An Application of Image Recognition Technology Based on Deep Learning in Safety Review of Reservoir Dam Metal Structure

Kai Wang

1 Institute of Disaster Prevention and Mitigation, Zhejiang Institute of Hydraulics & Estuary, Hangzhou, Zhejiang, 310020, China

*Corresponding author’s e-mail: yuwg@zjwater.gov.cn

Abstract. The safety review calculation of the gate is an important job in the safety evaluation of metal structures. In accordance with the requirements of the specification, this paper has developed an intelligent software for metal structures safety review calculation, which can automate the calculation process and generate a review evaluation report. In order to simplify the labor of the manual input of a large number of gate review calculation parameters, this paper used a deep-learning-based image recognition technology to partially realize the collection of the information relevant to the calculation from the drawings of metal structures, laying the foundation for later realization of the fully intelligent gate safety review.

1. Introduction
As a water-retaining structure that intercepts the flow of river and channels to raise the water level or regulate the flow, reservoir dam is an important engineering measure to control the temporal and spatial distribution of water resources and optimize the allocation of water resources. Zhejiang Province has many mountains and few plains, which is suitable for the construction of reservoirs to give full play to the comprehensive benefits of flood control, irrigation, and power generation. At present, Zhejiang Province has 5 large (1) reservoirs, 29 large (2) reservoirs, 158 medium reservoirs, and at least 4105 small reservoirs. A comprehensive and reasonable assessment of the safety status of these reservoirs and dams can provide good technical support for the scientific dispatch and safe operation of reservoirs, and is of great significance for ensuring social public safety. Therefore, the Ministry of Water Resources promulgated and implemented the Reservoir Dam Safety Appraisal Measures on August 1, 2003. The appraisal content mainly includes evaluation of the engineering quality, dam operation management, the standard for flood control, structural safety of the dam, seepage safety, seismic safety, metal structure safety and comprehensive dam safety[1].

As an important job of the safety appraisal of reservoirs and dams, the safety review calculation of gates is an important component of the safety evaluation of metal structures. According to the Guidelines for Safety Evaluation of Reservoirs and Dams (SL 258), the calculation and analysis of gates should focus on reviewing the strength, stiffness and stability of the structure[1]. The method of calculation, load combination and control standards shall be implemented in accordance with the Code for Design of Steel Gates for Water Conservancy and Hydropower Engineering (SL 74). Specifically, the gate calculation review content mainly includes: gate panel thickness and converted stress, main beam strength and deflection, secondary beam strength and deflection, radial steel gate arm strength, radial steel gate main frame stability, lifting lugs and hinge strength, the force of opening and closing gates,
etc[2]. In order to simplify the above calculation process and improve the efficiency of metal structure safety review, the author developed the Metal Structure Safety Review Intelligent Entry and Report Generation Software (Registered Copyright No. 6064630). The advantages of the software are as follows:

Firstly, it is developed in the Java programming language, and has all material mechanics formulas and parameter tables coded in the program, which overcomes the shortcomings of traditional methods using Excel spreadsheets that are poor in encapsulation and scalability.

Secondly, the SQLite3 embedded database is used to store the section steel parameters into the database, which is convenient for users to quickly check and use the geometric and mechanic parameters of a specific steel section.

Thirdly, applying the Browser-Server architecture, the HTML, CSS, and JavaScript languages are used to develop the Internet user interface. Users of this software only need to install the chrome browser in order to run the software without having to setup an installation environment.

Finally, it can intelligently generate safety review reports for gates.

In the process of undertaking the safety review job of the metal structure of the dam safety assessment of water conservancy projects, the author found that the archives kept by the reservoir management unit are mostly paper documents. Even electronic documents are mostly scanned images, making it difficult for the computer to quickly process information. Usually engineers have to identify important information from these gate drawings and enter the data manually. Although the calculation software has been developed in accordance with relevant specifications, which has improved the work efficiency to great extent, the manual input of parameters of the materials, positions, and cross-sectional shapes of the horizontal and vertical beams, rubber seal strips and other components of the gate is still tedious and error-prone. According to preliminary statistics, the review calculation of a plane gate requires at least 86 key parameters such as materials, locations, and calculation conditions, and the review of a radial steel gate requires at least 105 key parameters. If paper document information can be intelligently imported into metal structure safety review calculation software through computer technology, work efficiency will be greatly improved.

