defect and damage detection as well as assessment. In addition, Tizani et al. have presented general directions and challenges related to computing in civil and building engineering [5].

The crack and pothole are major distresses, and several methods for automatic detection have been proposed [6–9]. Especially, Zou et al. have proposed CrackTree algorithm, which can consider some problems such as low contrast, intensity inhomogeneity and shadows with similar intensity to the cracks [6]. Thus, it can accommodate a wide variety of situations, and it is the contribution of this method. Furthermore, Li et al. have proposed a method, which integrates the processing of two-dimensional images and global positioning systems data to realize the automatic pothole detection [8]. The contribution is the construction of the integrated processing of multisensory data according to the merits of the individual sensors. Moreover, Koch et al. have proposed a method, which can detect potholes by comparing candidate pothole region with the texture of the surrounding non-defect pavement [9].

Image processing methods for various kinds of distresses have been proposed in [10,11]. Specifically, Jahanshahi et al. have evaluated several parameters that affect the performance of color wavelet-based texture analysis algorithms for detecting corrosion by using the depth perception [10]. This method improved the reliability of the corrosion detection algorithm. German et al. have proposed a novel method that automatically detects spalled regions on the surface of reinforced concrete columns and measures their properties in distress images [11]. Its contribution is the combination use of a global adaptive local entropy-based thresholding algorithm and a template matching.

3. Inspection data used in our study

This section explains the inspection data used in our study. Fig. 1