Image recognition technology based on deep learning provides effective information technology methods for solving the problems above. In recent years, breakthroughs in deep learning technology have brought about considerable progress in image recognition technology, especially Optical Character Recognition (OCR) technology. Zhi Tian et al. proposed Connectionist Text Proposal Network (CTPN), which can accurately locate text areas in natural scene pictures[3]. Jianqi Ma et al. used Rotation Region Proposal Networks (RRPN), which can effectively perform horizontal text detection[4]. Yuchen Dai et al. adopted Fused Text Segmentation Networks (FTSN)[5], and Yuliang Liu et al. adopted Deep Matching Prior Network (DMPNet), which can support oblique text detection[6]. Baoguang Shi et al. used Convolutional Recurrent Neural Network (CRNN) and Connected Temporal Classifier (CTC) and achieved effective recognition of variable length text[7].

In this paper, CTPN, CRNN and CTC are used in combination to extract important engineering information relevant to gate safety review calculation from the metal structure drawings, which can realize the automation of the calculation input process to a certain extent.

2. Model Architecture of Neural Networks

2.1. Connectionist Text Proposal Network (CTPN).

A 3×3 spatial window is slide through the last convolutional map (conv5) of the VGG16 model. The sequential windows in each row are connected by a Bi-directional LSTM (BLSTM) recurrency. The RNN layer is connected to a 512D fully-connected layer, followed by the final output layer, which jointly predicts text/non-text scores, y-axis coordinates and side-refinement offsets of k anchors.
2.2. Convolutional Recurrent Neural Network (CRNN)
CRNN mainly includes three parts:
(1) Convolutional layers, which are mainly used to extract spatial features of images.
(2) Recurrent layers, which use deep bidirectional LSTM to extract sequence features.
(3) Transcription layers, which are used to convert the prediction result of each frame into the final tag sequence.

2.3. Connected Temporal Classifier (CTC)
Concatenated temporal classifier (CTC), a tool for sequence modeling which is widely used in the fields of speech recognition and optical character recognition, is currently the main method for variable-length annotation learning in the field of machine learning. Its main advantage is that there is no need to carry out cumbersome position labeling of sequence fragments, and the optimization of the objective function can be smoothly realized by labelling sequence tags.

3. Training dataset
The experiment uses International Conference on Document Analysis and Recognition Dataset 2013 (ICDAR 2013) as training data for CTPN network. The multi-directional text recognition data set (HUST-TR 400) is used as the training data of the CRNN+CTC network.

4. Graphs for Inference
This test uses drawing pictures of the layout, the master plan and the detailed structure of the radial steel gate of the spillway of the Baixi Reservoir in Ningbo City, Zhejiang Province. The table area or text area of the pictures above are used to test the image recognition performance of this model.
5. Experiment result and analysis

5.1. Connectionist Text Proposal Network (CTPN)

After 70,000 training steps of this model, during which the gradient optimization process is involved, the model training loss convergence curve is shown in Figure 3. The figure shows that: (a) the initial loss of the model at the beginning of this training (0 to 10000) is relatively large, ranging from 0.8 to 1.0; (b) the loss of the model in the middle period of this training (10000 to 67000) is gradually reduced, though overfitting occurs at individual step lengths, where the loss is large; (c) the model loss gradually stabilizes below 0.2 in the final training period (67000 to 70,000).

After training, we applied the model to the pictures for detection, and the detection performance can be shown from Figure 4 through 7. From Figure 4 through 7 we do the analysis as follows:

Firstly, the scene text detection model has an excellent performance on the recognition job of the plain text information area without tables (Figure 6), and can achieve accurate recognition.

Secondly, the model does not perform well enough in the detection of the text area in the table, as shown in Figure 5, where picture areas of many serial numbers and quantity information cannot be accurately identified.

The reason for the different performance of the model upon plain text pictures and table text pictures is that the images of the CTPN model training data set are all text areas in the natural scene, and there is no table boundaries around them. Therefore, the model has a better recognition effect on Figure 6, while the detection performance of Figure 4, 5, and 7 needs to be improved.

![Figure 3. Convergence of CTPN model during training.](image)

![Figure 4. Table area of sluice layout graph.](image)

![Figure 5. Table area of sluice structure layout graph.](image)
5.2. Convolutional Recurrent Neural Network (CRNN) and Connected Temporal Classifier (CTC)

The convolutional recurrent neural network (CRNN) and Connected Temporal Classifier (CTC) uses a pre-trained Chinese character recognition model to perform Chinese character recognition on the text area extracted by the CTPN model in 4.1. The pictures of texts after recognition are shown in Figure 8.

As shown in Figure 8 (a) to (d), we can get:

Firstly, the overall performance of this model when doing text recognition on plain text information areas without tables (Figure 8(c)) and regular table text areas (Figure 8(a) to (b)) is good, based on which we can effectively extract important engineering information.

Secondly, the model does not perform well when doing recognition on text areas in the irregular table. As shown in Figure 8(d), because the table is arranged irregularly, multiple columns of information are mixed together, and engineering information cannot be effectively processed. However, the overall accuracy of Chinese character recognition is high.
[a]

(b)

说明

1. 洪图中柱号、桩号以 m 计，尺寸以 mm 计。
2. 镀焊板片应进行整体组装，组装的允许偏差应符合《水利水电工程钢闸门制造安装及验收规程》（DL/T5189-94）的 8.5 条的规定。
3. 抗剪板应在新支撑与门叶钢结构相连接后焊并，焊缝为三面连续角焊缝，焊缝高度 16mm，抗剪板与接触板面必须接触紧密。
4. 停车在正门后组合时对门叶板面，采用焊，焊缝高度为 16mm。
5. 镀焊前必须进行表面处理，其质量要求达到不低于 S02 号级，清理度值在 40-70pm。
6. 焊条为 E307-16，0 级号焊条，5 毫米焊条，干板厚度 30mm，中间筋 843 环氧云铁焊条 1234，干板厚度 100mm。面焊，670-3 氧化橡胶焊一道，干焊 100mm。

Description

1. In the figure, the elevation and pile number are in m, and the size is in mm.
2. The weld panel should be assembled as a whole before leaving the factory, and the allowable deviation of the assembly should comply with the provisions of Section 8.5 of the "Specification for Manufacturing, Installation and Acceptance of Steel Gates for Water Conservancy and Hydropower Engineering" (DL/T5189-94).
3. The shear plate should be welded after the diagonal branch pipe is connected to the door leaf structure. The weld is a three-sided continuous fillet weld. The height of the weld is 16mm, and the end surface of the shear plate and the contact plate must be in close contact.
4. The lifting lugs are welded to the door leaf panel during assembly at the gate site, and the weld height is 16mm.
5. Surface pretreatment must be carried out before the gate is painted. The cleanliness level of the base metal is not lower than S02 level, and the roughness value is 40-70pm. The primer is a 702 epoxy zinc-rich paint, and the dry film thickness is 30mm. One middle layer of 842 epoxy mica primer, dry film thickness of 100mm, top coat: 670-3 oxide rubber paint, dry thickness of 100um.
6. Conclusion
We utilized a model for engineering information extraction, which combined CTPN and CRNN and CTC. CTPN is used to perform text areas detection and CRNN and CTC are used jointly to achieve text recognition results based on text areas detection results mentioned before. In summary, this model has better performance on pictures of plain text areas and regular table areas than those of irregular table areas. At this stage, the experiment results achieved so far showed that this model can be reliably used to extract important engineering information in order for the computer to process so as to automatically feed these parameters into another calculation model, which is used to perform the mechanical calculation on gates.
7. Future works
The CTPN model training data set can be improved in the following works. Adding table pictures for text areas detection training may achieve better results, thus improving the overall performance of the whole model.

Acknowledgments
This paper is funded by Youth Fund of Zhejiang Institute of Water Resources & Estuary(Project No.A18004).

References
[1] Ministry of Water Resources of the People's Republic of China. (2017) SL 258-2017. China Water Conservancy and Hydropower Press, Beijing City.
[2] Ministry of Water Resources of the People's Republic of China. (2019) SL 74-2019. China Water Conservancy and Hydropower Press, Beijing City.
[3] Tian Z., Huang W., He T., He P., Qiao Y. (2016) Detecting Text in Natural Image with Connectionist Text Proposal Network. Leibe B., Matas J., Sebe N., Welling M. (eds) Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science. 9912:56-72.
[4] Ma J., Shao W., Ye H., Wang L., Zheng Y. (2018) Arbitrary-Oriented Scene Text Detection via Rotation Proposals. IEEE Transactions on Multimedia. 20:3111-3122.
[5] Dai Y., Huang Z., Xu Y. Chen K., Guo J., Qiu W. (2018) Fused Text Segmentation Networks for Multi-oriented Scene Text Detection. in 24th International Conference on Pattern Recognition (ICPR), Beijing. pp: 3604-3609.
[6] Liu Y and Jin L. (2017) Deep Matching Prior Network: Toward Tighter Multi-oriented Text Detection. in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017 pp. 3454-3461.
[7] Shi B, Bai X and Yao C. (2017) An End-to-End Trainable Neural Network for Image-Based Sequence Recognition and Its Application to Scene Text Recognition. in IEEE Transactions on Pattern Analysis and Machine Intelligence.,39:2298-2304.